

On the Performance of Bayesian Generalized Dissimilarity Model Estimator

Evellin Dewi Lusiana^{1,2,*}, Suci Astutik³, Nurjannah³, Abu Bakar Sambah²

¹Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, Malang, 65145, East Java, Indonesia

²Faculty of Fisheries and Marine Science, Universitas Brawijaya, Malang, 65145, East Java, Indonesia

³Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, Malang, 65145, East Java, Indonesia

Received October 20, 2022; Revised December 23, 2022; Accepted January 16, 2023

Cite This Paper in the Following Citation Styles

(a): [1] Evellin Dewi Lusiana, Suci Astutik, Nurjannah, Abu Bakar Sambah, "On the Performance of Bayesian Generalized Dissimilarity Model Estimator," *Mathematics and Statistics*, Vol. 11, No. 1, pp. 160 - 168, 2023. DOI: 10.13189/ms.2023.110118.

(b): Evellin Dewi Lusiana, Suci Astutik, Nurjannah, Abu Bakar Sambah (2023). On the Performance of Bayesian Generalized Dissimilarity Model Estimator. *Mathematics and Statistics*, 11(1), 160 - 168. DOI: 10.13189/ms.2023.110118.

Copyright©2023 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract The Generalized Dissimilarity Model (GDM) is an extension of Generalized Linear Model (GLM) that is used to describe and estimate biological pairwise dissimilarities following a binomial process in response to environmental gradients. Some improvement has been made to accommodate the uncertainty quantity of GDM by applying resampling scheme such as Bayesian Bootstrap (BBGDM). Because there is an ecological assumption in the GDM, it is reasonable to use a proper Bayesian approach rather than resampling method to obtain better modelling and inference results. Similar to other GLM techniques, the GDM also employs a link function, such as the logit link function that is commonly used for the binomial regression model. By using this link, a Bayesian approach to GDM framework which called Bayesian GDM (BGDM) can be constructed. In this paper, we aim to evaluate the estimators' performance of Bayesian Generalized Dissimilarity Model (BGDM) in relative to BBGDM. Our study revealed that the performance of BGDM estimator outperformed that of BBGDM, especially in term of unbiasedness and efficiency. However, the BGDM estimators failed to meet consistency property. Moreover, the application of the BGDM to a real case study indicates that its inferential abilities are superior to the preceding model.

Keywords GLM, Logit Link Function, Pair-site Dissimilarity, Parameter Estimation

1. Introduction

The Generalized Dissimilarity Model (GDM) is a powerful and innovative method for describing and estimating beta diversity, i.e., shifts in biodiversity over regions, periods, and environmental changes [1,2]. It is an extension of the Generalized Linear Model (GLM) that refines the weaknesses of previous methods used to investigate diversity, such as the Mantel test, linear matrix regression, and non-metric multidimensional scaling (NMDS), all of which failed to accommodate the nonlinearities discovered in pairwise dissimilarities analyses [3]. The fitted dissimilarity function, or beta diversity function, must be increasingly monotonic within the GDM framework. It is a realization of the ecological assumption that compositional dissimilarity can only be exacerbated by increasing the distance between sites along an environmental gradient [4]. In addition, I-spline functions are utilized to better capture the nonlinearities of beta diversity [1].

Similar to other GLM techniques, the GDM also employs a link function that establishes the relationship between the expected compositional pair-site dissimilarity and a scaled combination of inter-location distances based on several geographical or environmental factors [2]. Furthermore, it was acknowledged that the pair-site

dissimilarity in the GDM follows a binomial process, preferred to use a logit link function because it is frequently applied to binomial GLM, especially in the logistic regression model [5].

In the classic GDM, the permutation test used in the model only allows for testing hypotheses and does not enable the characterization of the unpredictability in the predictions of dissimilarities [1]. Contrary to the earlier use of permutation methods to test for variable significance, a more general approach to quantifying unpredictability in statistical inference is necessary. It is especially crucial to measure unpredictability in dissimilarity estimations that are used to interpret the model, such as interval estimation [5]. Interval estimation is one of the most frequently employed techniques in statistics, as it provides a range of credible values for a parameter after accounting for uncertainty [6]. While interval estimation is not available in the classic GDM, the Bayesian Bootstrap GDM (BBGDM) proposed by Woolley et al. [5] enable the estimation of confidence intervals. This method integrates resampling technique which is Bayesian Bootstrap in the GDM framework. Therefore, BBGDM does not apply Bayesian approach that involves prior, likelihood, and posterior distribution for its parameter estimation.

On the other hand, Bayesian analysis has several advantages over the classical method, especially when there is prior knowledge regarding the model coefficients [7]. Because the GDM framework requires an underlying ecological assumption [1], it is reasonable to apply the Bayesian approach to this model. The monotonic fitted function requirement for the GDM can be fulfilled by using a non-negative prior distribution. Bayesian Generalized Dissimilarity Model (BGDM) demonstrated in this study is the first research that applies Bayesian technique for GDM parameter estimation. Specifically, we aim to evaluate the estimators' performance of BGDM through simulation and empirically study.

