ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.9, No.11, 2019

Application Of A Modified G-Parameter Prior ($g = 1/n^5$) In Bayesian Model Averaging To CO_2 Emissions In Nigeria

Akanbi Olawale Basheer (PhD) [1] Oladoja Oladapo Muyiwa[2*]

[1]Department of Statistics, University of Ibadan, Nigeria.

[2]Department of Mathematics and Statistics, First Technical University, Ibadan, Nigeria.

oladapo.oladoja@tech-u.edu.ng

Abstract

Bayesian Model Averaging (BMA) is a variable selection approach that takes model uncertainty into account by averaging over the weights by their posterior model probabilities. Of concern are the priors to be used for the quantities of interest and the model choice. Using uniform prior for the model choice as supported by the literature, this study elicits a modified g- parameter prior in BMA in a normal linear regression models to model CO_2 emissions in Nigeria. The functional form, $\frac{\phi_1(s_j)}{\phi_2(n)}$ of the g-prior, $g = \frac{1}{n^5}$ was elicited by taking the number of regressors constant aimed at capturing information for increasing sample size. The modified prior show consistency with models in literature. Five sub samples (50, 100, 1000, 10,000 and 100,000) generated from normal distribution each replicated 100 times were used to investigate the predictive performance and sensitivity of the modified prior. The asymptotic properties of the prior were derived. The best model for CO_2 emissions in Nigeria with 53% Posterior Model Probability (PMP) involves the Industrial, Residential & Commercial and Agricultural sector. And among these, the industrial sector contributed most with 99.98% Posterior Inclusion Probability (PIP).

Applying the modified g-prior with LPS 2.3260 from the prediction which is the close to the threshold of 2.335, industrial sector, residential & commercial sector and agricultural sector are best used for modeling CO_2 emissions in Nigeria.

Keywords: Variable Selection, Industrial Sector, Posterior Inclusion Probabilities.

1. Introduction

Carbon dioxide (CO_2), a greenhouse gas (GHG) is a gas essential for life - animals exhale it through respiration and plants sequester it to conduct photosynthesis. It exists in Earth's atmosphere in comparably small concentrations, but it is vital for sustaining life. It is a gas that uses up and discharges heat energy, forming the 'greenhouse effect'. Including other greenhouse gases on the earth surface, like nitrous oxide and methane, CO_2 is useful in sustaining a habitable temperature for the planet, that is, if there were absolutely no GHGs, our planet would simply be too cold. In actual sense, carbon dioxide appears on the earth surface through burning fossil fuels (coal, natural gas, and oil), solid waste, trees and wood products, formed from the remains of plants and animals, and also as a result of certain chemical reactions (like manufacture of cement). Carbon dioxide, being used up by plants as part of the biological carbon cycle during photosynthesis is removed from the atmosphere through this process. While CO_2 emissions come from different natural sources, various forms of human activities that are emissions related are causal factors for the rise that has occurred on the planet since the advent of technology and industry. There exists both natural and human sources of carbon dioxide emissions. Natural sources include decomposition, ocean release and respiration. Human sources come from activities like cement production, deforestation as well as the burning of fossil fuels like coal, oil and natural gas. Due to variety of human actions, the atmospheric concentration of carbon dioxide on the planet surface has been on the rise in large amounts since the advent of technology and industry and has now achieved harmful levels not seen in the last 30 thousand centuries. Human sources of carbon dioxide emissions are much smaller than natural emissions but they have disrupt the natural balance that existed for many thousands of years before the influence of humans.

This study elicits a new modified g-parameter prior to examine uncertainties in the phase of big data analytics when there is rapid increase in sample size and applied it to model CO_2 emissions in Nigeria.

