World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Bayesian three-stage a-optimal design for generalized linear models

Ali H [1,*], Nwaosu S. C [2], Lasisi K. E [2] and Abdulkadir A [2]

[1] *Department of Mathematics, University of Jos, P. M. B. 2048, Jos, Plateau, Nigeria.*
[2] *Department of Mathematical Sciences, Abubakar Tafawa Balewa University, P. M. B. 0248, Bauchi, Nigeria.*

World Journal of Advanced Research and Reviews, 2023, 19(01), 1150–1165

Publication history: Received on 02 June 2023; revised on 21 July 2023; accepted on 24 July 2023

Article DOI: https://doi.org/10.30574/wjarr.2023.19.1.1380

Abstract

Bayesian sequential designs are increasingly receiving attention in recent years, especially in clinical trials and biomedical research. Bayesian sequential design process can utilize the available prior information of the unknown parameters so that a better design can be achieved. In this paper, a hybrid computational method, which consists of the combination of a rough global optima search and a more precise local optima search, is proposed to efficiently search for the Bayesian A-optimal designs for multi-variable generalized linear models. Specifically, Poisson regression models and logistic regression models are investigated. Designs are examined for a range of prior distributions and the equivalence theorem is used to verify the design optimality. Design efficiency for various models were examined. Furthermore, the idea of the Bayesian sequential design is introduced and the Bayesian three-stage A-optimal design approach is introduced for generalized linear models. With the incorporation of the first stage data information into the second stage, the second stage data information into the third stage, the three-stage design procedure improved the design efficiency and produce more accurate and robust designs. The Bayesian three-stage A-optimal designs for Poisson and logistic regression models are evaluated based on simulation studies. The Bayesian three-stage A-optimal design is superior to the two-stage A-optimal design approach in terms of design optimality and efficiency criteria.

Keywords: A-Optimality; A-Efficiency; Logistic; Poisson; Fisher Information Matrix

1 Introduction

Generalized Linear Model is a flexible generalization of the ordinary linear model, and allows the response variable to have an error distribution other than normal. Thus, it can be used for the regression of discrete responses, including binary and count data which is what we focus on. It is perhaps not too much of an exaggeration to say that the rich development of normal-theory linear models has been a major force behind much of statistics and scientific discovery in the last century. When data was not quite normal then transformations were used to make the data normally distributed with constant variance. For example, Box and Cox (1964) is a widely used method that has garnered more than 100,309 citations according to the Web of Science. The need to make the data more like the assumptions of the linear model decreased as modern computing power increased. Generalized linear models (GLMs) were developed by Nelder and Wedderburn (1972) as a generalization of the linear model to a larger variety of response distributions and non-linear relationships between the mean and the variance. GLMs are usually restricted to responses from the natural exponential family, but their adoption allowed practitioners to better explore new kinds of data. Responses composed of success and failures, counts of pits on a polished surface, and even time between failures all can be represented using this technique. Software to fit GLMs is ubiquitous. Applications in the literature are numerous. The complexity of GLMs has made experimental design a difficult task. Bayesian *D*-optimality Chaloner and Larntz (1989) is one approach for designing experiments for GLMs. This approach is presently implemented in commercial software. Philosophically, the approach is a hybrid of frequentist and Bayesian approaches, somewhat to its own detriment. Issues are identified when the Bayesian prior covers models with no active effects. The asymptotic criterion is related to drastically different small-

* Corresponding author: Ali H https://orcid.org/0000-0001-5517-8226

sample performance. Ways to determine if a design is suffering from these maladies are suggested. GLMs rely strongly on independence of observations, but in many industrial experiments this is simply not true. Generalized linear mixed models (GLMMs) are a generalization of generalized linear models in which correlated random effects are introduced into the GLM to induce correlation among the observations. Some need exists for general methods to design exact experiments for GLMMs. Sung, *et al* (2019) developed sequential designs for conditionally natural exponential responses, and Waite and Woods (2015) developed approximate-block exact-point designs for conditionally natural exponential responses. Additional discussion of Waite and Woods (2015) is in Appendix B. Additional work for specific responses include Ouwens *et al.* (2006), and Niaparast and Schwabe (2013). Woods and Van de Ven (2011) provided a general method but restricted the model to the form in generalized estimating equations analysis.

Optimal design is an area of experimental design where a criterion, which will be referred to as a design criterion, is optimized with respect to the design points. These design criteria often deal with the variance of parameter estimates or some information measure of the design. In this work, a set of four design criteria were investigated to identify a way to construct optimal designs for GLMs having random blocks. The four criteria ranged from naively ignoring the dependency structure to approximations of the likelihood or score functions. The investigation focused on five models: a second order normal model, a first and second order binomial model, and a first and second order Poisson model, selected for having sufficiently interesting features in the design space, Atkinson & Woods (2015).

2 Material and methods

Ryan (2003) proposed a Bayesian generalization of Lauter's criterion for robust optimal design when there is model uncertainty. This criterion is then placed in a Bayesian decision theoretical framework and is shown to suggest a generalized robust Bayesian A-optimal design criterion. Using generalized Bayesian A-optimality to find the optimal design in a water contamination problem, they obtain optimal designs that are quite different from those resulting from standard Bayesian A-optimality.

Even through there are many challenges, some results have been established for optimal designs under generalized linear models. Yang and Stufken (2009) provided a new algebraic approach for locally optimal design for GLMs with two parameters, which is called the complete class approach. With their results, finding an optimal design can, for many optimality criteria, be restricted to a small class of designs, making it a more tractable problem. Further results using this approach are obtained for locally optimal design for GLMs with group effects Stufken and Yang (2012) and with more than one covariate Yang, *et al* (2011). Yang *et al* (2011) studied A-optimal designs and Dp-optimal designs were studied in Wu and Stufken (2014).

A-optimality: A design is A-optimal for $g(\theta)$ if it minimizes the trace of the inverse of the information matrix for $g(\theta)$, or equivalently the covariance matrix for $g(\theta)$. So ξ_0 is A-optimal for $g(\theta)$ in S if and only if $Tr\left[I_{\xi o}\left(g(\theta)\right)\right] = \min_{\xi \in s} Tr\left[I_{\xi}^{-1}\left(g(\theta)\right)\right]$. Thus, an A-optimal design actually minimizes the average of the asymptotic variance of the Maximum Likelihood Estimates of the elements of $g(\theta)$.