2. Materials and Methods

2.1. Generalized Dissimilarity Model (GDM)

As part of GLM model, GDM employs a link function, μ_{uv} , that can be expressed as follows (1)

$$\mu_{uv} = 1 - e^{-\eta_{uv}} \tag{1}$$

where μ is the expected dissimilarity between points and η_{uv} is linear function of the environmental dissimilarity across point u and v . The function of η_{uv} can be written as equation (2)

$$\eta_{uv} = \beta_0 + \sum_{j=1}^p \sum_{l=1}^{k_j} \beta_{jl} |I_{jl}(x_{ju}) - I_{jl}(x_{jv})| \tag{2}$$

with restriction $\beta_{jl} \geq 0$

where: $I_{jl}(x_j)$ denoted the l -th basis I-spline function of x_j and β_{jl} is coefficient of the l -th basis I-spline function

of x_j , while x_j is the predictor or environmental variable.

The estimation of GDM parameters is carried out like other GLMs models by using the maximum likelihood estimation (MLE) method and Iteratively Re-weighted Least Square (IRLS) algorithm. However, to ensure that the regression coefficient is positive ($\beta_{jl} \geq 0$), modifications were made by replacing least-square regression in IRLS with non-negative least squares (NNLS) [1].

2.2. Bayesian Bootstrap Generalized Dissimilarity Model (BBGDM)

Inference and predictive uncertainty are difficult to characterize using GDM, thus reducing the usefulness of GDM in ecological risk assessment and conservation determination. GDM uses a permutation test approach to assessing the significance of the predictor variables. This permutation test overcomes the data-dependence problem (because dissimilarity is calculated from a smaller number of observations), but does not provide a quantification of predictive uncertainty. Hence, the Bayesian bootstrap approach is used so that uncertainty in observations can be accommodated in the entire analysis process, including for predictions [5].

The integration of Bayesian bootstrap into the GDM parameter estimation is done by adding weights to the biological dissimilarity based on the number of location points. This is done by generating as many as B bootstrap weighting vectors, $\{g_b\}_{b=1}^B$, where each of weighting vector g_b is random variable generated from Dirichlet distribution with a size $n \times 1$ (n =number of points) [5]. Next, transform the weighting vector for that location into a weight for biological dissimilarity by taking the upper triangle of $g_b g_b'$. Then, GDM modeling is carried out using these weight vectors to get the estimated parameter values and repeated for as many as B bootstrap weights. The parameter estimator obtained is the empirical average of the estimator sampling distribution of each repetition [5].

The response variable of GDM measuring pairwise biological dissimilarities (y_{uv}) can be defined as the number of species that are not shared between sites u and v . This event is represented by a binomial process that can be expressed as follows (3) [5].

$$y_{uv} \sim Bin(ns_{uv}, \pi_{uv}), \tag{3}$$

where ns_{uv} is the total number of species combined at sites u and v , y_{uv} is the number of species that are not shared between sites u and v , and π_{uv} is the expected dissimilarity between sites u and v , i.e., a link function. Here, we used a logit link function. The logit function also maps linear predictors into the range of 0 and 1, which is an appropriate value interval for Beta diversity indices such as the Jaccard index and the Bray-Curtis index [5]. Furthermore, the Bayesian bootstrap estimate of $\hat{\mu}_{uv}$ for the logistic model can be easily expressed as the following

equation (4)

$$\hat{\mu}_{uv} = \frac{1}{B} \sum_{b=1}^B \frac{1}{1 + \exp(-\eta_{uv}^{(b)})} \quad (4)$$

where: $\eta^{(b)}$ is a linear function of the environmental variables dissimilarity as expressed in equation (2) in the b^{th} bootstrap.

2.3. Bayesian Generalized Dissimilarity Model (BGDM)

By using the approach of logit function used for BBGDM, let $\tilde{x}_{uv,jl} = |I_{jl}(x_{ju}) - I_{jl}(x_{jv})|$, then equation (2) can be written as equation (5):

$$\eta_{uv} = \beta_0 + \sum_{j=1}^p \sum_{l=1}^{k_j} \beta_{jl} \tilde{x}_{uv,jl} \quad (5)$$

Suppose that there are n pair-sites, where $i = 1, 2, \dots, n$. Suppose also that $\beta^t = (\beta_0, \beta_{11}, \dots, \beta_{pk_j})$ and $\sum_{j=1}^p k_j = m$, then a vector for the GDM coefficient can be defined as $\tilde{\beta} = \{\beta_k: k = 0, 1, \dots, m\}$, which has elements corresponding to β . Hence, equation (5) can be more simply written as equation (6):

$$\eta = \tilde{x}_i \tilde{\beta} \quad (6)$$

As a result, the appropriate logit link function for the GDM can be expressed as [8]

$$\pi(\tilde{x}_i) = \frac{\exp(\tilde{x}_i \tilde{\beta})}{1 + \exp(\tilde{x}_i \tilde{\beta})} \quad (7)$$

