Dimensions of Digital Marketing and its Impact on Enhancing the Mental Image of the Organization

Bahr Kadhim Mohammad

mahemmed@qu.edu.iq

Saif Sabar Hatem

saif1992112@gmail.com

University of Al Qadisiyah

Article history: Received: 4/12/2024 Accepted: 15/12/2024 Available online: 25 /3 /2025 Corresponding Author : Saif Sabar Hatem

Abstract : When the response variable follows a normal distribution, the ordinary least squares (OLS) method is considered the most efficient and effective for estimating features. However, in real-world applications, we often encounter non-normal distributions, which is very common. In such cases, the OLS method becomes inefficient and does not provide accurate estimates. In this paper, we propose that the error distribution in a (23) factorial design is skewed and follows an Epsilon-Skew Normal (ESN) distribution. We employ the Modified Maximum Likelihood Estimator (MMLE) to estimate the parameters. We use simulation to validate the proposed method, along with real medical data to confirm the effectiveness of our approach. Additionally, we apply the Likelihood Ratio Test to the real data to assess the validity of the method.

Introduction: Experimental design has recently become a crucial instrument in the process of product creation. This design offers a methodical and empirical technique to evaluate the influence of regulated input parameters on the response variable being investigated. It facilitates the comprehension of the relationships between the input variables and the identification of the most favorable combination of variable levels to attain the optimal outcome, hence enhancing the quality or quantity of the end output. Usually, it is assumed that a mathematical model exhibits a linear correlation between the input factors and the response variable. Subsequently, the study is conducted utilizing the method referred to as study of Variance (ANOVA).

This technique is suitable when the replies adhere to a normal distribution, are not influenced by each other, and possess identical variances. However, in many real-life situations, the dependent variable may not follow a normal distribution, and the relationship between the dependent variable and the independent factors may not be linear. Assuming a normal distribution and independent observations, the equations obtained from Maximum Likelihood Estimation (MLE) are linear and can be solved. Conversely, if the data does not adhere to a normal distribution, the equations obtained from maximum likelihood estimation (MLE) become nonlinear, which presents greater difficulty in managing them. This study presupposes that the response variable adheres to an epsilon-skew normal distribution. Reviewing previous studies on this topic could provide other viewpoints.

In (Daniel, 1960), the identification of outliers in experimental design was discussed, noting that field experiments might include observations deviating from the general data pattern. Researchers suggested using alternative methods to the least squares (LS) method, such as M-estimation and the Huber function (Huber, 1981; Huber, 1973). In (Oyeyemi, 2004), a model was proposed using B techniques and Box-Cox transformation or Generalized Linear Models (GLM) to handle non-normal data, with a comparison of response plots through confidence intervals. In (Hutson, 2004), a regression model was presented that accounts for the error term following the Skew Epsilon Normal (ESN) distribution, including the estimation of skewness. In (Wang, Li, & Jiang, 2007), An approach for regression shrinkage that is strong and resilient against outliers was proposed. It utilizes LAD-Lasso and the Huber loss function for consistent variable selection. In (Kulkarni & Patil, 2010), 2ⁿ factorial experiments with a Poisson response variable were modified, comparing LOG, SQRT, ANOVA, and GLM methods, with the modified GLM showing the best performance. In (Patil & Kulkarni, 2011), GLM and LOG transformation were proposed for analyzing experiments with non-normal response variables and small sample sizes, compared with ANOVA based on confidence intervals There is no text provided. In an article (Jafari & Hashemi, 2011) the optimal D-design of the basic linear regression is presented with the assumption that errors follow a normal skewed distribution. On the other hand, a study (Kane and Mandal 2016) used the adaptive Lasso method to evaluate its effectiveness in complex designs using small samples. (Mohammed et al., 2017) aimed to analyze a response variable following the logarithm-epsilon-normal deviant (LESN) distribution, where the maximum probability estimation (MLE) method was used to estimate parameters and verify the effectiveness of the proposed methodology through simulation and factual data. On the other hand, the study

(Omondi et al.,2021) focused on the use of two-level experimental design and simulation to identify the main factors that affect product durability in different environments, with the aim of enhancing reaction efficiency and extending product life.

In this paper, we propose that the error distribution in a (2^3) factorial design is skewed and follows an Epsilon-Skew Normal (ESN) distribution. We employ the Modified Maximum Likelihood Estimator (MMLE) to estimate the parameters. We use simulation to validate the proposed method, along with real medical data to confirm the effectiveness of our approach. Additionally, we apply the Likelihood Ratio Test to the real data to assess the validity of the method.

The structure of this paper consists of several sections. The second section reviews the basic principles and benefits of two-level experiment designs. The third section provides a summary of the normal-skewed logarithm (ESN) distribution. Section 4 deals with the development of accurate and efficient estimators for Design 2^3 coefficients, which rely on modified maximum probability estimators (MMLE) to analyze factor experiments in which response variables follow the ESN distribution. Section 5 provides a summary of the simulation results and a detailed analysis of a particular sample data. Finally, Section VI provides a brief conclusion of the article.

2. Full factorial experiment design with two levels

Factor experiments, first introduced by Fischer (1935) and later developed by Yates (1937), are highly effective and widely usable designs in most research. These experiments study the effect of more than one factor on the trait under study, in order to obtain information about each factor of the experiment on the one hand, and to identify the interaction between these factors on the other. The design of 2^k factor experiments is widely used, where k refers to the number of factors involved in the study, while 2 represents the number of levels for each factor. These experiments have many benefits, especially in exploratory studies, because they are able to uncover important patterns that provide guidance for further experiments, as explained (Box et al., 1978), (Montgomery, 1984), and (Hinkelmann &Kempthorne1994).