In spite of these and other contributions in recent years, results and theorems for optimal designs under generalized linear models are still very limited. Most of them focus on the model with only main effects as its factorial effects and only a few have considered models with interactions. Besides, most known results only provide some general properties of optimal designs to help search for them. The explicit expressions of optimal designs are pretty much unknown and can only be obtained through algorithms. In addition, full factorial designs are usually considered in known theorems, but in reality, we cannot conduct a full factorial design due to the limitation of time, people or money. Therefore, many areas still require further study, especially models with interactions and optimal fractional factorial designs.

In this paper, we explore Bayesian A-optimal designs under generalized linear models focusing on the Logistic and Poisson models, which are commonly used for binary and count responses will identify explicitly such models. Under the three-stage Bayesian optimal design setting, a hybrid computational technique will be introduced in order to attain a global optimal point and hence the equivalence theorem will be used to identify the Bayesian optimal design.

2.1 Efficiency of the Bayesian a-optimal design

The goal of the Bayesian A-optimal design is to find design points at which the trace of the Fisher information matrix evaluated at the true parameter values is maximized. The A-efficiency is defined as the ratio of the determinant of the

Fisher information matrix with the chosen design points to that with the true A-optimal design points at the true parameter values, i.e.,

$$\xi_2(X) = \frac{tr(X, \beta_{true})}{I(X_{A-opt}, \beta_{true})} \qquad (1)$$

2.2 Three-stage designs

In this section, a three-stage experimental design procedure which utilizes Bayesian techniques is implemented for the Logistic regression model and the Poisson regression model. This procedure is shown in figure 2 which is a practical extension to the two-stage designs of Wang (2018). In addition, a section is devoted to the evaluation of the procedure as is a section concerning the sample distribution between the stages. The two-stage procedure in part originated with some other two-stage work by Minkin (1987) and Myers, Myers, *et al* (1987). Myers *et al* (1987) found and studied two-stage D-Q (non-Bayesian) optimal designs which were found to be considerably more robust and nearly as efficient as the one-stage designs. The main extension from their work and Woods's work is in the use of Bayesian design and Bayesian estimation, hence the extension to the three-stage designs.

2.3 Three-stage design procedure

The two-stage procedure uses two design optimality criteria, one in the second stage and one in the third. The second stage design should be quite robust to poor parameter guesses while not necessarily being very efficient if parameter knowledge is good.

The set of nonlinear equations one solves to obtain the maximum likelihood estimates (MLE) of μ and β is given by

$$\sum_{i=1}^{k}\left(r_i - n_j P_i\right)\left(x_j - \mu\right) = 0 \qquad \ldots(2)$$

$$-\sum_{i=1}^{k}\left(r_i - n_j P_i\right)\beta = 0 \qquad \ldots(3)$$

$$\sum_{i=1}^{k}\left(r_i - n_j Q_i\right)\left(x_j - \mu\right) = 0 \qquad \ldots(4)$$

And

$$-\sum_{i=1}^{k}\left(r_i - n_j Q_i\right)\beta = 0 \qquad \ldots(5)$$

Where k is the number of design points, r_i is the number of successes at the i[th] design point out of n_j P is the centered logistic model and Q is the centered Poisson model. The likelihood function apart from constants for the centered logistic regression model and Poisson regression model is given by

$$L(\mu, \beta; r) \propto \prod_{i=1}^{k} P_i^{r_i}\left(1 - P_i\right)^{n_i - r_i} \qquad \ldots(6)$$

$$L(\mu, \beta; r) \propto \prod_{i=1}^{k} \frac{e^{\beta_i r_i \mu_i} e^{-r_i \mu_i}}{\beta_i!} \qquad \ldots(7)$$

The expression for a joint independent normal prior apart from the constants is then

$$\pi(\mu, \beta) \propto EXP\left[-\tfrac{1}{2}\left(\tfrac{\beta - \mu_\beta}{\sigma_\beta}\right)^2\right] EXP\left[-\tfrac{1}{2}\left(\tfrac{\mu - \mu_\mu}{\sigma_\mu}\right)^2\right] \qquad \ldots(8)$$

Where β and μ are random variables and μ_β, σ_β and μ_μ, σ_μ are their respective means and standard deviations. Consequently, the posterior likelihood, again apart from the constants, is

$$P_{os}L \propto EXP\left[-\tfrac{1}{2}\left(\tfrac{\beta-\mu_\beta}{\sigma_\beta}\right)^2\right]EXP\left[-\tfrac{1}{2}\left(\tfrac{\mu-\mu_\mu}{\sigma_\mu}\right)^2\right]\prod_{i=1}^{k}P_i^{r_i}\left(1-P_i\right)^{n_i-r_i} \quad \dots (9)$$

$$P_{os}L \propto EXP\left[-\tfrac{1}{2}\left(\tfrac{\beta-\mu_\beta}{\sigma_\beta}\right)^2\right]EXP\left[-\tfrac{1}{2}\left(\tfrac{\mu-\mu_\mu}{\sigma_\mu}\right)^2\right]\prod_{i=1}^{k}\frac{e^{\beta_i r_i \mu_i}e^{-r_i\mu_i}}{\beta_i!} \quad \dots (10)$$

Taking the log, the posterior likelihood can be written in the following form

$$Log\left(P_{os}L\right) \propto -\tfrac{1}{2}\left(\tfrac{\beta-\mu_\beta}{\sigma_\beta}\right)^2 - \tfrac{1}{2}\left(\tfrac{\mu-\mu_\mu}{\sigma_\mu}\right)^2 + \sum_{i=1}^{k}\left[r_i Log\left(P_i\right) + n_i Log\left(1-P_i\right) - r_i Log\left(1-P_i\right)\right] \quad \dots (11)$$