The GDM likelihood function for a logit link function is defined as follows (8):

$$\begin{aligned} l(\tilde{X}, \tilde{\beta}) &= f(\mathbf{y} | \tilde{\beta}) \\ &= \prod_{i=1}^n \binom{ns}{y_i} \pi(\tilde{x}_i)^{y_i} [1 - \pi(\tilde{x}_i)]^{ns-y_i} \\ &= \prod_{i=1}^n \binom{ns}{y_i} \left(\frac{\exp(\tilde{x}_i \tilde{\beta})}{1 + \exp(\tilde{x}_i \tilde{\beta})} \right)^{y_i} \left(\frac{1}{1 + \exp(\tilde{x}_i \tilde{\beta})} \right)^{ns-y_i} \end{aligned}$$

Bayesian framework for GDM (BGDM) parameter estimation is constructed by involving three components, which are prior distribution, likelihood function, and posterior distribution. In the case of GDM parameter estimation, all regression coefficients ($\tilde{\beta}$), except the intercept, must be non-negative [1]. Therefore, it is proposed to use truncated distribution as the prior ($p(\mathbf{y})$) such as truncated normal and truncated Student-t. Meanwhile, equation (8) is used as the likelihood function.

Finally, the posterior distribution of $\tilde{\beta}$ is defined as:

$$p(\tilde{\beta} | \mathbf{y}) = \frac{f(\mathbf{y} | \tilde{\beta}) p(\tilde{\beta})}{p(\mathbf{y})}, \quad (9)$$

where $f(\mathbf{y} | \tilde{\beta})$ is the likelihood function of the GDM, $p(\tilde{\beta})$ is the prior distribution, and the constant $p(\mathbf{y})$ is restricted as follows:

$$0 < p(\mathbf{y}) = \int_{\mathbb{R}^m} p(\mathbf{y} | \tilde{\beta}) p(\tilde{\beta}) d\tilde{\beta} < \int_{\mathbb{R}^m} p(\tilde{\beta}) d\tilde{\beta} = 1 \quad (10)$$

Furthermore, the mean posterior is formulated in equation (11) below:

$$E(\tilde{\beta} | \mathbf{y}) = \int_{\mathbb{R}^m} \tilde{\beta} p(\tilde{\beta} | \mathbf{y}) d\tilde{\beta} \quad (11)$$

The mean posterior for coefficient β_k ($k \neq 0$) can be written as in equation (11), while the mean posterior for β_0 is expressed in equation (12) below:

$$E(\beta_k | \mathbf{y}) = \int_0^\infty \beta_k p(\beta_k | \mathbf{y}) d\beta_k \quad (12)$$

The integrations in equations (12) are highly complex, so the Markov Chain Monte Carlo (MCMC) technique must be used to perform them [9]. In this research, we adopted the Hamiltonian Monte Carlo (HMC) algorithm implemented in the RStan Package [10].

2.4. Simulation Study Scenario

In this research, we conducted a simulation study to compare the performance of BGDM and BBGDM parameter estimates. We generated a dataset with varied samples ranging from 50 to 5000 for two predictors, namely, X_0 (intercept) and X_1 . The continuous predictor X_1 is generated from the standard normal distribution. Meanwhile, the response variable Y is generated from the binomial distribution. Then the data are fitted to BBGDM and BGDM model. In specific, BGDM was fitted using 1000 iteration and two chains. Moreover, the resulting posterior mean will be evaluated with a trace plot [11]. We set all parameter values to 1. The performance of parameter estimates, in term of unbiasedness and consistency properties, is evaluated by using Root Mean Square Error (RMSE) criteria, in which estimator with lower RMSE had better performance [12]. The empirical SE is a measure of the precision or efficiency of the estimator [13].

2.5. Case Study

We perform an empirical study using the Dune dataset [14] which is commonly used for studying the ordination method [15,16]. The dataset consists of 30 species abundant at 20 sites. The dataset extracted from R package `vegan::dune`. In addition, it also contains environmental variables, such as soil thickness. We fit the BGDM for pair-site biological dissimilarity, in this case, the Bray-Curtis dissimilarity [17], using soil thickness as the predictor. Then we compare the results with BBGDM [5].

3. Results

3.1. Simulation Study Results

The trace plot evaluation of the BGDM parameter estimation for the data simulation under the normal prior as depicted in Figure 1 produces random scatterings around the mean values for the posterior samples. Note that the chains merge well, making the trace plots look good. As a result, it can be said that the posterior mean of the BGDM does exist.