The main factor effect is defined as the change in response due to a change in the level of the factor participating in the experiment. An interaction effect arises when the influence of one element is modified by the quantities of other constituents. Factorial effects encompass the cumulative consequences of both main effects and interaction effects. A full factorial design is able to accurately estimate the effects of all primary factors and interactions of higher order.

A design that consists of three components, A, B, and C, each with two levels, is commonly known as 2^3 -factorial design. The design yields eight treatment options, which can be visually depicted as a cube, as illustrated in Figure A. The symbols "+" and "-" are employed as orthogonal coding symbols to denote the high and low levels of the factors, respectively. The eight runs in this design are sequentially listed in Figure B within the "design matrix" as follows: (1), a, b, ab, c, ac, bc, and abc. It is crucial to acknowledge that these symbols also represent the total of all data obtained for any particular treatment combination.



There are three commonly used notations for the runs in 2^k designs. The first notation is the "+" and "-" notation, which is commonly referred to as geometric coding, orthogonal coding, or effects coding. The second method involves utilizing lowercase letter labels to designate the therapy combinations. The ultimate representation employs the digits 1 and 0 to signify elevated and diminished factor levels, correspondingly, rather than the symbols "+" and "-". The many notations for the 2^3 design are depicted below:

Run	A	B	С	Labels	A	B	С
1	_	_	_	(1)	0	0	0
2	+	_	_	a	1	0	0
3	_	+	_	Ь	0	1	0
4	+	+	_	ab	1	1	0
5	_	_	+	С	0	0	1
6	+	_	+	ac	1	0	1
7	_	+	+	bc	0	1	1
8	+	+	+	abc	1	1	1

Table 1: The design matrix 2^3

This type of factorial experimental design offers several methodological advantages: - The design is orthogonal and fulfills the requirement $X^T X = nl$ where:

- $X=(x_{ij})$: The design matrix consists of elements x_{ij} , where x_{ij} represents factor j at level i. The symbol I represents the identity matrix of size $q \times q$.

- The estimates of parameter β_j (j=1, 2, ..., q) obtained using orthogonal designs are unbiased and have reduced variances.

The orthogonal design matrix (X) enables the independent measurement of the major effects of components without any overlap between the effects.

- The potential to calculate interactions of varying orders, such as first order, second order, and so on. One alternative method to describe the idea of main effects and interaction effects in two-level designs, as outlined by (Montgomery ,2009), is by use a regression model. Consider a full factorial design to investigate the effects of three antiviral medications, namely A, B, and C, each having two levels. There are a total of 8 treatment combinations or levels in this case. The regression model commonly employed to analyze both the primary effects and interaction effects is:

 $Y_{i} = \beta_{0} + \beta_{1}x_{1} + \beta_{2}x_{2} + \beta_{3}x_{3} + \beta_{1}2x_{1}x_{2} + \beta_{13}x_{1}x_{3} + \beta_{23}x_{2}x_{3} + \beta_{123}x_{1}x_{2}x_{3} + \varepsilon_{i}$

The variable Y_i represents the response variable, while the parameters β represent the factors (A, B, C) as X₁, X₂, and X₃ accordingly. The word εi represents a random mistake. Frequently, when constructing a statistical model in desired factor trials, the objective is to discover a model that closely approximates the estimated values of the response variable to the actual values. The variables (X₁, X₂, X₃) are assigned a numerical code of 1 for the high level and -1 for the low level of their respective factors. The interaction between X₁ and X₂ is represented as X₁X₂, and the remaining interaction effects are specified in a similar manner.

3.Epsilon-skew-normal distribution

The Gaussian distribution, which was popularized by (Gauss ,1809) but known for almost a century before that, continues to be widely used because of its analytical elegance, simplicity, and the central limit effect identified by (De Moivre,1733). Nevertheless, there have consistently been uncertainties and hesitations regarding the unrestricted utilization of the normal distribution and reluctance towards approaches that rely on assumptions of normalcy. Several distributions that are similar to the normal distribution, partially comprise it, and possess some of its desirable characteristics, have been important in providing alternatives to the normal distribution.

Out of these options, the epsilon-skew-normal (ESN) distribution, represented as ESN (θ , σ , ε), is particularly notable. This distribution is defined by a singular peak at a certain value, with a probability mass of (1 - ε /)2 concentrated around this peak. The probability density function (PDF) for the ESN (0, 1, ε) distribution is defined as follows:

$$f_0(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} exp(-\frac{x^2}{2(1-\varepsilon)^2} ifx < 0 \\ \frac{1}{\sqrt{2\pi}} exp(-\frac{x^2}{2(1+\varepsilon)^2} ifx \ge 0 \end{cases}$$
(1)
The sumulative distribution function (CDE) is given by:

The cumulative distribution function (CDF) is given by:

$$F_{0}(x) = \begin{cases} (1-\varepsilon)\phi\left(\frac{x}{1-\varepsilon}\right)ifx < 0\\ -\varepsilon + (1+\varepsilon)\phi\left(\frac{x}{1+\varepsilon}\right)ifx \ge 0 \end{cases}$$
(2)

When the value of ε is equal to zero, this distribution aligns perfectly with the conventional normal distribution. For extremely large or small values of ε , it undergoes a transformation and becomes half-normal distributions that are already known. The ESN (θ , σ , ε) distribution has been confirmed to have a mean of θ , making it well-suited for

statistical applications that necessitate a versatile distribution that combines the characteristics of the normal distribution with skewness.

The epsilon-skew-normal distribution is an instance of extending distributions to incorporate extra characteristics beyond the conventional normal distribution. Azzalini (1985, 1986) and other researchers have demonstrated that this distribution provides substantial versatility in data modeling and analysis, making it a powerful statistical tool.