And

$$Log\left(P_{os}L\right) \propto -\tfrac{1}{2}\left(\tfrac{\beta-\mu_\beta}{\sigma_\beta}\right)^2 - \tfrac{1}{2}\left(\tfrac{\mu-\mu_\mu}{\sigma_\mu}\right)^2 + \sum_{i=1}^{k}\left[\beta_i r_i \mu_i - e^{r_i\mu_i} - Log\left(\beta_i!\right)\right] \quad \dots (12)$$

Differentiating expressions (12) with respect to β yields

$$\frac{\partial Log\left(P_{os}L\right)}{\partial\beta} = \left(\frac{\mu_\beta-\beta}{\sigma_\beta^2}\right) + \sum_{i=1}^{k}\left(\frac{r_i}{P_i}\frac{\partial P_i}{\partial\beta} - \frac{n_i}{(1-P_i)}\frac{\partial P_i}{\partial\beta} + \frac{r_i}{(1-P_i)}\frac{\partial P_i}{\partial\beta}\right) \quad \dots (13)$$

$$\frac{\partial Log\left(P_{os}L\right)}{\partial\beta} = \left(\frac{\mu_\beta-\beta}{\sigma_\beta^2}\right) + \sum_{i=1}^{k}\left(r_i\mu_i + r_i\mu_i\frac{\partial Q_i}{\partial\beta} + \frac{1}{(\beta_i!)}\frac{\partial Q_i}{\partial\beta}\right) \quad \dots (14)$$

Where $\frac{\partial P_i}{\partial\beta} = P_i\left(1-P_i\right)\left(x_i-\mu\right).$ $\frac{\partial Q_i}{\partial\beta} = e^{r_i\mu_i}.$

Simplifying expression 13 we obtain

$$\frac{\partial Log\left(P_{os}L\right)}{\partial\beta} = \left(\frac{\mu_\beta-\beta}{\sigma_\beta^2}\right) + \sum_{i=1}^{k}\left(r_i - n_iP_i\right)\left(x_i-\mu\right). \quad \dots (15)$$

$$\frac{\partial Log\left(P_{os}L\right)}{\partial\beta} = \left(\frac{\mu_\beta-\beta}{\sigma_\beta^2}\right) + \sum_{i=1}^{k}\left(r_i - n_iQ_i\right)\left(x_i-\mu\right). \quad \dots (16)$$

Differentiating expression 12 with respect to μ yields

$$\frac{\partial Log\left(P_{os}L\right)}{\partial\mu} = \left(\frac{\mu_\mu-\mu}{\sigma_\mu^2}\right) + \sum_{i=1}^{k}\left(\frac{r_i}{P_i}\frac{\partial P_i}{\partial\mu} - \frac{1}{(1-P_i)}\frac{\partial P_i}{\partial\mu} + \frac{r_i}{(1-P_i)}\frac{\partial P_i}{\partial\mu}\right) \quad \dots (17)$$

$$\frac{\partial Log\left(P_{os}L\right)}{\partial\mu} = \left(\frac{\mu_\mu-\mu}{\sigma_\mu^2}\right) + \sum_{i=1}^{k}\left(\beta_i r_i + \mu_i r_i\frac{\partial Q_i}{\partial\mu} + \frac{1}{(\beta_i!)}\frac{\partial Q_i}{\partial\mu}\right) \quad \dots (18)$$

Where $\frac{\partial P_i}{\partial\mu} = -\beta\left(P_i\left(1-P_i\right)\right).$ $\frac{\partial P_i}{\partial\mu} = e^{r_i\beta_i}$

After some simplification, expression 17 becomes

$$\frac{\partial Log(P_{os}L)}{\partial \mu} = \left(\frac{\mu_\mu - \mu}{\sigma_\mu^2}\right) - \sum_{i=1}^{k} \beta\left(r_i - n_i P_i\right) \quad \text{...(19)}$$

$$\frac{\partial Log(P_{os}L)}{\partial \mu} = \left(\frac{\mu_\mu - \mu}{\sigma_\mu^2}\right) - \sum_{i=1}^{k} \beta\left(r_i - n_i Q_i\right) \quad \text{...(20)}$$

Equating expressions 19 and 20 to zero produces the set of nonlinear equation which when solved yield the posterior mode. The pair of equations

$$\left(\frac{\mu_\beta - \beta}{\sigma_\beta^2}\right) + \sum_{i=1}^{k} \left(r_i - n_i P_i\right)\left(x_i - \mu\right) = 0 \quad \text{...(21)}$$

$$\left(\frac{\mu_\mu - \mu}{\sigma_\mu^2}\right) - \sum_{i=1}^{k} \beta\left(r_i - n_i P_i\right) = 0 \quad \text{...(22)}$$

$$\left(\frac{\mu_\beta - \beta}{\sigma_\beta^2}\right) + \sum_{i=1}^{k} \left(r_i - n_i Q_i\right)\left(x_i - \mu\right) = 0 \quad \text{...(23)}$$

$$\left(\frac{\mu_\mu - \mu}{\sigma_\mu^2}\right) - \sum_{i=1}^{k} \beta\left(r_i - n_i Q_i\right) = 0, \quad \text{...(24)}$$

Is an adjusted version of the set given by 8 - 12. The adjustment terms keep the estimation of μ and β from straying too far away from their respective prior means in units of prior variance. If a joint uniform distribution is used as the prior, the posterior mode is simply given by the maximum likelihood estimates subject to the constraints that $\beta_L \le \hat{\beta} \le \beta_U$ and $\mu_L \le \hat{\mu} \le \mu_U$. The estimates become the nearest endpoint of the prior if the maximum likelihood estimates should fall outside the uniform prior.