Figure 2 plots the estimators of BBGDM and BGDM that fitted in simulated data with three spline transformed predictors, with means displayed as black lines. It is clear that for β_1 , mean of BBGDM and BGDM estimator are quite far from the true value $\beta_1 = 1$, which indicates some biases. This pattern is also applied for other parameter estimates fitted with BBGDM. In contrast, BGDM estimator for β_2 and β_3 is close to the true values. Hence, unbiasedness of BBGDM estimators is better than BGDM.

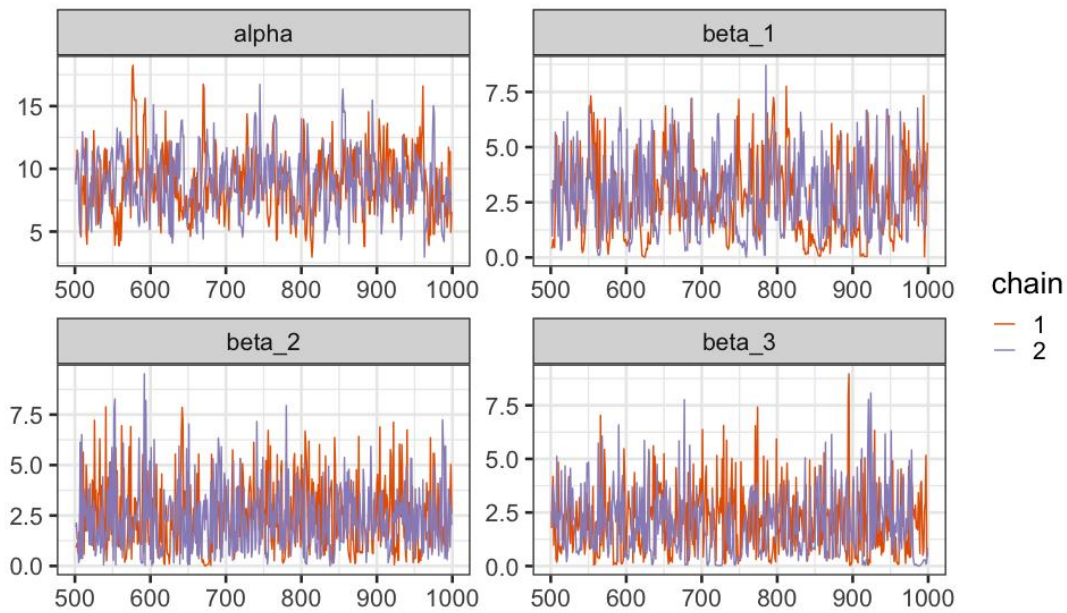
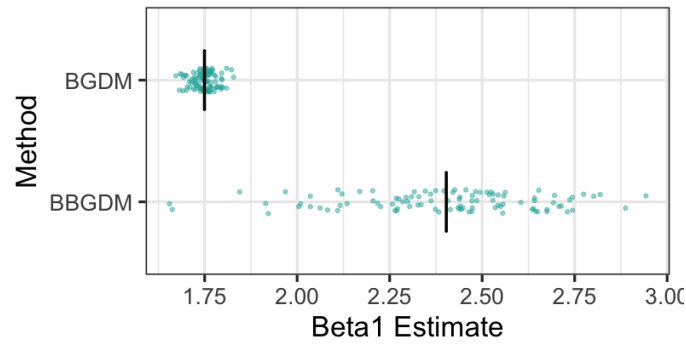
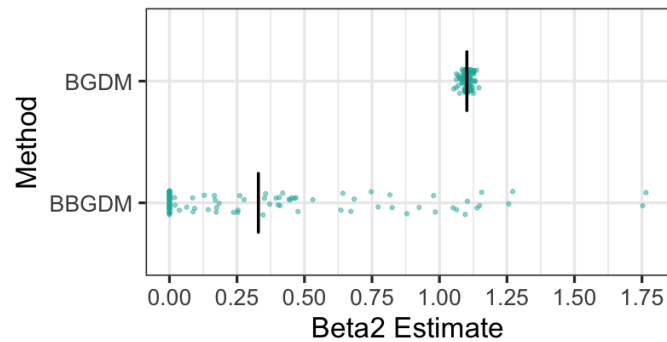


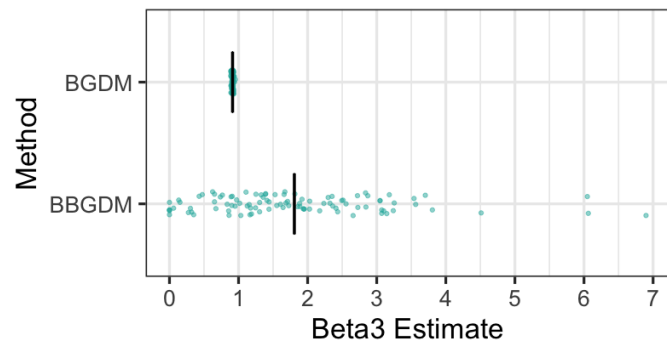
Figure 1. Trace plot of BGDM parameter estimation for data simulation