4. The 2³ Factorial Design with Epsilon-skew-normal

The three factor factorial experimental design model can be written as $\begin{aligned} y_{ijkl} &= \hat{\mu} + \hat{T}_i + \hat{\beta}_j + \hat{\gamma}_k + (\hat{T}\hat{\beta})_{ij} + (\hat{T}\hat{\gamma})_{ik} + (\hat{\beta}\hat{\gamma})_{ij} + (\hat{\beta}\hat{\gamma})_{ik} + (\hat{T}\hat{\beta}\hat{\gamma})_{ijk} + e_{ijkl} \\ i &= 1, \dots, afactorA, j = 1, \dots, bfactorB, k = 1, \dots, cfactorC, l = 1, \dots, nreplicate \end{aligned}$ (3) In model (3), we assume that this model is fixed effect and then, $\sum_{i=1}^{a} T_i = 0, \sum_{i=1}^{b} \beta_i = 0, \sum_{k=1}^{c} \gamma_k = 0, \sum_i (T\beta)_{ii} = 0 \forall j$ $\sum_{i} (T\beta)_{ij} = 0 \forall i, \sum_{i} (T\gamma)_{ik} = 0, \forall k \sum_{k} (T\gamma)_{ik} = 0 \forall i, \sum_{j} (\beta\gamma)_{jk} = 0 \forall k$ $\sum_{k} (\beta \gamma)_{jk} = 0 \forall j, \sum_{i} (T\beta \gamma)_{ijk} = 0, \forall jk \sum_{j} (T\beta \gamma)_{ijk} = 0 \forall ik, \sum_{k} (T\beta \gamma)_{ijk} = 0 \forall ijk$ Suppose that the error from model (3) follow the

$$f(x;\varepsilon) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-x^2}{2(1+\varepsilon)^2}\right]; x < 0\\ \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-x^2}{2(1-\varepsilon)^2}\right]; x \ge 0 \end{cases}$$
(4)

So, if we use the error distribution as ESN $(0, \sigma^2, \epsilon)$, then the pdf we will be as follows,

$$f(e;\varepsilon) = \begin{cases} \frac{1}{\sqrt{2\pi}} exp\left[\frac{-e^2}{2(1+\varepsilon)^2}\right]; x < 0\\ \frac{1}{\sqrt{2\pi}} exp\left[\frac{-e^2}{2(1-\varepsilon)^2}\right]; x \ge 0 \end{cases}$$
(5)
Where

Where

$$e_{ijkl} = y_{ijkl} - \mu - T_i - \beta_j - \gamma_k - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}$$

$$f(x; \sigma^2 \varepsilon) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[\frac{-(y_{ijkl} - \mu - T_i - \beta_j - \dots - (T\beta\gamma)_{ijk})^2}{2(1+\varepsilon)^2 \sigma^2}\right]; e > 0 \\ \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[\frac{-(y_{ijkl} - \mu - T_i - \beta_j - \dots - (T\beta\gamma)_{ijk})^2}{2(1-\varepsilon)^2 \sigma^2}\right]; e \le 0 \end{cases}$$
(6)

The following section discuss the parameter estimation by using the maximum likelihood (MLE) method based on the distribution (6)

The ln likelihood function (ln L) for the probability density function (6) with e_{ijkl} ESN (0, σ , ε) is defined as follows:

$$L = \prod_{i=1}^{n} f(e_{i}; \sigma^{2}; \varepsilon) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^{2}}} exp\left[\frac{-\sum_{i=1}^{a} \sum_{i=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} Z_{ijkl}^{k}}{2(1+\varepsilon)^{2}}\right]; Z_{ijkl} > 0\\ \frac{1}{\sqrt{2\pi\sigma^{2}}} exp\left[\frac{-\sum_{i=1}^{a} \sum_{k=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} Z_{ijkl}^{k}}{2(1-\varepsilon)^{2}\sigma^{2}}\right]; Z_{ijkl} \le 0 \end{cases}$$
(7)

Where

$$Z_{ijkl} = \frac{y_{ijkl} - \mu - T_i - \beta_j - \gamma_k - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma}$$

In real data analysis, the non-normality is more frequent phenomena, so using the ordinary last square methods gives, likelihood estimation for this type of data. To find an explicit solution for the parameters in (7), we will express (7) in terms of the order statistics as follows:

$$L = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[\frac{-\sum_{i=1}^{a} \sum_{l=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} g_1(Z_{ijk(L)}^2)}{2(1+\varepsilon)^2}\right]; Z_{ijk(L)} > 0\\ \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[\frac{-\sum_{i=1}^{a} \sum_{l=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} g_2(Z_{ijk(L)}^2)}{2(1-\varepsilon)^2\sigma^2}\right]; Z_{ijk(L)} \ge 0 \end{cases}$$
Where:
$$q(Z_{ijk(L)}) = Z^2;$$
(8)

$$g(Z_{ijk(L)}) = Z^{2};$$

$$Z_{ijk(L)} = \frac{y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\tau}$$

here, $y_{ijk(L)}$ is the order statistics of the sample observation,

 $y_{iik(L)}$; l = 1, 2, ..., n

Now, the ln likelihood equation will be as follows:

$$lnL = \begin{cases} nln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)exp\left[\frac{-\sum_{i=1}^{a}\sum_{l=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{a}g_{1}(Z_{ijk(L)})}{2(1+\varepsilon)^2}\right]; Z_{ijk(L)} > 0\\ nln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)exp\left[\frac{-\sum_{i=1}^{a}\sum_{l=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}g_{2}(Z_{ijk(L)})}{2(1-\varepsilon)^2\sigma^2}\right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(9)