The likelihood for the second stage is the same as the likelihood for the double stage experiment.

$$L_2\left(\mu, \beta; r_2 / x_2\right) = \prod_{i=1}^{k_2} \binom{n_{2i}}{r_{2i}} P_{2i}^{\,r_{2i}} \left(1 - P_{2i}\right)^{n_{2i} - r_{2i}} \quad \text{...(25)}$$

$$L_2\left(\mu, \beta; r_2 / x_2\right) = \prod_{i=1}^{k_2} \frac{e^{\beta_i r_i \mu_i} e^{-r_i \mu_i}}{\beta_i!} \quad \text{...(26)}$$

Where

$$P_{2i} = \frac{1}{1 + \exp^{\left(\xi - \beta\left(x_{2i} - \mu\right)\right)}}$$

Is the probability of success at the i^{th} design point in the second stage, r_2 is a $(k_2 \times 1)$ vector of responses in the second stage in which each element is the number of successes out of $n_{2i,}$ the sample size at the i^{th} design point, k_2 is the number of design points in the second stage, and x_2 is a $(k_2 \times 1)$ vector of design point locations. We know that the total or joint likelihood can be expressed as a product of the second stage likelihood and the third stage likelihood conditioned on the second.

$$L_{2,3}\left(\mu,\beta;r_2,r_3/x_2,x_3\right)=L_2\left(\mu,\beta;r_2/x_2\right)L_{3/2}\left(\mu,\beta;r_3/r_2,x_2,x_3\right) \quad \ldots(27)$$

Where the conditional likelihood for the third stage given the second stage is given by

$$L_{3/2}\left(\mu,\beta;r_3/r_2,x_3\right)=\prod_{i=1}^{k_3}\binom{n_{3j}}{r_{3j}}P_{3j}^{\,r_{3j}}\left(1-P_{3j}\right)^{n_{3j}-r_{3j}} \quad \ldots(28)$$

$$L_{3/2}\left(\mu,\beta;r_3/r_2,x_2,x_3\right)=\prod_{i=1}^{k_3}\frac{e^{\beta_{3i}r_{3i}\mu_{3i}}e^{-r_{3i}\mu_{3i}}}{\beta_{3i}!} \quad \ldots(29)$$

And

$$P_{3j}=\frac{1}{1+\exp^{\left(\xi-\beta\left(x_{3i}-\mu\right)\right)}}$$

Is again the probability of success at the j^{th} design point in the third stage, r_3 is a $\left(k_3\times1\right)$ vector of responses in the third stage in which each element is the number of successes out of n_{3j}, the sample size at the j^{th} design point, and k_3 is the number of design points in the third stage.

In this dissertation, the number of design points in the third stage, k_3 is three. In order for the entire three-stage design to be highly efficient, it must resemble the optimal 3-level designs. In order to address its goal of robustness, the third stage needs just 3 levels to complete a three-stage design which has the complexion of an optimal 3 level design.

Taking the natural logarithm of equation 29 results in

$$LogL_{2,3}=LogL_2+LogL_{3/2} \quad \ldots(30)$$

After expanding,

$$LogL_{2,3}=\sum_{i=1}^{k_2}\left[\log\binom{n_{2i}}{r_{2i}}+r_{2i}\log\left(P_{2i}\right)+\left(n_{2i}-r_{2i}\right)\log\left(1-P_{2i}\right)\right]+\sum_{j=1}^{k_3}\left[\log\binom{n_{3j}}{r_{3j}}+r_{3j}\log\left(P_{3j}\right)+\left(n_{3j}-r_{3j}\right)\log\left(1-P_{3j}\right)\right] \quad \ldots(31)$$

Equation 31 can then be used directly to find the Fishers information matrix. However, the total or joint Fishers information is merely the sum of the individual information matrices formed from equations 24 and 28, the two likelihoods.

The total information matrix for three stage procedure is then

$$I_{2,3}\left(\mu,\beta\right)=\begin{pmatrix}I_{22} & I_{23}\\ I_{23} & I_{33}\end{pmatrix}$$

Where

$$I_{22} = \beta^2 \sum_{i=1}^{k_2} n_{2i} p_{2i} \left(1 - p_{2i}\right) + \beta^2 \sum_{j=1}^{k_3} n_{3j} p_{3j} \left(1 - p_{3j}\right),$$

$$I_{23} = -\beta \sum_{i=1}^{k_2} \left(x_{2i} - \mu\right) n_{2i} p_{2i} \left(1 - p_{2i}\right) - \beta \sum_{j=1}^{k_3} \left(x_{3j} - \mu\right) n_{3j} p_{3j} \left(1 - p_{3j}\right) \text{ and}$$

$$I_{33} = \sum_{i=1}^{k_2} \left(x_{2i} - \mu\right)^2 n_{2i} p_{2i} \left(1 - p_{2i}\right) + \sum_{j=1}^{k_3} \left(x_{3j} - \mu\right) n_{3j} p_{3j} \left(1 - p_{3j}\right)$$

The inverse of the information matrix is given by,

$$I_{2,3}^{-1}\left(\mu, \beta\right) = \begin{pmatrix} U_{22} & U_{23} \\ U_{23} & U_{33} \end{pmatrix}$$

Where

$$U_{22} = \frac{I_{33}}{D}, \qquad U_{23} = \frac{I_{23}}{D}, \qquad U_{33} = \frac{I_{22}}{D}, \qquad \text{and } D \text{ is the determinant of } I_{2,3}\left(\mu, \beta\right).$$

3 Results and discussion

Implementation of Bayesian Three-Stage A-Optimal Designs

Here, we will show examples of Bayesian three-stage D-optimal and A-optimal designs for Logistic regression and Poisson regression models. The A-optimality criteria for the second stage design is defined in the previous section, which is to choose the design measure η maximizing $\phi_1\left(\eta\right)$ and minimizing $\phi_2\left(\eta\right)$. For the third stage, the design criteria are to maximize and minimize the equations defined in (7-8).

The hybrid computational method discussed previously is used to find the three-stage optimal designs efficiently.

Thus, we would expect that three-stage designs work better for larger and even smaller samples. In this simulation, we choose the total number of sample size ***n*** as 200. Suppose that we have ***n*** experimental runs in total and we use $\boldsymbol{n_2}$ experimental runs at the second stage and $\boldsymbol{n_3=n-n_2}$ runs at the third stage. One important question is how to allocate the ***n*** runs optimally to get the best result. In this chapter, the equal sample size is chosen for the second and third stage, i.e., $\boldsymbol{n_2 = n_3 = 100}$.