(a)



(b)

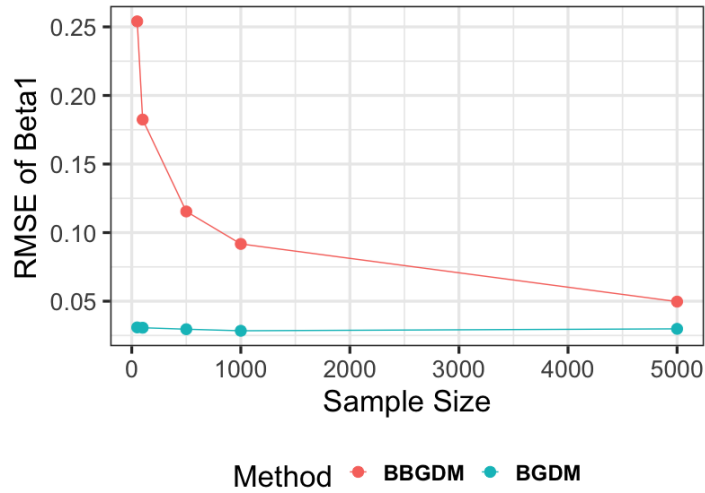


(c)

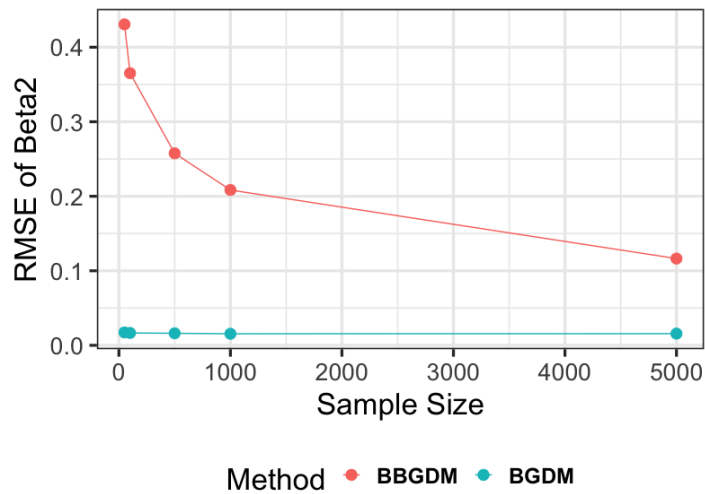
Figure 2. Estimator unbiasedness of BBGDM and BGDM ($n=50$) for spline transformed predictor (a) First spline; (b) Second spline; (c) Third spline

Investigation on parameter estimates performance at various sample sizes of BBGDM and BGDM is presented in Figure 3. The RMSEs of BBGDM estimators are decreased along with the increasing number of sample size. Meanwhile, BGDM estimators have stable RMSE throughout the analysis. Finally, it can be seen that BGDM estimators outperformed those of BBGDM as their RMSEs are lower at all sample sizes. In term of consistency property, BBGDM method has consistent estimators which are depicted by their RMSE value that tends to decrease as

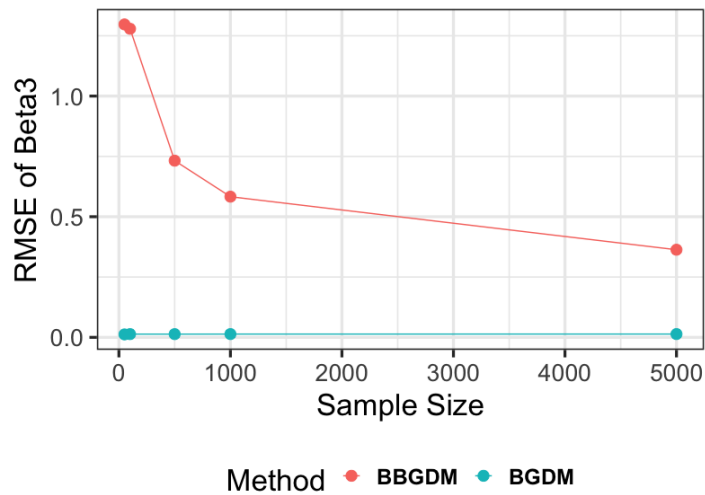
the increasing number of sample size. However, taking a closer look to the RMSE of BGDM estimators (Figure 4), the RMSEs do not subside as the sample size enlarged. As a result, BBGDM estimators seem not fulfill the consistency characteristics. Inconsistency of Bayesian estimation can be occurred due to misspecification in simple linear regression model [18]. Meanwhile, the use of Hamiltonian Monte Carlo (HMC) is also exhibited to affect the consistency of posterior inference [19].