The following are the partial derivatives of ln likelihood function of (9) with respect to the model parameters: $\left(\sum_{i}\sum_{j}\sum_{k}\sum_{l}\left[-y_{ijk(L)}-\mu-T_{i}-\beta_{j}-\gamma_{k}-(T\beta)_{ij}-(T\gamma)_{ik}-(\beta\gamma)_{jk}-(T\beta\gamma)_{ijk}\right], \tau < 0$

$$\frac{\partial lnL}{\partial \mu} = \begin{cases} \frac{1}{2(1+\varepsilon)^2 \sigma} \left[\frac{1}{\sigma} \frac{\sigma}{\sigma} \right]; Z_{ijk(L)} < 0\\ \frac{\sum_i \sum_j \sum_k \sum_L}{2(1-\varepsilon)^2 \sigma} \left[\frac{-y_{ijk(L)} - \mu - T_i - \beta_j - \gamma_k - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(10)

$$\frac{\partial lnL}{\partial T_{i}} = \begin{cases} \frac{\sum_{j} \sum_{k} \sum_{l}}{2(1+\varepsilon)^{2}\sigma} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} < 0 \\ \frac{\sum_{j} \sum_{k} \sum_{L}}{2(1-\varepsilon)^{2}\sigma} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(11)

$$\frac{\partial \ln L}{\partial \beta_{j}} = \begin{cases} \frac{\sum_{j} \sum_{k} \sum_{l}}{2(1+\varepsilon)^{2} \sigma} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} < 0 \\ \frac{\sum_{j} \sum_{k} \sum_{l}}{2(1-\varepsilon)^{2} \sigma} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(12)

$$\frac{\partial lnL}{\partial \gamma_k} = \begin{cases} \frac{\sum_j \sum_k \sum_l}{2(1+\varepsilon)^2 \sigma} \left[\frac{-y_{ijk(L)} - \mu - T_i - \beta_j - \gamma_k - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} < 0 \\ \frac{\sum_j \sum_k \sum_l}{2(1-\varepsilon)^2 \sigma} \left[\frac{-y_{ijk(L)} - \mu - T_i - \beta_j - \gamma_k - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(13)

$$\frac{\partial \ln L}{\partial (T\beta)_{ij}} = \begin{cases} \frac{\sum_k \sum_l}{2(1+\varepsilon)^2 \sigma} \left[\frac{y_{ijk(L)} - \mu - T_i - \beta_j - \gamma_k - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} < 0 \\ \frac{\sum_k \sum_l}{2(1-\varepsilon)^2 \sigma} \left[\frac{-y_{ijk(L)} - \mu - T_i - \beta_j - \gamma_k - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{ijk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(14)

$$\frac{\partial \ln L}{\partial (\beta \gamma)_{jk}} = \begin{cases} \frac{\sum_{i} \sum_{l} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} < 0\\ \frac{\sum_{i} \sum_{l} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(15)

$$\frac{\partial lnL}{\partial (T\gamma)_{ik}} = \begin{cases} \frac{\sum_{j} \sum_{l}}{2(1+\varepsilon)^{2}\sigma} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} < 0\\ \frac{\sum_{j} \sum_{l}}{2(1-\varepsilon)^{2}\sigma} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(16)

$$\frac{\partial lnL}{\partial (T\beta\gamma)_{ijk}} = \begin{cases} \frac{\sum_{l} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} < 0\\ \frac{\sum_{l} \left[\frac{-y_{ijk(L)} - \mu - T_{i} - \beta_{j} - \gamma_{k} - (T\beta)_{ij} - (T\gamma)_{ik} - (\beta\gamma)_{jk} - (T\beta\gamma)_{ijk}}{\sigma} \right]; Z_{ijk(L)} \ge 0 \end{cases}$$
(17)

Based on Mutholkar & Hutson (2000); Al-Mousawi et al. (2012), the MLE of $\theta = (\mu, \tau, \beta, \gamma, \tau \beta, \tau \gamma, \beta\gamma, \tau \beta\gamma)$ can be found by considering that the order statistics $y_{ijk(1)}, y_{ijk(2)}, ..., y_{ijk(n)}$ of the response variable y_{ijkL} that following epsilon skew normal ($\theta, \varepsilon, \sigma^2$) distribution. Now, let $y_{ijk(0)} = -\infty$ and $y_{ijk(n+1)} = \infty$. Define $t = t(y_{ijk(1)}, ..., y_{ijk(n)}, \theta)$, where $y_{ijk(t)} < \theta < y_{ijk(t+1)}$, with t is a random integer Let us define the Epsilon skew normal in (5) as:

$$f(y_{ijkL}) = \frac{1}{\sqrt{2\pi\sigma^2}} \left\{ exp\left[\frac{-(y_{ijkL}-\theta)^2}{2(1-\varepsilon)^2\sigma^2}\right] I(y_{ijkL} \ge \theta) + \left[\frac{-(y_{ijkL}-\theta)^2}{2(1-\varepsilon)^2\sigma^2}\right] I(y_{ijkL} < \theta) \right\}$$
(18)
If $0 \le t \le n$, then the ln-likelihood function of (19) can be defined by:

$$L(\theta, \sigma^{2}, \varepsilon) = \begin{cases} \frac{-n}{2} \log(2\pi\sigma^{2}) - \left(\frac{1}{4\sigma^{2}}\right) \exp\left\{\sum_{l=1}^{n} \left(y_{ijkl} - \theta\right)^{2}\right\}; ift = 0, n\\ \frac{-n}{2} \log(2\pi\sigma^{2}) - \left(\frac{1}{2\sigma^{2}}\right) \left\{\sum_{l=1}^{t} \frac{\left(y_{ijkl} - \theta\right)^{2}}{(1+\varepsilon)^{2}} + \sum_{l=t+1}^{t} \frac{\left(y_{ijkl} - \theta\right)^{2}}{(1+\varepsilon)^{2}}\right\}; if1 \le t < n \end{cases}$$
(19)