Logistic Regression Models

The model can be written as

$$y_{ij} \sim Logistic\left(n_i\right), \qquad \ldots (32)$$

$$\text{Where } n_i = \frac{1}{1 + \exp\left(-x_i^T \beta\right)}$$

We define y_{ij} in (32) to be the response for the j^{th} replicate of the i^{th} design point and assume it follows a logistic distribution with n_i as the mean; x_i is the regressor vector at the i^{th} point and β is the parameter vector. For the one-variable and two-variable models,

$$n_i = \frac{1}{1+\exp\left(-\beta_0 - \beta_1 x_i\right)} \quad \dots (33)$$

$$n_i = \frac{1}{1+\exp\left(-\beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}\right)} \quad \dots (34)$$

Denote by p_i the proportion of whole sample size at i[th] design point x_i, with $\sum_{i=1}^{k} p_i = 1$. Also denote $w_i = n_i\left(1-n_i\right)$ for i = 1, 2, . . . , k. The Fisher information matrix I(β, η) for the logistic regression model can be written as

$$I\left(\beta,\eta\right) = \sum_{i=1}^{k} p_i w_i x_i x_i^T \quad \dots (35)$$

where x_i is a p × 1 design vector of the i[th] design points.

The regressors here are not limited and for this study we coded x as (0,1).

The tables below show three applications of the Bayesian A-optimal designs for the Logistic regression models. We will use A-efficiency measure described in the previously to evaluate the efficiencies of the Bayesian three-stage design.

Table 1 A one variable Logistic regression model for D-optimal design, n = 200. The true model is $n = 1 \big/ \left(1+\exp\left(-\left(-1.0725+0.0279 Age\right)\right)\right)$

$\beta_0 \sim N(1,3) I[-1,2], \beta_1 \sim N(-4,2) I[-6,2]$						
Point	**True D-optimal Design**		**Second-Stage Design**		**Third-Stage Design**	
	Pi	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**
1	0.5	0.319	0.421	0.812	0.272	0.673
2	0.5	-0.252	0.212	0.385	0.473	0.328

Table 2 A one variable Logistic regression model for D-optimal design with different priors.

$\beta_0 \sim N(1,3) I[-1,2], \beta_1 \sim N(-4,2) I[-6,-1]$						
Point	**True D-optimal Design**		**Second-Stage Design**		**Third-Stage Design**	
	Pi	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**
1	0.5	0.319	0.242	0.415	0.651	0.881
2	0.5	-0.252	0.232	0.379	0.471	0.865
$\beta_0 \sim N(1,3) I[-1,2], \beta_1 \sim N(-2,2) I[-6,-1]$						
Point	**True D-optimal Design**		**Second-Stage Design**		**Third-Stage Design**	
	P_i	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**
1	0.5	0.319	0.113	0.578	0.563	0.792

2	0.5	-0.252	0.282	-0.432	0.687	0.654

Table 3 A one variable Logistic regression model for A-optimal design. n = 200 and n_2 = 100

$\beta_0 \sim N(1,3)I[-1,2], \beta_1 \sim N(-4,2)I[-6,2]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	P_i	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**
1	0.5	0.521	0.232	0.573	0.546	0.745	0.572	0.867
2	0.5	0.238	0.421	0.348	0.529	0.213	0.048	0.291
$\beta_0 \sim N(1,3)I[-1,2], \beta_1 \sim N(-4,2)I[-6,-1]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	P_i	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**
1	0.5	0.521	0.352	0.752	0.598	0.794	0.795	0.623
2	0.5	0.238	0.423	-0.483	0.531	-0.206	0.606	-0.002

Table 4 A one variable Logistic regression model for A-optimal design. n = 200 and n_2 = 100 with different priors

$\beta_0 \sim N(1,3)I[-1,2], \beta_1 \sim N(-2,2)I[-6,-1]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	P_i	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**
1	0.5	0.521	0.322	0.641	0.385	0.706	0.542	0.787
2	0.5	0.238	0.311	-0.182	0.254	-0.013	0.568	0.365

Table 5 A two variable Logistic regression model for A-optimal design. n = 200 and n_2 = 100. The true model is n = 1/(1+exp(-(-1.1107+0.0279*Age + 0.0018*BMI)))

$\beta_0 \sim N(1,3)I[-1,2], \beta_1 \sim N(-4,2)I[-6,2]$						
Point	**True A-optimal Design**		**Second-Stage Design**		**Third-Stage Design**	
	P_i	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**
1	0.3	(1,1)	0.1	(1,1)	0.612	(1,1)
2	0.3	(1,0.4)	0.1	(1,0.432)	0.452	(1,0.532)
3	0.3	(0,1)	0.1	(0,1)	0.489	(0.142,1)

Table 6 A two variable Logistic regression model for A-optimal design. n = 200 and n_2 = 100 with different priors.

$\beta_0 \sim N(1,3)I[-1,2], \beta_1 \sim N(-2,2)I[-6,-1]$						
Point	**True A-optimal Design**		**Second-Stage Design**		**Third-Stage Design**	
	P_i	**X_i**	**P_i**	**X_i**	**P_i**	**X_i**
1	0.3	(1,1)	0.1	(1,1)	0.087	(1,1)
2	0.3	(1,0.4)	0.1	(1,0.043)	0.217	(1,0.321)
3	0.3	(0,1)	0.1	(0,1)	0.342	(0.241,1)

3.1 Poisson Regression Models

A Poisson regression model, is useful in modelling a random variable of counts, is a generalized linear model with unknown parameters in the information matrix. The Poisson regression model may be written as

$$y_{ij} \sim Poisson(\lambda_i),$$
$$where$$
$$\lambda_i = \exp(x_i^T \beta), \quad \ldots$$
$$for$$
$$i = 1, 2, \ldots, k \quad (36)$$

where y_{ij} is the response for the j^{th} replicate of the i^{th} design point, x_i the regressor vector at the i^{th} point, β the parameter vector, and λ_i the Poisson mean at i^{th} point. We are going to investigate the following one-variable, and two-variables, respectively,

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i}) \quad \ldots (37)$$

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}) \quad \ldots (38)$$

Denote by n_i the proportion of whole sample size at i^{th} design point x_i with $\sum_{i=1}^{k} n_i = 1$.

