Bayesian Analysis of the Level of Resistance of Four Varieties of Maize to Downy Mildew Disease

*Adeleke, M.O.

Department of Statistics, University of Ilorin, Ilorin, Nigeria

Abstract

This paper is based on Bayesian analysis of an experiment conducted by Kwara State Agricultural Development Project to determine the level of resistance of four maize varieties to Downy mildew disease. The four varieties of maize namely DMLSR-White, DMRESR-Yellow, 43DMR-SR-White and TZSR-White were planted in six different farms known to be affected by the disease. Three variables were observed from each farm; the height of the plants, the weight of cobs, and the weight of grains. The design corresponding to this experiment is a randomized complete block design, with the varieties as the variable of interest while the farms serve as the blocks. Bayesian analysis was carried out on each of the three variables, which involves estimating the values of parameters in the model from samples generated from the posterior distribution of the parameters. Gibbs sampler was used to estimate the values of these parameters and inference was made based on 95% equal tails credible interval. It was discovered that the four maize varieties are equally resistant to the disease although the result revealed that the prevalent of the downy mildew disease is significantly different among the farms.

1.0   Introduction

Maize also called Corn is a grass domesticated by indigenous people in Mesoamerica in prehistoric times. The Aztecs and Mayans cultivated the crop in numerous varieties throughout Central and Southern Mexico to cook or grind [1, 2]. In Nigeria, maize is one of the staples that are widely grown. However, its production is more in the Northern part than in the Southern part of the country, which is largely forest in nature. Due to the scanty rainfall in the northern part of the country, maize growers support their production with fertilizer and irrigation. Maize as a cereal crop is high yielding, easy to process and readily digested. It is a versatile crop that grows across a range of agro-ecological zones all over the world. Over 50% of the population in the Sub-Saharan Africa use maize as a staple food. In Nigeria, particularly, more than 60% of maize produced is consumed by the industrial sector.

The downy mildew disease is a fungal disease that affects grain crops and it is prevalent in Asia and Africa. It is a disease in many tropical regions of the world accounting for about 70% loss in maize production. The symptoms of the disease vary depending on the age of the plant and prevailing climatic conditions. For maize below four weeks, there is stunted and chlorotic plant, and sometimes premature death. At older age, the plants experience molting, chlorotic streaking with lesion and white striped leaves observed in lower leaf surface, which eventually shred. Infected plants are usually stunted and have malformed reproductive organ (bushy tassels and eaves) which may not seed [3].

In hypothesis testing, hypotheses are set about some unknown parameters and appropriate tests are carried out on them. Bayesian method of analysis assumes all unknown parameters are random variables and therefore have probability distributions. This is in contrast to the frequentist approach that assumes the unknown parameters are constant. Therefore, appropriate probability distributions (also known as the prior distributions) must be assigned to the parameters.

---

* Corresponding Author: Tel: +234(0)7030759054; Email: adeleke.mo@unilorin.edu.ng

Bayesian method of analysis uses information from the data (likelihood) as well as the prior information on the parameter(s) of interest to arrive at a final inference about the parameter(s) which is known as the posterior. It estimates the values of the parameters in the model by summarizing samples generated from the posterior distribution of the parameters. In most cases, it is easier to work with conjugate priors [4], because the prior and posterior distributions belong to the same family of distribution. This makes it easier to sample observations from the posterior distribution.

Based on this background, the present study was designed to determine which of four maize varieties (DMLSR – White, DMRESR – Yellow, 43DMRSR – White and TZSR – White) is most resistant to downy mildew disease in some infected areas in Tsaragi, Kwara State. The resistance will be assessed via the height of plants, weight of cobs and weight of grains from each variety. The result of the work could also serve as prior information for future works.

2.0   Materials and Methods

2.1   Source of Data

The data used in this study are secondary data obtained from Kwara State Agricultural Development Project (KWADP). Six farms which consist of four plots each were used. On each farm, the four maize varieties were planted on each of the four plots. Each plot was made up of 5 rows of 5 m long, the spacing of the maize was 1.2 m by 0.3 m with two plants per stand resulting in a population density of about 50 plants per plot.

The data of interest in this work is a product of a randomized complete block design. The four varieties are the treatments of interest while the farms serve as the blocking factor. This data was analyzed using a Bayesian approach [5, 6].

2.2   The Statistical Model

The model for the randomized complete block design is given as follows [7]

$$y_{ij} = \mu + \tau_i + \beta_j + e_{ij} \tag{1}$$

Equation (1) above can be written in matrix form as follows

$$Y = XB + \varepsilon \tag{2}$$

$$\varepsilon \sim N(0, \sigma^2)$$

$$B = \begin{bmatrix} \mu \\ \tau \\ \beta \end{bmatrix} \text{ and } X = [1 \quad \Delta^T \quad D^T]. \tag{3}$$

Where Y is a n×1 vector of the observations of the dependent variable.