Let

$$D(\theta,\varepsilon) = \frac{1}{(1+\varepsilon)^2} \sum_{l=1}^{t} \left[y_{ijkl} - \theta \right]^2 + \frac{1}{(1+\varepsilon)^2} \sum_{l=t+1}^{t} \left[y_{ijkl} - \theta \right]^2$$

See Mudholkar & Hutson (2000) for more information and the prove of likelihood function in (19). It is north to nothing that for t=0 and t=n, the ln-likelihood function comes straightly from the half-normal distributions. Now,

based on the likelihood function in equation (19), and using the order statistics to find the maximum likelihood estimators for $(\theta, \sigma^{2}, \epsilon)$.

$$\varepsilon = \frac{\left[\sum_{l=1}^{t} (y_{ijkl} - \theta)\right]^{\frac{1}{3}} - \left[\sum_{l=t+1}^{n} (y_{ijkl} - \theta)\right]^{\frac{1}{3}}}{\left[\sum_{l=1}^{t} (y_{ijkl} - \theta)\right]^{\frac{1}{3}} + \left[\sum_{l=t+1}^{n} (y_{ijkl} - \theta)\right]^{\frac{1}{3}}}$$

$$\sigma^{2} = \frac{1}{4} \left\{ \left[\sum_{l=1}^{t} (y_{ijkl} - \theta) \right]^{\frac{1}{3}} + \left[\sum_{l=t+1}^{n} (y_{ijkl} - \theta) \right]^{\frac{1}{3}} \right\}$$
5. SIMULATION STUDY

5. SIMULATION STUDY

In order to illustrate the Modified maximum likelihood estimation (MMLE) of the parameters $(\mu, \tau_i, \beta_j, \gamma_k, \sigma, \varepsilon)$ of the epsilon-skew-normal distribution, we conduct a simulation study using the R software. The goal of this simulation is to study the reliability factorial experiment (2³), and to estimate the parameters when the response variable follows the ESN distribution. At first we assume that *Y ESN*, and repeat each experiment (IT=1000) for all the simulation experiments. Then we identify some of the default values for the parameters that we need in this distribution, as well as changing the number of repetitions for each experiment (r=1, 2, 4, 10) and the default values for the skew parameter

($\varepsilon = 0.2, 0.5, 0.8$). When giving default values for the factors (A, B, C) we represent the parameters ($\beta_1, \beta_2, \beta_3$). After estimating the parameters, we find the design matrix and calculate the confidence intervals (upper and lower) in order to get the parameters estimation.

Table (1) summarize the MLE for the parameters β_0 , β_1 , β_2 , β_3 , $\varepsilon \wedge \sigma$, and we can see that the parameter estimates are close to the default values in the designed algorithm for this paper. We note that there is a simple difference, and sometimes an increase in variance (σ^2) when we increase the number of repetitions (r). Figures () and () graphically illustrate the factors A, B, C and support the results obtained in terms of stability and convergence to the simulation experiments

Table (2) Estimation of para	neters for model design 2	³ Factors when n=50
------------------------------	---------------------------	--------------------------------

REP.		E	${ otin} eta_0 = \mu$	$eta_1= au$	$\beta_2=eta$	$m{eta}_3=\gamma$	${oldsymbol{eta}}_{12}= aueta$	$\hat{m{eta}}_{13}= au\gamma$	$\hat{\beta}_{23} = \beta \gamma$	$\hat{m{\beta}}_{123} = aum{eta}\gamma$	ð	ê
	•	Lower	3.817	0.527	1.864	-	2.362	-	0.607	0.901	4.893	0.818
	0.	Upper	4.595	1.314	1.735	2.158	1.722	2.470	1.120	1.014	8.189	1.105
	4	Estimat	4.206	0.921	1.799	1.005	2.042	0.829	0.864	0.957	6.541	0.961
	0	Lower	4.167	-	0.429	0.525	-	0.128	-	0.222	5.996	0.764
1	U. 5	Upper	8.334	3.027	2.014	2.003	3.331	4.433	2.195	3.187	10.17	1.768
	3	Estimat	6.250	1.409	1.222	1.264	1.519	2.280	0.922	1.705	8.084	1.266
	•	Lower	1.788	-	-	0.126	0.042	0.159	0.151	-	7.119	1.242
	U. o	Upper	6.356	2.995	2.986	1.757	4.577	2.805	5.411	4.854	13.97	2.077
	ð	Estimat	4.072	1.072	1.218	0.942	2.309	1.482	2.781	2.355	10.54	1.659
	Δ	Lower	4.121	0.763	-	0.562	-	-	0.834	0.844	3.932	0.422
	0. 2	Upper	7.971	3.278	2.001	2.104	4.498	3.331	1.124	1.001	6.173	0.988
	4	Estimat	6.046	2.021	0.343	1.333	1.996	1.193	0.979	0.923	5.052	0.705
	Δ	Lower	3.487	-	-	0.102	0.285	0.116	-	-	4.965	0.662
2	U. 5	Upper	6.555	2.323	3.252	1.133	4.730	2.185	1.987	2.241	9.116	1.443
	5	Estimat	5.021	1.010	1.551	0.618	2.508	1.150	0.928	1.031	7.041	1.052
	Δ	Lower	4.580	-	0.017	-	0.193	0.045	-	-	5.093	1.109
	U. 8	Upper	5.621	4.682	4.115	2.425	2.661	1.263	3.556	1.989	10.22	1.899
	0	Estimat	5.100	2.244	2.066	0.892	1.427	0.654	1.513	0.852	7.657	1.504
4	0.	Lower	2.174	0.426	0.122	0.884	0.024	0.605	1.052	0.158	3.126	0.422
4	2	Upper	6.336	3.551	5.653	6.336	1.554	4.338	5.987	3.748	5.715	0.988