(a)



(b)



(c)

Figure 3. RMSE comparison of parameter estimates between BBGDM and BGDM for spline transformed predictor (a) First spline; (b) Second spline; (c) Third spline

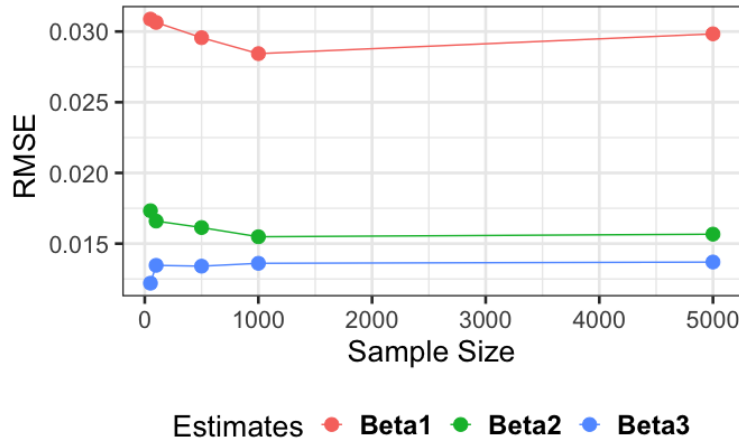


Figure 4. RMSE of BGDM estimators for consistency property

Table 1. Empirical standard error of parameter estimates

Sample Size (n)	BBGDM			BGDM		
	β_1	β_2	β_3	β_1	β_2	β_3
50	0.2540	0.4307	1.2970	0.0309	0.0173	0.0122
100	0.1824	0.3651	1.2789	0.0306	0.0166	0.0135
500	0.1154	0.2576	0.7323	0.0296	0.0161	0.0134
1000	0.0918	0.2085	0.5831	0.0284	0.0155	0.0136
5000	0.0497	0.1164	0.3632	0.0298	0.0157	0.0137

Table 1 presents the empirical standard error of BBGDM and BGDM estimators. It can be seen that the standard errors of BGDM estimators are all smaller than those of BBGDM. Therefore, estimation of BGDM is more efficient.

3.2. Case Study Results

Table 2. Comparison of GDM, BBGDM, and BGDM estimates for Dune data

Parameter	BBGDM		BGDM	
	Estimates	Confidence Interval 95%	Estimates	Credible Interval 95%
Intercept	0.62	[0.25, 0.94]	0.65	[0.52, 0.77]
β_1	0.00	[0.00, 0.65]	0.07	[0.00, 0.23]
β_2	0.17	[0.00, 1.1]	0.28	[0.01, 0.65]
β_3	0.45	[0.00, 1.61]	0.57	[0.07, 1.11]

The estimation results for the BBGDM and BGDM are presented in Table 2. It is clear that the intercept estimates for these three methods are similar. The same goes for β_2 . On the other hand, the estimates for β_1 and β_3 are quite

different.

4. Discussions

The result of our simulation study was regarding the unbiasedness, consistency, and accuracy or efficiency of BGDM estimators. Unbiasedness is one of the important properties in statistical modelling and parameter estimation technique. It denotes if the expected value of an estimator's sampling distribution is equal to the unknowable actual value of the population parameter [20]. Moreover, identifying that an estimator is unbiased on mean is not very useful if the values of estimators fluctuate between samples. Hence, estimators are chosen not only for their unbiasedness, but also for how widely their estimates vary from sample to sample [21]. RMSE is a metric that commonly used to evaluate bias and estimation accuracy [21,22]. Our result suggested that BGDM estimator is more accurate and has minimum standard error. An unbiased estimator which has minimum standard error than other is called as efficient estimator [21]. Therefore, BGDM estimator is relatively efficient to BBGDM estimator. However, as shown in Figure 2, the BGDM estimators were not consistent. If the estimates of the parameter are inconsistent, this severely weakens any conclusions that may be drawn from the study [23].

The case study analysis involves two parameter estimation techniques of GDM that are BBGDM, and BGDM. Interval estimation is one of the most frequently employed techniques in statistics, as it provides a range of credible values for a parameter after accounting for uncertainty [6]. While interval estimation is not available in the classic GDM, the BBGDM proposed by Woolley et al. [5] and the BGDM developed in this study enable the estimation of confidence intervals and credible intervals, respectively. In Table 1, the interval estimations produced by these models indicate the uncertainty predictions regarding biological dissimilarity in the Dune dataset. Moreover, they can also be used for significance testing, where an interval estimate for a model parameter not containing zero suggests the significance of that parameter, i.e., the significant effect of the predictor variable corresponding to the model parameter on the response variable [24]. Therefore, based on Table 1, it can be said that in the BBGDM, soil thickness does not significantly affect biological dissimilarity. However, the BGDM's credible interval shows otherwise.