1 is a n×1 vector of 1's

τ is a t×1 vector of the treatment effects

Δ is a t×n matrix known as the design matrix of the treatment, its elements Δ_{ij} is 1 when observation y_{ij} is from treatment i and 0 otherwise.

β is a b×1 vector of the block effects

D is a b×n matrix known as the design matrix of the block, it's elements d_{ij} is 1 when observation y_{ij} is from block j and 0 otherwise.

2.3   Prior Distributions for Unknown Parameters

Mathematically, the posterior is written as: posterior $\propto$ likelihood $\times$ prior. The parameters to be estimated are B and σ^2, as depicted in equation (2). Since there is no information beforehand about the range of values of the parameters in this study, the concept of uninformative prior shall be employed [8]. It has been shown that when uninformative prior is used, the data dominate the posterior. Therefore, the estimates obtained from the posterior will be similar to maximum likelihood estimate in frequentist approach [6]. The posterior distribution of the parameters is given as:

$$p(B, \sigma^2|Y) \; \alpha \; \pi(B, \sigma^2). f(Y|B, \sigma^2) \qquad\qquad (4)$$

Where $\pi(B, \sigma^2)$ is the prior information on the parameters and $f(Y|B, \sigma^2)$ is the likelihood.

$\pi(B, \sigma^2)$ can be further simplified into $\pi(B|\sigma^2). \pi(\sigma^2)$

Since the data is assumed to follow a normal distribution, the appropriate conjugate prior distributions of these parameters can be worked out [9]:

$$B|\sigma^2 \sim N(\tau, \gamma^2) \, and \, S = 1/\sigma^2 \sim \chi^2_v \qquad\qquad (5)$$

This implies that σ^2 has an inverse chi-square distribution.

2.4   Sampling from Posterior Distribution

The response variable follows a normal distribution; therefore the posterior distribution has a form of a normal-inverse chi-square distribution. Markov Chain Monte Carlo (MCMC) will be used to sample data points from the posterior distribution in order to summarize and estimate the parameters of interest (B). Numerous MCMC algorithms have been developed, each having very similar outputs but some are better in certain situations [4,8]. In this work the Gibbs sampler was employed.

The Gibbs sampler is an MCMC algorithm that generates samples from the posterior distribution. It generates samples for each parameter from the conditional distribution of the parameters because it is easier to derive conditional distribution from a joint distribution than deriving a marginal distribution from a joint distribution [9]. One major reason for adopting Gibbs sampler in this work is because the number of parameters to be estimated is fixed. This makes it feasible to derive the conditional distribution of the parameters. The sampling was implemented on R software using the package 'bayesm' [10]. Summaries of the samples drawn will be computed and inference will be made based on the 95% equal tails credible interval.

3.0   Results and Discussion

The farm and variety that gave the lowest yield are used as the reference categories for farm and variety respectively. The estimates of the parameters presented are:

$\tau'_i = \tau_i - \tau_l, \; i = 1, 2, 3, 4 \; (i \neq l)$ for the varieties; and

$\beta'_j = \beta_j - \beta_l, \; j = 1, 2, ..., 6 \; (j \neq l)$ for the farms.

Where "l" represents the variety/ farm with the lowest yield.

The results from the MCMC algorithm for each of the variables considered are presented in Tables 1, 2 and 3. Table 1 indicates that the mean height of each variety of the plants does not differ significantly from each other at 0.05 level of significance because zero (0) is included in the 95% credible intervals for the varieties. Also, the mean height in some farms differs significantly from other farms, which inform the inclusion of farms as blocks in the model.

Table 1: Posterior Distribution for the Analysis of Height of the Plants

		Estimate of the parameters	Standard error	95% credible interval	
				Lower	Upper
Variety	1	0.153	0.00081	-0.091	0.40
	2	0.023	0.00094	-0.226	0.27
	4	0.090	0.00089	-0.153	0.33
Farm	2	0.151	0.00108	-0.15	0.45
	3	0.551	0.00111	0.255	0.85
	4	0.515	0.00108	0.215	0.82
	5	0.430	0.00102	0.134	0.73
	6	0.217	0.00119	-0.082	0.52

The effect of maize varieties on the weight of cobs is presented in Table 2. It was observed that there is no significant difference in the mean weight of cobs between the varieties considered based on the 95% credible intervals. However, the mean weight of cobs differs significantly in some of the farms, which further justifies the use of farms as blocks. Table 3 shows that the mean weight of grains for all the four varieties in this study does not differ significantly from each other at 5% level of significance.