2. Bayesian Model Averaging

Bayesian Model Averaging (BMA) is an empirical tool to deal with model uncertainty in various milieus of applied science. In general, BMA is employed when there exist a variety of models which may all be statistically reasonable but most likely result in different conclusions about the key questions of interest to the researcher. BMA assigns probabilities on the model space and deals with model uncertainty by mixing over models, using the posterior model probabilities as weights. Typically, though not always, BMA focuses on which predictors to include in the analysis. The implementation of BMA, which was first proposed by Leamer (1978), for linear regression models is as follows. Consider a linear regression model with a constant term, θ_0 and s explanatory variables $x_1, x_2, \ldots, x_s$,

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_s x_s + \varepsilon \quad \ldots\ldots \tag{1}$$

where y denotes the observed data on the dependent variable. ε is the error term which is independent and identically distributed as $N(0, \sigma^2_{\varepsilon})$. In matrix form, it can be best written as

$$y = \Theta' X + \varepsilon \quad \ldots\ldots. \tag{2}$$

Given the number of regressors, we will have 2^s different combinations of right hand side variables indexed by Z_j for j = 1, 2, 3 … 2^s. Once the model space has been constructed, the posterior distribution for any coefficient of interest, say θ_h , given the data D is

$$P_r(\theta_h \mid D) = \sum_{j:\theta_h \in Z_j} P_r(\theta_h \mid Z_j) P_r(Z_j \mid D) \quad \ldots\ldots.. \tag{3}$$

BMA uses each model's posterior probability, $P_r(Z_j|D)$, as weights. The posterior model probability of Z_j is the ratio of its marginal likelihood to the sum of marginal likelihoods over the entire model space and is given by

$$P_r\left(Z_j \mid D\right) = P_r\left(D \mid Z_j\right)\frac{P_r(Z_j)}{P_r(D)} = P_r\left(D \mid Z_r\right)\frac{P_r(Z_j)}{\sum_{i=1}^{2^s} P_r(D \mid Z_i) P(Z_i)} \quad \ldots\ldots\ldots.. \tag{4}$$

Where,

$$P_r\left(D \mid Z_j\right) = \int P_r(D \mid \theta_j, Z_j) P_r\left(\theta_j \mid Z_j\right) d\theta_j \ldots\ldots \tag{5}$$

and θ_j is the vector of parameters from the model Z_j, $P_r(\theta_j|Z_j)$ is a prior probability assigned to the parameters of model Z_j, and $P_r(Z_j)$ is the prior probability that Z_j is the true model. It indicates how probable the researcher thinks model Z_j is before looking at the data. $P_r(D|Z_j)$ reflects the probability of the data given the model Z_j.

2.1. The Modified Parameter Prior

Prior elicitation can be extremely critical for the outcome of BMA analyses. The g - parameter prior elicited must be positive and should be as small as possible so that data information can spread across the models (i.e. 0 < g < 1). Literature reveals that some g - classes in existence can still be improved upon. Thus, we elicited a modified g - parameter prior which explained information of the prior as contained in a single observation and approximated by the SIC, BIC or HQ. The modified prior elicited for the rth term of the sample size is

$$g = \frac{1}{n^r}, r \geq 3 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad .(6)$$

The rationale behind the modification is for capturing information for fast increasing sample size. It has a mean of zero and always approximate to SIC or BIC. Since g is neither greater than one nor less than zero especially for a large sample size (n). Using FLS (2001a), we are of the assumption that

$$g = \frac{\phi_1(s_j)}{\phi(n)} \quad \text{with}$$

$$\lim_{n\to\infty} \phi_2(n) = \infty \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad \ldots (7)$$

Now, the assumption that there is a true model Z_s in Z generates the data under the following conditions.

$$\lim_{n\to\infty} \frac{\phi_2^{'}(n)}{\phi(n)} = 0 \ldots\ldots\ldots\ldots\ldots \quad \ldots\ldots (8)$$

together with either

$$\lim_{n\to\infty} \frac{n}{\phi_2(n)} \in [0,\infty) \ldots\ldots\ldots\ldots\ldots \quad .. (9)$$

or $\phi_1(*)$ is a non-decreasing function.

are satisfied, which ensures that the posterior distribution of the models is consistent. Where, ϕ_1 is the numerator function, ϕ_2 is the denominator function and $\phi_2^{'}(n)$ is the first order derivative of the function $\phi_2(n)$.