		Estimat	4.255	1.988	2.888	3.610	0.789	2.472	3.520	1.953	4.420	0.705
	0	Lower	3.051	0.169	0.073	0.224	0.149	0.093	0.448	0.101	4.102	0.662
	U. 5	Upper	9.188	4.822	2.585	3.998	2.889	1.220	2.098	3.400	8.748	1.443
	3	Estimat	6.119	2.496	1.329	2.111	1.519	0.657	1.273	1.751	6.425	1.052
	0	Lower	3.466	0.110	0.612	-	-	0.129	-	0.172	4.211	0.332
	0. 8	Upper	8.778	3.748	5.366	1.118	3.707	2.899	3.172	3.889	8.002	1.899
		Estimat	6.122	1.929	2.989	0.491	1.761	1.514	1.262	2.031	6.106	1.116
	0	Lower	3.582	-	0.134	0.190	0.784	-	-	-	1.878	0.326
	U. 2	Upper	6.988	3.142	3.667	3.442	4.338	2.885	1.288	5.228	3.131	0.762
	2	Estimat	5.285	1.196	1.901	1.816	2.561	1.077	0.554	2.349	2.504	0.544
	0	Lower	4.552	0.362	0.537	0.100	0.292	0.238	-	0.990	2.911	0.445
10	U. 5	Upper	10.14	5.556	3.919	3.692	4.115	5.512	3.511	2.669	7.664	1.288
	3	Estimat	7.348	2.959	2.228	1.896	2.204	2.875	1.627	1.830	5.287	0.867
	0	Lower	4.434	0.741	-	-	0.212	-	-	-	2.200	0.311
	U. 8	Upper	8.557	6.335	3.788	4.822	3.338	4.889	4.112	2.178	6.655	1.544
	0	Estimat	6.496	3.538	1.798	2.140	1.775	2.049	1.876	1.003	4.427	0.927



Figure (1) Trace plot, replication =2, n=50



Figure (2) Trace plot, replication = 4, n=50



Figure (3) Trace plot, replication = 10, n=50 and

The results presented in Table (2), supported by Figure (1),(2)and(3) show the trace plot values of the factors (A, B and C) in terms of their stability, approximation, and sometimes fluctuations during the simulation experiments conducted with changes in the skew parameter (ϵ). The levels of the factors (high and low) are also varied in the replications. And with a sample size of n=50.

In the first case, when (r = 1) and the skew parameter values are ($\epsilon = 0.2, 0.5, 0.8$), we observe that the trace plot values for the factors (A,B and C) are close to the default values of the simulations, with the best case occurring at ($\epsilon = 0.2$), where the results show values concentrated near the averages of the factors.

In the second case, with (r = 2) and skew parameter values of $(\varepsilon = 0.2, 0.5, 0.8)$, we find that the trace plot values for the factors (A, B and C) remain close to the default values of the simulations. The best case again occurs at $(\varepsilon = 0.2)$, where the figure shows values centered around the averages of the factors.

In the third case, when (r = 4) and skew parameter values of ($\epsilon = 0.2, 0.5, 0.8$), we notice that the trace plot values for the factors (A,B and C) are closer and more stable compared to the previous cases relative to the default simulation values, with the best case occurring at

 $(\epsilon = 0.8).$

In the fourth case, when (r = 10) and skew parameter values of ($\epsilon = 0.2, 0.5, 0.8$), we find that the trace plot values for the factors (A,B and C) are closer and more stable than in previous cases concerning the default simulation values. The optimal case occurs at ($\epsilon = 0.8$), where the figure illustrates values concentrated near the averages of the factors. 6. Real Data

Viral hepatitis is an infectious disease caused by viruses that damage liver cells. The resulting damage may be temporary or permanent. Viral hepatitis is characterized by the presence of inflammatory cells within the liver tissue. Viral hepatitis causes jaundice in children. There are five types of hepatitis (A, B, C, D, E). There are also other types that are not classified or have an unclear association with the disease, such as hepatitis G. The hepatitis virus class Viral hepatitis causes acute or chronic liver infection and inflammation, resulting in a serious global public health problem. Hepatitis B and C viruses are two major causes of severe illness and death. The global burden of disease caused by acute hepatitis B and C, liver cancer, and cirrhosis is high (about 27% of all deaths) and is expected to occupy a high position in the list of causes of death in the next two decades. It is estimated that 57% of cases of cirrhosis and 78% of cases of primary liver cancer are caused by infection with hepatitis B or C viruses. According to the World Health Assembly (2010), the transmission of hepatitis virus can be prevented thanks to the following effective strategies in the field of public health (vaccination against hepatitis virus infection, implementation of safe health care, food, and water supply).

Based on the factors causing the disease, a factorial experiment was conducted (2^3) to identify the factors affecting the incidence of viral hepatitis disease on a sample consisting of patient records collected from Al-Rifai Hospital/Dhi Qar, which represents the response variable (y) with interactions of these factors at two levels for each factor, a high level (+1) and a low level (-1). There are 4 common factors produced by three main factors, and there are 3 two-way interactions, and one three-way interaction. Tables (3) and (4) show the factors and levels for each factor, describe the main factors, and interacting factors.