Also denote by $w_i = x_i$ for $i = 1, 2, \ldots, k$.

Table 7 A one variable Poisson regression model for A-optimal design. n = 200 and n_2 = 100. The true model is ƛ = exp(1.2589-0.0011*Age)

$\beta_0 \square N(1,3)I[-1,2], \beta_1 \square N(-4,2)I[-6,2]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	Pi	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**
1	0.5	0	0.321	0	0.382	0	0.582	0
2	0.5	1	0.272	1	0.294	1	0.321	1
$\beta_0 \square N(1,3)I[-1,2], \beta_1 \square N(-4,2)I[-6,-1]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	Pi	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**
1	0.5	0	0.324	0	0.533	0	0.621	0
2	0.5	1	0.432	1	0.540	1	0.702	1

Table 8 A one variable Poisson regression model for A-optimal design. n = 200 and n_2 = 100 with different priors.

$\beta_0 \square N(1,3)I[-1,2], \beta_1 \square N(-2,2)I[-6,-1]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	P_i	X_i	P_i	X_i	P_i	X_i	P_i	X_i
1	0.5	0	0.152	0	0.241	0	0.407	0
2	0.5	1	0.074	1	0.102	1	0.191	1

Table 9 A two variable Poisson regression model for A-optimal design. n = 200 and n_2 = 100. The true model is ƛ = exp(1.3288 - 0.0013*Age – 0.0077*Study)

$\beta_0 \square N(1,3)I[-1,2], \beta_1 \square N(-4,2)I[-6,2]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	Pi	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**
1	0.3	(0,1)	0.231	(0,1)	0.272	(0,1)	0.632	(0,1)
2	0.3	(1,0)	0.218	(1,0)	0.187	(1,0)	0.752	(1,0)
3	0.3	(1,1)	0.178	(1,1)	0.062	(1,1)	0.392	(1,1)
$\beta_0 \square N(1,3)I[-1,2], \beta_1 \square N(-2,2)I[-6,-1]$								
Point	**True A-optimal Design**		**First-Stage Design**		**Second-Stage Design**		**Third-Stage Design**	
	Pi	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**	**Pi**	**Xi**
1	0.3	(0,1)	0.201	(0,1)	0.272	(0,1)	0.721	(0,1)
2	0.3	(1,0)	0.327	(1,0)	0.232	(1,0)	0.762	(1,0)
3	0.3	(1,1)	0.301	(1,1)	0.420	(1,1)	0.672	(1,1)

3.2 Evaluation of Bayesian Three-Stage A-Optimal Designs (CHD)

3.2.1 A-Efficiency

The goal of A-optimal design is to minimize the trace of the inverse of the Fisher information matrix at the true parameter values. Thus, a logical way to evaluate a three-stage design is to compare the trace of the inverse of the Fisher information matrix of the three-stage design to that of the two-stage true A-optimal design.

We define A-efficiency as follows:

$$A-eff_{3rdstage} = \frac{tr\left(X_{3rdstage}, \beta_{true}\right)}{I\left(X_{A-opt}, \beta_{true}\right)} \quad \ldots (39)$$

Table 10 Comparisons of the A-efficiency between the Bayesian one-stage, two-stage and three-stage A-optimal designs for one-variable Logistic regression models. n = 200, n_2 = 100 and N = 100.

$n = 1 \big/ \left(1+\exp\left(-\left(-1.0725+0.0279Age\right)\right)\right)$			
Prior	**One Stage**	**Two Stage**	**Three Stage**
$B_0 \sim N(1,2)\ I[0,2]$ $B_1 \sim N(3,2)\ I[0,6]$	0.8529	0.9219	0.9832
$B_0 \sim N(1,2)\ I[0,2]$ $B_1 \sim N(2,6)\ I[0,6]$	0.8254	0.8021	0.9781
$B_0 \sim N(1,2)\ I[0,2]$ $B_1 \sim N(1,3)\ I[0,6]$	0.8926	0.8835	0.9920
$B_0 \sim U[0,2]$ $B_1 \sim U[-4.5,5.5]$	0.6732	0.8751	0.9103
$B_0 \sim U[0,2]$ $B_1 \sim U[0,5]$	0.9721	0.9722	0.9702

Table 11 Comparisons of the A-efficiency between the Bayesian one-stage, two-stage and three-stage A-optimal designs for one-variable Logistic regression models. n = 200, n_2 = 100 and N = 100 with different priors.

$n = 1 \big/ \left(1+\exp\left(-\left(-1.0725+0.0279Age\right)\right)\right)$			
Prior	**One Stage**	**Two Stage**	**Three Stage**
$B_0 \sim N(1,3)\ I[-1,2]$ $B_1 \sim N(-4,2)\ I[-6,2]$	0.8893	0.8890	0.9874
$B_0 \sim N(1,3)\ I[-1,2]$ $B_1 \sim N(-4,2)\ I[-6,-1]$	0.9805	0.9821	0.9903
$B_0 \sim N(1,3)\ I[-1,2]$ $B_1 \sim N(-2,2)\ I[-6,-1]$	0.8531	0.8865	0.9721
$B_0 \sim U[-1,2]$ $B_1 \sim U[-8,0]$	0.8843	0.8904	0.9904
$B_0 \sim U[-1,2]$ $B_1 \sim U[-6,0]$	0.8932	0.9327	0.9684

Table 12 Comparisons of the A-efficiency between the Bayesian one stage, two-stage and three-stage A-optimal designs for two-variable Logistic regression models. n = 200, n_2 = 100 and N = 100

n = 1/(1+exp(-(-1.1107+0.0279*Age + 0.0018*BMI)))			
Prior	**One Stage**	**Two Stage**	**Three Stage**
$B_0 \sim N(1,2)\ I[0,2]$ $B_1 \sim N(3,2)\ I[0,6]$	0.7983	0.8807	0.9023
$B_0 \sim N(1,2)\ I[0,2]$ $B_1 \sim N(2,6)\ I[0,6]$	0.8805	0.8832	0.8987
$B_0 \sim N(1,2)\ I[0,2]$ $B_1 \sim N(1,3)\ I[0,6]$	0.8795	0.8953	0.9945
$B_0 \sim U[0,2]$ $B_1 \sim U[-4.5,5.5]$	0.8634	0.8742	0.9832
$B_0 \sim U[-2,2]$ $B_1 \sim U[0,5]$	0.8906	0.8998	0.9964