In comparison with the elicited parameter prior, it is shown that the numerator function ϕ_1 = 1 and its denominator function, $\phi_2(n) = n^r$ then the first derivative for the denominator function is $\phi_2^{'}(n) = rn^{r-1}$. That is,

$$g = \frac{\phi_1}{\phi_2(n)} = \frac{1}{n^r} \ldots\ldots\ldots\ldots\ldots \quad (10)$$

For $r \geq 3$, the conditions are satisfied as established below,

a. $$\lim_{n\to\infty} \frac{\phi_2^{'}(n)}{\phi_2(n)} = \lim_{n\to\infty} \frac{rn^{r-1}}{n^r} = \lim_{n\to\infty} \frac{r}{n} = 0 \ \ldots\ldots\ldots\ldots \quad (11)$$

And either,

b. $$\lim_{n\to\infty} \frac{n}{\phi_2(n)} = \lim_{n\to\infty} \frac{n}{n^r} = \lim_{n\to\infty} \frac{1}{n^{r-1}} = 0 \in [0,\infty) \ \ldots\ldots \quad .. (12)$$

Or,

c. $\phi_1(*) = 1$ (Constant) is a non-decreasing function.

The asymptotic properties of the modified parameter prior is derived as follows

I. Distribution of the modified parameter prior

$$\theta_j \mid g \sim N(0, \sigma^2 (\frac{1}{n^r} X_j^{'} X_j)^{-1}) \ \ldots\ldots \quad . (13)$$

where, σ is the precision for the model; $X_j^{'} X_j$ is the standardized design matrix of n × k for model j; and n is the number of observations.

II. Posterior Probability of the Parameter

From (4)

$$\theta \mid D, Z_j \sim f_t^{s_j}(c, \overline{\theta}_j, V(\theta \mid D, Z_j)) \ldots \quad .. (14)$$

Where $f_t^{s_j} \sim$ a multivariate t distribution, and c = degree of freedom.

Mean = $$\overline{\theta}_j = [(1 + \frac{1}{n^r} X_j^{'} X_j)]^{-1} X_j^{'} y, c > 1 \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (15)$$

Then,

$$V(\theta_j \mid D, Z_j) = \frac{(1+\frac{1}{n^r})^{-2} (X_j^{'} X_j)^{-1} [y^{'} \Omega_{X_j} y + \frac{1}{n^r} (y - \overline{y}_{i_n})^{'} (y - \overline{y}_{i_n})]}{c-2}, c > 1 \ \ldots\ldots\ldots \quad (16)$$

III. Marginal Likelihood of Model j

The marginal likelihood of model j is defined as

$$P(D \mid Z_j) \alpha (\frac{1}{n^r + 1})^{s_j/2} [\frac{1}{\frac{1}{n^r+1} + 1} y^{'} \Omega_{X_j} y + \frac{1}{n^r + 1} (y - \overline{y}_{i_n})^{'} (y - \overline{y}_{i_n})]^{-\frac{n-1}{2}} \ \ldots\ldots \quad \ldots.. (17)$$

IV. Posterior Model Probability

Each prior model probability is assumed for both models (j, q), then we have

$$P(D|Z_j)=\frac{(\frac{1}{n^r+1})^{s_j/2}[\frac{1}{\frac{1}{n^r+1}+1}y'\Omega_{X_j}y+\frac{1}{n^r+1}(y-\bar{y}_{i_n})'(y-\bar{y}_{i_n})]^{-\frac{n-1}{2}}}{\sum_{i=1}^{2^s}P(D|Z_i)} \quad \ldots \quad \ldots. (18)$$

V. Bayes Factor for Models in (j, q) models in BMA

Equal hyper parameter prior is assumed for the two models (j and q), then we have

$$V_{jq}=(\frac{1}{n^r+1})^{\frac{s_j-s_q}{2}}[\frac{\frac{1}{n^r}+1[y'\Omega_{X_q}y+\frac{1}{n^r}(y-\bar{y}_{i_n})'(y-\bar{y}_{i_n})]}{\frac{1}{n^r}+1[y'\Omega_{X_j}y+\frac{1}{n^r}(y-\bar{y}_{i_n})'(y-\bar{y}_{i_n})]}]^{-\frac{n-1}{2}} \quad \ldots\ldots\ldots\ldots \quad (19)$$

VI. Relationship of the modified g to information criterion

$$p\lim\frac{InV_{jq}}{\frac{n}{2}In(\frac{y'\Omega_{X_q}y}{y'\Omega_{X_j}y})+\frac{s_q-s_j}{2}In(n^r)}=p\lim\frac{InV_{jq}}{SC_{jq}}=1 \quad \ldots\ldots\ldots\ldots\ldots\ldots. (20)$$