Although the frequentist confidence interval is the most widely used interval estimate, its limitations and practical issues are well recognized. First of all, confidence intervals contradict likelihood's basic concept [25,26]. Second, the relationship between the width of the interval and the estimate's precision is occasionally inverted, demonstrating that the width of a confidence interval does not, in general, indicate the accuracy of our knowledge about the parameter [27]. Third, statements about the probability of the parameter being within a given range cannot be made using confidence intervals [28]; they only permit claims about the random quantity (interval) surrounding the actual, fixed parameter. Because a parameter is not a random variable according to the frequentist philosophy, probability statements about the parameter being in a specific range, which are frequently desired in practice, are not possible. Some of the problems mentioned above are solved by Bayesian credible intervals, including compliance with the likelihood principle [26,29] and forming straightforward probabilistic statements regarding the desired parameter [30]. Because credible intervals also link the width of the interval to the accuracy of the estimate obtained, they are frequently promoted as alternatives to standard frequentist confidence intervals [31,32]. It also has been shown that Bayesian credible intervals are more precise than frequentist confidence intervals [7].

5. Conclusions

Parameter estimation for the GDM can be accomplished via a Bayesian technique, as in other GLM methods which called BGDM. Because the response variable for the GDM can be described as a binomial process, the use of logit link function is appropriate. Our simulation study results

suggest that the BGDM estimator exhibits a non-zero bias but smaller than the BBGDM estimator. Moreover, the performance of BGDM estimator is surpassed that of BBGDM based on RMSE criteria as well as more efficient. However, these estimators are inconsistent. The fitting of the BGDM to the Dune dataset indicates that this method is able to produce reliable interval estimation quantifying the uncertainty in the statistical inference.

REFERENCES

- [1] S. Ferrier, G. Manion, J. Elith, K. Richardson, "Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment", *Diversity and Distributions*, vol. 13, no. 3, pp. 252–264, 2007. DOI: 10.1111/j.1472-4642.2007.00341.x.
- [2] K. Mokany, C. Ware, S. N. C. Woolley, S. Ferrier, M.C. Fitzpatrick, "A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment", *Global Ecology and Biogeography*, vol. 31, no. 4, pp. 802–821, 2022. DOI: 10.1111/geb.13459.
- [3] M.C. Fitzpatrick, N.J. Sanders, S. Normand, J.-C. Svenning, S. Ferrier, A.D. Gove, R.R. Dunn, "Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients", *Proceedings of the Royal Society B: Biological Sciences*, vol. 280, no. 1768, pp. 20131201, 2013. DOI: 10.1098/rspb.2013.1201.
- [4] F. B. R. Lasram, T. Hattab, G. Halouani, M. S. Romdhane, F. le Loc'h, "Modeling of beta diversity in Tunisian waters: Predictions using generalized dissimilarity modeling and bioregionalisation using fuzzy clustering", *PLoS ONE*, vol. 10, no. 7, pp. 1–16, 2015. DOI: 10.1371/journal.pone.0131728.
- [5] S. N. C. Woolley, S.D. Foster, T. D. O'Hara, B.A. Wintle, P.K. Dunstan, "Characterising uncertainty in generalised dissimilarity models", *Methods in Ecology and Evolution*, vol. 8, no. 8, pp. 985–995, 2017. DOI: 10.1111/2041-210X.12710.
- [6] L. Hespanhol, C. S. Vallio, L. M. Costa, B. T. Saragiotto, "Understanding and interpreting confidence and credible intervals around effect estimates", *Brazilian Journal of Physical Therapy*, vol. 23, no. 4, pp. 290–301, 2019. DOI: 10.1016/j.bjpt.2018.12.006.
- [7] K. Gray, B. Hampton, T. Silveti-Falls, A. McConnell, C. Bausell, "Comparison of bayesian credible intervals to frequentist confidence intervals", *Journal of Modern Applied Statistical Methods*, vol. 14, no. 1, pp. 43–52, 2015. DOI: 10.22237/jmasm/1430453220.
- [8] A. Agresti, *Categorical Data Analysis*, John Wiley & Sons, New York, 2002.
- [9] C. Geyer, "Introduction to MCMC methods", *Handbook of Markov Chain Monte Carlo*, no. March, pp. C, 2011.
- [10] S. Thomas, W. Tu, "Learning Hamiltonian Monte Carlo in R", *The American Statistician*, vol. 75, no. 4, pp. 403–413, 2021. DOI: 10.1080/00031305.2020.1865198.