Table 13 Comparisons of the A-efficiency between the Bayesian one stage, two-stage and three-stage A-optimal designs for two-variable Logistic regression models. n = 200, n_2 = 100 and N = 100 with different priors.

n = 1/(1+exp(-(-1.1107+0.0279*Age + 0.0018*BMI)))			
Prior	**One Stage**	**Two Stage**	**Three Stage**
$B_0 \sim N(1,3)\ I[-1,2]$ $B_1 \sim N(-4,2)\ I[-6,-1]$	0.8845	0.8931	0.9852
$B_0 \sim N(1,3)\ I[-1,2]$ $B_1 \sim N(-2,2)\ I[-6,-1]$	0.8956	0.8894	0.9890
$B_0 \sim N(1,3)\ I[-1,2]$ $B_1 \sim N(-1,3)\ I[-6,-1]$	0.8826	0.8935	0.9847
$B_0 \sim U[-1,2]$ $B_1 \sim U[-8,0]$	0.9004	0.9097	0.9906
$B_0 \sim U[-1,2]$ $B_1 \sim U[-6,0]$	0.9306	0.9598	0.9978

Table 14 Comparisons of the A-efficiency between the Bayesian one-stage, two-stage and three-stage A-optimal designs for one-variable Poisson regression models. n = 200, n_2 = 100 and N = 100.

λ = exp(1.2589-0.0011*Age)			
Prior	**One Stage**	**Two Stage**	**Three Stage**
$B_0 \sim N(1,2)\ I[0,3]$ $B_1 \sim N(0.5,6) I[0,5]$	0.8972	0.8997	0.9945
$B_0 \sim N(1,2)\ I[0,3]$ $B_1 \sim N(1,6)\ I[0,5]$	0.9042	0.9132	0.9894
$B_0 \sim N(1,2)\ I[0,3]$ $B_1 \sim N(1,3)\ I[0,5]$	0.8807	0.8935	0.9725

$B_0 \sim U[-1,3]$ $B_1 \sim U[-5.5,6.5]$	0.8846	0.8952	0.9957
$B_0 \sim U[-1,3]$ $B_1 \sim U[0,5]$	0.8849	0.8972	0.9877

Table 15 Comparisons of the A-efficiency between the Bayesian one-stage, two-stage and three-stage A-optimal designs for one-variable Poisson regression models. n = 200, n_2 = 100 and N = 100 with different priors.

λ̂ = exp(1.2589-0.0011*Age)			
Prior	**One Stage**	**Two Stage**	**Three Stage**
B_0~N(1,3) I[-1,3] B_1~N(-4,6) I[-6,0]	0.8763	0.8876	0.9907
B_0~N(1,3) I[-1,3] B_1~N(-4,3) I[-6,0]	0.8732	0.8895	0.9896
B_0~N(1,3) I[-1,3] B_1~N(-2,3) I[-6,0]	0.8807	0.8974	0.9963
$B_0 \sim U[-1,3]$ $B_1 \sim U[-8,0]$	0.8794	0.8805	0.9842
$B_0 \sim U[-1,3]$ $B_1 \sim U[-6,0]$	0.8746	0.8945	0.9942

Table 16 Comparisons of the A-efficiency between the Bayesian one-stage, two-stage and three-stage A-optimal designs for two-variable Poisson regression models. n = 200, n_2 = 100 and n_3 = 100

λ̂ = exp(1.3288 - 0.0013*Age – 0.0077*Study)			
Prior	**One Stage**	**Two Stage**	**Three Stage**
B_0 ~ N(0,3) I[-1,1] B_1~N(0.5,3) I[-1,1] B_2~N(1,2) I[0,3]	0.8974	0.8979	0.9896
B_0 ~ N(0,3) I[-1,1] B_1 ~ N(0.5,1) I[-1,1] B_2~N(1,2) I[0,3]	0.8867	0.8945	0.9947

Table 17 Comparisons of the A-efficiency between the Bayesian one-stage, two-stage and three-stage A-optimal designs for two-variable Poisson regression models. n = 200, n_2 = 100 and n_3 = 100 with different priors.

λ̂ = exp(1.3288 - 0.0013*Age – 0.0077*Study)			
Prior	**One Stage**	**Two Stage**	**Three Stage**
B_0~N(2,1) I[1,3] B_1~N(-1,1) I[-2,0] B_2~N(-3,1) I[-4,0]	0.8732	0.8721	0.9768
B_0~N(2,1) I[1,3]	0.8977	0.8989	0.9975

B_1~N(-1,1) I[-2,0] B_2~N(-3,1) I[-6,0]			

4 Conclusion

Tables 6-8 show the simulated Bayesian three-stage A-efficiency values for the Poisson regression models given in Tables 9–10 and more. In Tables 8 and 9, we list the simulated Bayesian three-stage A-efficiency values for the logistic regression models given in Tables 2–3 and more.

Using the simulated results from Tables 1–17, we have the following comments and conclusions.