VII. Predictive Distribution of Model q

The log predictive score follows a multivariate student's t distribution

$$i_j^*=\frac{1}{\frac{1}{n^r+1}+1}y'\Omega_{X_j}y+\frac{1}{n^r+1}(y-\bar{y}_{i_n})'(y-\bar{y}_{i_n}) \quad \ldots\ldots \quad . (21)$$

3. Simulation and Analysis

Using literature works of Hoeting *et al.* (1999), Fernandez *et al.* (2001a), Ley and Steel (2008), Eicher *et al.* (2011) and Olubusoye and Akanbi (2015) as guide, the simulation experiments are performed by designing a matrix H, for the regressors is an n × S matrix where S = 15 is a fixed number of regressors for sample size n, derived as follows: the first 10 columns in H, represented by ($h_{(1)}$,..., $h_{(10)}$) are drawn from independent Normal density and the subsequent five columns ($h_{(11)}$,..., $h_{(15)}$) are built standard form.

($h_{(11)}$,…, $h_{(15)}$) = ($h_{(1)}$... $h_{(10)}$)(0.3 0.5 0.7 0.9 1.1)'(1 1 1 1 1)+ ε …… . (22)

where ε = n × 5 matrix of independent standard normal variable. It is observed that (3.1) shows a correlation between the first five regressors and the last five regressors. The last five has the form of small to moderate correlations between ($h_{(1)}$... $h_{(5)}$) and ($h_{(11)}$,..., $h_{(15)}$). Each of the regressors will be demeaned after generating H, hence leading to matrix X* = ($X^*_{(1)}$... $X^*_{(15)}$) which satisfies i'X* = 0. A vector of size n is generated using one of the models:

Model 1:

y = 4 + 2$X^*_{(1)}$ - $X^*_{(5)}$ + 1.5$X^*_{(7)}$ + $X^*_{(11)}$ + 0.5$X^*_{(13)}$ + π ……… . (23)

Model 2:

y = 1 + π , $\pi \sim N(0,\sigma^2)$ …….. (24)

where σ^2 is assumed known to be an arbitrary value of 2.5^2. A uniform prior is used over the model space Z such that the model prior is:

P (Z_j) = 2^{-S}, j = 1... S……………….. (25)

where S is the number of available regressors.

Sample sizes of 50, 100, 1,000, 10,000 and 100,000 are used in the simulation experiment for each replication. 19 different vectors of regressors are generated for the forecast of model 1 with 100 replications. The Markov Chain Monte Carlo with 50,000 iterations after a burn-in drawing of 20,000 are used for the model simulation.

3.1. Posterior Model Inference

The posterior probability, which is a conditional probability of an unknown quantity, assigned to the model generating the data usually determine the performance of the Bayesian methodology. The true model probability is expected to be high for small or moderate values of sample size and it usually converges to 1 as sample size increases. Table 3.1 shows the means and standard deviations across the 100 replications for the true model 1 posterior probability. The row displays each of the r^{th} term of the modified g- prior from $r \geq 3$. It is obvious from this table that Model 1 is consistent. The prior performs well from n = 1000. It is seen that the prior performs best when r = 5. This implies that, as r increases the higher the rate of convergence to the true value of the parameter for Model 1.

Table 3.1: Means and Stds of the Posterior Probability of True Model 1 using the modified prior

N	50		100		1000		10000		100000	
Prior	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
$1/n^3$	0.5648	0.2079	0.7735	0.1893	0.9996	0.0008	0.9999	0.000	1.000	0.000
$1/n^4$	0.6301	0.2129	0.8217	0.1724	0.9997	0.0011	1.0000	0.000	1.000	0.000
$1/n^5$	0.7137	0.2009	0.8304	0.1963	0.9998	0.0476	1.0000	0.000	1.000	0.000

Table 3.2 shows the quartiles of the ratio between the posterior probability of the correct model and the highest posterior probability of any other model. The quartile ratio increases as sample size increases for the model 1. This shows how best the true model is as compared to the next best model.