Factors	Factors levels						
ractors	High level: +1	Low level: -1					
A= Vaccination	Full and early vaccination of children and coverage of health workers.	Inadequate vaccination					
B= Healthcare	Ensuring safe blood supplies and safe injection practices	Lack of preventive measures or poor control of transmission in health facilities.					
C=Food and water supply	Ensuring safe food and water according to health standards.	Lack of effective safeguards or measures for food and water safety.					

Table (3) The Factors and Levels for each factor

Factors		Descriptive of the factors (main and interactions)		
	А	Vaccination		
Main Factors	В	Healthcare		
	С	Food and water supply		
	AB	The interaction between vaccination and health care		
Two-factors interactions	AC	The interaction between vaccination and food and water supply		
	BC	The interaction between health care and food and water supply		
Three-factors interactions	ABC	The interaction between vaccination, health care and food and water supply		

Table 5. the results of response variable () for each level with interaction

Factor Effect	μ	Α	В	С	AB	AC	BC	ABC	Response	
									Replication 1	Replication 2
1	+1	-1	-1	-1	+1	+1	+1	-1	1.654	1.195
2	+1	+1	-1	-1	-1	-1	+1	+1	1.693	1.285

3	+1	-1	+1	-1	-1	+1	-1	+1	1.272	1.199
4	+1	+1	+1	-1	+1	-1	-1	-1	1.716	1.281
5	+1	-1	-1	+1	+1	-1	-1	+1	1.336	1.178
6	+1	+1	-1	+1	-1	+1	-1	-1	1.390	1.347
7	+1	-1	+1	+1	+1	-1	+1	-1	1.445	1.462
8	+1	+1	+1	+1	+1	+1	+1	+1	1.582	1.354

Testing the data distribution

There are many statistical tests used to determine the distribution of the data of the phenomenon studied including the Kolmogorov-Smirnov test, Shapiro -Wilk test and the cumulative probability of standard distribution. As well as the Boxplots chart, this is often used to illustrate the distribution of factors data by levels. Therefore, the last method was chosen to test these data, as it shows the appropriateness of the data and its distribution by factor levels.



Figure (4) Boxplot charts for viral hepatitis according to factors (A, B, C) with two levels for each factor and interaction

Through the forms in Figure (3.19), we observe that many levels of factors are asymmetric. The graph shows that the vaccination factor (A) and the healthcare application factor (B) are skewed to the right. Signs of this include the high level (+1) and the low level (-1). We note that the right tail is shorter, while the left tail is longer. Additionally, the last

factor, as well as the two-way interactions and the three-way interaction, also exhibit asymmetry. These factors, along with their levels, affect the distribution of the response variable (hepatitis), making it asymmetric and tending towards the (ESN) distribution.



Histogram of Skew Normal Distribution for response variabl

Figure (5): Histogram of Skew normal distribution for response variables.

Figure (5) shows that the distribution does not appear to be symmetric, indicating a possible skew in the data distribution. Specifically, the distribution seems to be tilted to the right, with the right tail being longer than the left tail. We also observe that the data follows a skewed normal distribution, specifically following an epsilon-skew normal distribution (ESN).

Table (6) MLE (Coefficient for the viral hepati	itis (main factors and interactions
-----------------	----------------------------------	-------------------------------------

Parameter	MLE Coeff.	Lower CI	Upper CI	Likelihood Ratio	P value
β_0	2.772	2.431	3.113	11.364	0.3342
$\hat{\beta}_1(A)$	0.445	0.395	0.495	7.7743	<mark>0.0037</mark>
$\hat{\beta}_2(B)$	0.375	0.337	0.413	6.9822	<mark>0.0129</mark>
$\hat{\beta}_3(\mathcal{C})$	0.224	0.196	0.252	4.4682	0.3251
$\hat{\beta}_4(AB)$	0.482	0.428	0.536	5.5478	0.0148
$\hat{\beta}_5(AC)$	0.442	0.388	0.496	3.6594	0.1224
$\hat{\beta}_6(BC)$	0.371	0.333	0.409	6.1131	<mark>0.0155</mark>
$\hat{\beta}_7(ABC)$	0.482	0.434	0.530	5.9877	0.0171
ð	0.1533	0.138	0.169	-	-
Ê	0.0874	0.077	0.098	-	-

The table (6) summarizes the maximum likelihood estimates (MLE) of the parameters, along with the confidence interval bounds, likelihood ratios, and p-values.

1. Estimated Coefficients (β)

The estimated coefficients indicate the effects of various factors and their interactions on the dependent variable. Below is a summary of each coefficient's effect:

• β_0 (Intercept): The estimated value is 2.772, representing the baseline mean of the model in the absence of other factor effects.

• β_1 (A): With an estimated value of 0.445, this coefficient represents the individual effect of factor A, which stands for Vaccination. This effect is statistically significant with a p-value of 0.0037, indicating a notable impact of this factor.

• Remaining Coefficients: These represent the individual effects of factors B and C and the two- and three-way interactions between factors such as AB, AC, BC, and ABC, reflecting the complexity of factor interactions.

2. Confidence Interval Bounds

Confidence intervals provide an assessment of the accuracy of each estimate, helping to understand the stability of each estimate:

• β_0 : The confidence interval ranges from 2.431 to 3.113, suggesting a 95% confidence that the true value lies within this range.

• β_1 (A): With a confidence interval ranging from 0.395 to 0.495, this reinforces the importance of factor A and suggests a stable estimate, given its relatively narrow bounds.

3. Likelihood Ratio Test and P-values

The likelihood ratio test indicates the significance of factors or interactions in explaining the dependent variable. Here is a summary of factor significance based on p-values:

• β_1 (A): With a likelihood ratio of 7.7743 and a p-value of 0.0037, factor A (Vaccination) has a significant effect on the dependent variable at the 5% significance level.