- [11] X. Fernández-i-Marín, "Ggmcmc: Analysis of MCMC samples and Bayesian inference", *Journal of Statistical Software*, vol. 70, 2016. DOI: 10.18637/jss.v070.i09.
- [12] T. Chai, R.R. Draxler, "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): Arguments Against Avoiding RMSE in The Literature", *Geoscientific Model Development*, vol. 7, pp. 1247–1250, 2014.
- [13] T. P. Morris, I. R. White, M. J. Crowther, "Using simulation studies to evaluate statistical methods", *Statistics in Medicine*, vol. 38, no. 11, pp. 2074–2102, 2019. DOI: 10.1002/sim.8086.
- [14] R. H. G. Jongman, C. J. F. ter Braak, O. F. R. van Tongeren, *Data Analysis in Community and Landscape Ecology*, 1995.
- [15] A. Hoegh, D. W. Roberts, "Evaluating and presenting uncertainty in model-based unconstrained ordination", *Ecology and Evolution*, vol. 10, no. 1, pp. 59–69, 2020. DOI: 10.1002/ece3.5752.
- [16] B. van der Veen, F. K. C. Hui, K. A. Hovstad, E. B. Solbu, R. B. O'Hara, "Model-based ordination for species with unequal niche widths", *Methods in Ecology and Evolution*, vol. 12, no. 7, pp. 1288–1300, 2021. DOI: <https://doi.org/10.1111/2041-210X.13595>.
- [17] C. Ricotta, "Of beta diversity, variance, evenness, and dissimilarity", *Ecology and Evolution*, vol. 7, no. 13, pp. 4835–4843, 2017. DOI: 10.1002/ece3.2980.
- [18] M. Ekström, M. Nilsson, "A Comparison of Model-Assisted Estimators, With and Without Data-Driven Transformations of Auxiliary Variables, With Application to Forest Inventory", *Frontiers in Forests and Global Change*, vol. 4, 2021. DOI: 10.3389/ffgc.2021.764495.
- [19] P. Grünwald, T. van Ommen, "Inconsistency of Bayesian Inference for Misspecified Linear Models, and a Proposal for Repairing It", 2014. <http://arxiv.org/abs/1412.3730>.
- [20] M.E. Ezell, K. C. Land, Ordinary Least Squares (OLS), in: K.B.T.-E. of S.M. Kempf-Leonard (Ed.), Elsevier, New York, 2005: pp. 943–950. DOI: <https://doi.org/10.1016/B0-12-369398-5/00171-7>.
- [21] J. Shabbir, R. O. Id, "Use of an efficient unbiased estimator for finite population mean", pp. 1–13, 2022. DOI: 10.1371/journal.pone.0270277.
- [22] M. Harwell, "A strategy for using bias and RMSE as outcomes in Monte Carlo Studies in statistics", *Journal of Modern Applied Statistical Methods*, vol. 17, no. 2, 2018. DOI: 10.22237/jmasm/1551907966.
- [23] D. E. Giles, "On the Inconsistency of Instrumental Variables Estimators for the Coefficients of Certain Dummy Variables", *Journal of Quantitative Economics*, vol. 15, no. 1, pp. 15–26, 2017. DOI: 10.1007/s40953-016-0042-7.
- [24] T. R. Knapp, "Significance Test, Confidence Interval, Both or Neither?", *Clinical Nursing Research*, vol. 26, no. 3, pp. 259–265, 2017. DOI: 10.1177/1054773817708652.
- [25] R. Kelter, "The evidence interval and the Bayesian evidence value: On a unified theory for Bayesian hypothesis testing and interval estimation", *British Journal of Mathematical and Statistical Psychology*, vol. 75, no. 3, 2022. DOI: <https://doi.org/10.1111/bmsp.12267>.
- [26] J. O. Berger, R. L. Wolpert, *The likelihood principle : a review, generalizations, and statistical implications*, Institute of Mathematical Statistics, 1988.
- [27] R. D. Morey, J.-W. Romeijn, J. N. Rouder, "The philosophy of Bayes factors and the quantification of statistical evidence", *Journal of Mathematical Psychology*, vol. 72, pp. 6–18, 2016. DOI: <https://doi.org/10.1016/j.jmp.2015.11.001>.
- [28] J. K. Kruschke, T. M. Liddell, "The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective", *Psychonomic Bulletin & Review*, vol. 25, no. 1, pp. 178–206, 2018. DOI: 10.3758/s13423-016-1221-4.
- [29] J. N. Rouder, "Optional stopping: No problem for Bayesians", *Psychonomic Bulletin & Review*, vol. 21, no. 2, pp. 301–308, 2014. DOI: 10.3758/s13423-014-0595-4.
- [30] E.-J. Wagenmakers, R.D. Morey, M.D. Lee, "Bayesian Benefits for the Pragmatic Researcher", *Current Directions in Psychological Science*, vol. 25, no. 3, pp. 169–176, 2016. DOI: 10.1177/09637214166643289.
- [31] R. McElreath, *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, CRC Press, 2016.
- [32] J. K. Kruschke, "Rejecting or Accepting Parameter Values in Bayesian Estimation", *Advances in Methods and Practices in Psychological Science*, vol. 1, no. 2, pp. 270–280, 2018. DOI: 10.1177/2515245918771304.