- In general, Bayesian three-stage A-optimal designs achieve better A-efficiency than the two-stage design.
- The three-stage design is more effective for the cases with less A-efficiency in the second stage. The A-efficiency difference due to the prior knowledge on parameters is reduced using the three-stage approach. For the designs with the two-stage A-efficiency about 70 – 80%, the three-stage design can increase it to around 96%. For those with the two-stage A-efficiency about 70 – 80%, the three-stage design can increase it to around 97%. For the designs with good A-efficiency in the second stage, the third-stage design does not have much space to improve it. For example, in Table (10) on first two cases of right side, the two-stage designs already achieve the A-efficiency of 90% and 92%, respectively, while the three-stage designs have the A-efficiency 97% and 98%, respectively.
- The mode of the truncated normal prior or the center of the uniform prior is a very important factor to the A-efficiency. In the situation that all the other factors (variance and range) are the same, the design can achieve greater A-efficiency when the mode of the truncated normal prior or the center of the uniform prior is on the true parameter values. This result can be easily seen from several examples in Tables 16 and 17.
- As the prior uncertainty (variance and range) increases, the A-efficiency of the two stage and three-stage design usually decreases. It is true in general when the mode of the truncated normal prior or the center of the uniform prior the A-efficiency is on the true parameter values. But when they are not on the true parameter values, the A-efficiency could be better in the case that larger uncertainty is assumed on the prior. This can be seen from the second and third examples on the left side of Table 12. In such cases, the true parameter values get higher probability densities with more spread prior so that a better A-efficiency can be achieved.

We have investigated the Bayesian three-stage A-optimal design for generalized linear models. Particularly, we discussed the one-variable and two-variable Logistic and Poisson regression models. We presented the procedure and algorithm of the Bayesian three stage A-optimal design. We illustrated the Bayesian two-stage A-optimal design procedure by using examples of representative Logistic and Poisson regression models and listed the true, second-stage and third-stage A-optimal design points with the proportions of the associated sample size. It can be concluded that if the second-stage design is not really optimal, once the data are observed, the third-stage design usually tries to balance the design from the second-stage and make the overall design close to the true A-optimal design.

Due to the dependence between the third-stage design and the second-stage data, it is very complicated to use the expected Fisher information matrix for the three-stage designs. We discussed how to use simulation to approximate the Fisher information matrix, which further enabled us to compute A-efficiency of the two-stage design. We evaluated the performance of the three-stage designs in terms of the A-efficiency for various Poisson and Logistic regression models. It has been demonstrated that three-stage designs are more efficient and robust than two-stage designs.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Atkinson, A. C. and D. C. Woods (2015). Designs for generalized linear models. In A. M. Dean, M. D. Morris, J. Stufken, and D. R. Bingham (Eds.), Handbook of Design and Analysis of Experiments. Boca Raton: Chapman & Hall/CRC.

[2] Box, G. E. P. and D. R. Cox (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society B* ***26(6),*** 211-246.

[3] Chaloner K. (1989). Bayesian design for estimating the turning point of a quadratic regression. Communications in Statistics. *Theory and Methods.**18**(4),* 1385-1400.

[4] Chaloner, K., and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference **21**(9)*, 191-208.

[5] Minkin, S. (1987). Optimal Designs for Binary Data. *Journal of the American Statistical Association, **82**(5),* 1098 - 1103.

[6] Myers, R. H., D. C. Montgomery, G. G. Vining, and T. J. Robinson. (1987). Generalized linear models with applications in engineering and the sciences. 2nd ed. Hoboken, NJ:Wiley.

[7] Nelder, J. and Wedderburn, R. (1972). Generalized Linear Models. *Journal of Royal Statistical Society-Series B, **135**(8)*, 370-384.

[8] Niaparast, A, Schwabe, H., Laine, M. (2013) Simulation-based optimal design using a response variance criterion. *Journal of Computational and Graphical Statistics. **20**(1)*, 24-52.

[9] Ouwens, E. S., Woods, D. C. and Pettitt S. U. (2006). Fully Bayesian experimental design for Engineering studies. *Entropy **17**(8)*, 1063-1089.

[10] Stufken, J. and M. Yang (2012). On locally optimal designs for generalized linear models with group effects. *Statistica Sinica **22**(2),* 1765-1786.

[11] Waite, T.W., and Woods, D. C. (2015). Singular prior distributions in Bayesian D-optimal design for nonlinear models. arXiv:1506.02916. 05.08.2020.

[12] Wang, Z. (2018). Locally D-optimal designs for generalized linear models (Doctoral dissertation, Arizona University).

[13] Woods D. C., Overstall, A. M., Adamou, M. and Waite T. W. (2018). Bayesian Design of Experiments for generalized linear models and dimensional analysis with industrial and scientific application. *Quality Engineering **29**(1),* 91-103.

[14] Woods, D. C. and P. van de Ven (2011). Blocked designs for experiments with non-normal response. *Technometrics. **53**(5),* 173-182.

[15] Wu, H. P. and Stufken, J. (2014). Locally φp-optimal designs for generalized linear models with a single-variable quadratic polynomial predictor. *Biometrika **101**(5)*, 365-375.

[16] Yang, M., B. Zhang and S. Huang, (2011). Optimal designs for generalized linear models with multiple design variables. *Statistica Sinica **21**(7),* 1415-1430.

[17] Sung, C.-L., Hung, Y., Rittase, W., Zhu, C., and Wu, C. (2019). A generalized Gaussian process model for computer experiments with binary time series," Journal of the American Statistical Association, 1–42.

Ryan, K. J. (2003). Estimating expected information gains for experimental designs with application to the random fatigue-limit model. Journal of Computational and Graphical Statistics 12, 585–603.

Author's short biography

Ali, H. is an academic staff in the department of Mathematics, Faculty of Natural Sciences, University of Jos. He holds Masters Degree in Statistics and currently undergoing his PhD in Biostatistics with bias in Bayesian inference. He has published over seven (7) articles in high impact factor journals, He is a member of many professional bodies among them are SIAM, PSSN, NMS etc.

Lasisi, K. E. is a professor of Statistics in the Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Nigeria. He has published over forty(30) articles in high impact factor journals, He is a member of many professional bodies amongst them are SIAM, PSSN, NMS etc.

Nwaosu, S. C. is a professor of Statistics in the department of Statistics, Joseph Sawuaan Tarka University of Agriculture, Makurdi, Nigeria. He has published over forty(40) articles in high impact factor journals, He is a member of many professional bodies amongst them are SIAM, PSSN, NMS etc.

Abdulkadir, A. is a professor of Statistics in the Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Nigeria. He has published over forty(30) articles in high impact factor journals, He is a member of many professional bodies amongst them are SIAM, PSSN, NMS etc.