• β_2 (B): With a likelihood ratio of 6.9822 and a p-value of 0.0129, factor B (Healthcare) also has a statistically significant effect on the dependent variable at the 5% level.

• Interactions (AB, BC, and ABC): Coefficients β_4 (AB), β_6 (BC), and β_7 (ABC) all show statistical significance at the 5% level, with p-values below 0.05, indicating the importance of interactions between these factors.

• Non-significant Coefficients: β_0 , β_3 (C), and β_5 (AC) have p-values greater than 0.05, suggesting they are not statistically significant at the 5% level, and thus may not have a substantial impact on the dependent variable.

The results indicate that factors A and B and the interactions AB, BC, and ABC have significant effects on the dependent variable, highlighting the importance of these factors in explaining the model's outcomes.

7. Conclusions

Based on the results of simulation experiments and real data, we conclude the following:

1. The Modified Maximum Likelihood Estimation (MMLE) method is a comprehensive and effective approach for accurately estimating the parameters of factorial experiment models. It demonstrates strong performance in capturing the effects of main factors and their interactions, particularly when the response variable follows a skewed normal distribution.

2. The Likelihood Ratio Test (LRT) proves to be effective in identifying important factors when the response variable follows a non-normal distribution.

3.In analyzing real data, the MMLE method efficiently identified key factors in the viral hepatitis dataset. These factors include vaccination (Factor A), healthcare (Factor B), the binary interactions between vaccination and healthcare (AB), and between vaccination and food and water supply (AC), as well as the triple interaction among vaccination, healthcare, and food and water supply (ABC).

8. Recommendations

Based on the results and conclusions, some recommendations for this study can be summarized

1. Using the proposed method when the response variable follows a skewed normal distribution.

2. Using tests of non-normal distributions with factorial experiments. Recommendations Conclusions and Recommendations

3. The researcher recommends employing Bayesian methods in analyzing factorial experiments that do not follow a normal distribution.

4. Employing the proposed method with industrial, agricultural, and medical datasets

References

• Azzalini, A. (1985). A class of distributions which includes the normal ones. Scandinavian Journal of Statistics, 12(2), 171–178.

• Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. Statistica, 46(2), 199–208.

• Box, G. E. P., Hunter, W. G., & Hunter, J. S. (1978). Statistics for experimenters (2nd ed.). Wiley and Sons.

• Daniel, C. (1960). Locating outliers in factorial experiments. Technometrics, 2(2), 149–156. https://doi.org/10.2307/1266562

- De Moivre, A. (1733). Approximatio ad summam terminorum binomii $(a + b)^n$ in seriem expansi.
- Gauss, C. F. (1809). Theoria Motus Corporum Coelestium. Perthes and Besser.
- Hinkelmann, K., & Kempthorne, O. (1994). Design and analysis of experiments (Vol. 1). Wiley.

• Huber, P. J. (1973). Robust regression: Asymptotics, conjectures, and Monte Carlo. The Annals of Statistics, 1(5), 799–821. <u>https://doi.org/10.1214/aos/1176342503</u>

• Huber, P. J. (1981). Robust statistics (1st ed.). Wiley.

• Hutson, A. D. (2004). Utilizing the flexibility of the epsilon-skew-normal distribution for common regression problems. Journal of Applied Statistics, 31(6), 673–683. <u>https://doi.org/10.1080/0266476042000188765</u>

• Jafari, H., & Hashemi, R. (2011). Optimal designs in a simple linear regression with skew-normal distribution for error term. Journal of Applied Mathematics, 1(2), 65–68.

• Kane, A., & Mandal, A. (2016). A new analysis strategy for designs with complex aliasing. Submitted to the American Statistician.

• Kulkarni, H. V., & Patil, V. V. (2010). A comparative study of analyses of 2n2ⁿ factorial experiments with a Poisson distributed response variable. Communications in Statistics - Simulation and Computation, 39(8), 1530–1547. https://doi.org/10.1080/03610911003608091

• Mohammed, B. K. (2018). Robust lasso variable selection for factorial experiments analysis with application. International Journal of Statistics and Applications, 8(2), 79–87.

• Mohammed, B. K., Roman, M., Odah, M. H., & Bager, A. S. M. (2017). Testing reliability: Factorial design with data from a log-epsilon-skew-normal distribution. Economic Computation & Economic Cybernetics Studies & Research, 51(3), 23–40.

- Montgomery, D. C. (1984). Design and analysis of experiments. Wiley.
- Montgomery, D. C. (2009). Design and analysis of experiments (9th ed.). Wiley.
- Mudholkar, G. S., & Hutson, A. D. (2000). The epsilon-skew-normal distribution for analyzing near-normal data. Journal of Statistical Planning and Inference, 83(2), 291–309. <u>https://doi.org/10.1016/S0378-3758(99)00174-6</u>

• Omondi, B., et al. (2021). 2k2^k factorial experiments in reliability analysis for Weibull and log-normal distributions. International Journal of Engineering and Innovative Research, 3(2), 115–120.

• Oyeyemi, G. M. (2004). Treatment of non-normal responses from designed experiments. Journal of the Nigerian Statistical Association, 17(1), 8–19.

• Patil, S. C., & Kulkarni, H. V. (2011). Analysis of 2n2ⁿ factorial experiments with exponentially distributed response variable. Applied Mathematical Sciences, 5(10), 459–476.

• Wang, H., Li, G., & Jiang, G. (2007). Robust regression shrinkage and consistent variable selection through the LAD-Lasso. Journal of Business & Economic Statistics, 25(3), 347–355. https://doi.org/10.1198/07350010600000251

• Wu, C., & Hamada, M. (2011). Experiments: Planning, analysis, and optimization. Wiley.