Table 2: Posterior Distribution for the Analysis of Weight of Cobs

		Estimate of the parameters	Standard error	95% credible interval	
				Lower	Upper
Variety	1	1.003	0.0057	-0.483	2.5
	2	0.343	0.0059	-1.149	1.8
	4	0.001	0.0056	-1.525	1.5
Farm	1	1.926	0.0072	0.071	3.7
	2	2.416	0.0075	0.552	4.2
	3	4.161	0.0073	2.305	6.0
	4	6.039	0.0071	4.191	7.8
	6	0.810	0.0069	-1.029	2.6

From the results presented, it can be inferred that Variety 3 (43DMRSR-White) yield the lowest value for all the variables considered in this experiment; therefore it serves as the reference category. The inclusion of farms in the model as an additional source of variation is justified because some of the credible intervals for the farms in all the tables does not include zero. Therefore, there is a significant difference in the effect of the farms for all the three variables considered as observed in Tables 1, 2 and 3. This result agrees with previous studies on the mean weight of 100 kernels of some varieties of maize [11].

Table 3: Posterior Distribution for the Analysis of Weight of Grains

		Estimate of the parameters	Standard error	95% credible interval	
				Lower	Upper
Variety	1	1.16	0.0048	-0.48	2.4
	2	0.50	0.0047	-0.71	1.7
	4	0.17	0.0043	-0.1	1.4
Farm	1	1.46	0.0056	0.00	2.9
	2	1.59	0.0057	0.15	3.0
	3	3.58	0.0056	2.1	5.0
	4	4.46	0.0057	3.0	5.9
	5	0.97	0.0061	-0.48	2.4

The International Institute of Tropical Agriculture (IITA) reported that three out of the four varieties compared in this experiment (DMLSR-White, DMRESR-Yellow and 43DMR-SR-White) were specifically bred to be resistant to Downy mildew disease [12, 13]. Therefore it is expected that these three varieties should produce similar outputs when planted in downy mildew infected areas. Although TZSR–White was bred to be resistant to streak [13], this work shows that it is equally resistant to downy mildew disease when its output was compared to those of varieties specifically bred to be resistant to the disease.

4.0   Conclusion

The severity of the downy mildew disease has a significant impact on the yield of maize as evident from the significant differences in the yields obtained from the six farms under study. Also, the four varieties of maize considered in this study does not exhibit any distinguishing characteristics in terms of height of the plants, weight of the cobs and weight of the grains when they are planted in areas affected by downy mildew disease. This indicates that the four varieties of maize investigated in the study are equally resistant to downy mildew disease. Farmers can therefore make decision on which of the varieties of maize to be planted based on availability and cost of acquiring the seed since the yield expected from the four varieties of maize are similar.

References

[1]   Gibson, L. and Benson, G. (2002). Origin, History and Uses of Corn. Department of Agronomy, Iowa State University. **(INCLUDE PUBLISHER, CITY).**
[2]   Roney, J. (2009).The Beginnings of Maize Agriculture.Archeology South West, Vol. 23, No. 1, pp. 4.
[3]   Bonde, M.R. (2008). Epidemiology of Downy Mildew Diseases of Maize, Sorghum and Pearl Millet. Tropical Pest Management, Vol. 1, No. 28, pp. 49-60.
[4]   Marin, J-M. and Robert, C. (2014). Bayesian Essentials with R, 2nd ed. New York, Springer.
[5]   Yahya, W.B., Olaniran, R.O. and Ige, S.O. (2014). On Bayesian Conjugate Normal Linear Regression and Ordinary Least Square Regression Methods: A Monte Carlo Study. Ilorin Journal of Science, Vol. 1, No. 1, pp. 216-231.
[6]   Cousineau, D. and Helie S. (2013). Improving maximum likelihood estimation using prior probabilities: A tutorial on maximum a posteriori estimation and an examination of the weibull distribution. Tutorials in Quantitative Methods for Psychology, Vol. 9, No. 2, pp. 61-71.
[7]   Mongomery, D.C. (1997). Design and Analysis of Experiments, 5th ed, New York, John Wiley and Sons Inc.
[8]   Hoff, P.D. (2009). A First Course in Bayesian Statistical Method, New York, Springer.
[9]   Koop, G., Poirier, D.J. and Tobias, J.L. (2007). Bayesian Econometric Methods, New York, Cambridge University Press.
[10]   Rossi, P.E., Allenby, G.M. and McColloch, R. (2005). Bayesian Statistics and Marketing, Chichester, John Wiley and Sons Ltd.

[11]   Iken, J.E., Amusa, N.A. and Obatolu, V.O. (2002). Nutrient Composition and Weight Evaluation of some Newly developed Maize Varieties in Nigeria. The Journal of Food Technology in Africa. Vol.7, No.1, pp. 27-29.

[12]   Adigun I.K. (1994). Report by Oyo State ADP on Campaign for Downy Mildew Disease Control, In: Archival Report of IITA: Proceedings of the Training Workshop for Downy Mildew Control, Akure. pp.8-9

[13]   Kling J.G. and Kim S.K. (1994). Evaluation and Multiplication of Maize Varieties, In: Archival Report of IITA: Proceedings of the Training Workshop for Downy Mildew Control, Akure, pp. 16-25.