Table 3.2 Quartiles of ratio of Posterior Model (1) Probability; True model vs Next best Model using the modified prior

N	50			100			1000			10000			100000		
Prior	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
$1/n^3$	0.4	1.6	3.87	1.7	4.9	9.96	15	120	320	150	2500	5302	1050	11210	23334
$1/n^4$	0.5	1.6	3.89	1.7	4.9	10.97	15	130	397	155	2532	5693	1051	11212	23410
$1/n^5$	0.6	1.7	3.89	1.7	4.9	10.97	17	140	409	160	2559	5899	1060	12203	25667

Table 3.3 displays means and standard deviations of the number of visited models over the 100 replications in the 50,000 iterations in the model space. As the sample size increases, knowing the model which produced the data is

one of the 2^{15} = 32,768 possible models investigated, it is expected to be as small as possible. The MC^3 sampler visits lesser models as n increases. At n = 100,000 the number of models visited is 8 as compared to thousands being visited when the sample size is 50.

Table 3.3: Means and Stds of Model 1 visited using the modified priors

N	**50**		**100**		**1000**		**10000**		**100000**	
Prior	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
$1/n^3$	4105.42	2212.142	1976.02	1701.635	16.21	9.675	8.79	1.372	8.53	1.141
$1/n^4$	3258.95	2161.26	1359.89	1543.604	11.11	11.820	8.53	1.141	8.53	1.141
$1/n^5$	2609.77	2013.68	1356.60	1462.843	13.61	7.418	8.53	1.141	8.53	1.141

Table 3.4 shows the degree of errors when the posterior model probability is allocated to the wrong sampling model. Notably, these tables indicate the means and standard deviations of the posterior probabilities of including each of the regressors. Since, model 1 contains regressors 1, 5, 7, 11 and 13 (indicated with asterisks in the table below). When n= 50, regressors 1 to 7 are usually included but as the sample size increases from 1000, the true regressors are usually included. Generally, the PIP of regressors not included in the correct model is approximately very small especially for large samples using the modified prior. It is also observed that the prior performs best when r = 5.

Table 3.5 shows the mean and standard deviations of the model 2 (null model) posterior probability. It is observed that the modified prior performs even when the sample size is small as compared with model 1 in Table 3.1. Also, like model 1, It is seen that the prior performs best when r = 5. This implies that, as r increases the higher the rate of convergence to the true value of the parameter for Model 2.

Table 3.6 displays the quartile ratios of the correct null model compared to any other models among the visited models for the elicited g- prior across the various sample size. The table below shows that as the sample size increases for the g- prior, the quartile ratio also increases for model 2. This shows how best the true model is as compared to the next best model.

Table 3.4: Means and Std of Posterior Probabilities of including each regressor, using the best modified prior $1/n^5$

N	50		100		1000		10000		100000	
Prior	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
*1	0.59	0.41	0.969	0.109	1.000	0.000	1.000	0.000	1.000	0.000
2	0.00	0.41	0.000	0.109	0.000	0.000	0.000	0.000	0.000	0.000
3	0.00	0.41	0.000	0.109	0.000	0.000	0.000	0.000	0.000	0.000
4	0.01	0.41	0.011	0.109	0.003	0.000	0.003	0.000	0.000	0.000
*5	0.00	0.41	0.018	0.109	0.997	0.000	1.000	0.000	1.000	0.000
6	0.00	0.41	0.000	0.109	0.000	0.000	0.000	0.000	0.000	0.000
*7	0.17	0.41	0.614	0.109	1.000	0.000	1.000	0.000	1.000	0.000
8	0.00	0.41	0.000	0.109	0.000	0.000	0.000	0.000	0.000	0.000
9	0.00	0.41	0.000	0.109	0.000	0.000	0.000	0.000	0.000	0.000
10	0.00	0.41	0.000	0.109	0.000	0.000	0.000	0.000	0.000	0.000
*11	0.39	0.41	0.865	0.109	1.000	0.000	1.000	0.000	1.000	0.000
12	0.03	0.41	0.009	0.109	0.000	0.000	0.000	0.000	0.000	0.000
*13	0.11	0.41	0.062	0.109	0.992	0.000	1.000	0.000	1.000	0.000
14	0.02	0.41	0.008	0.109	0.000	0.000	0.000	0.000	0.000	0.000
15	0.02	0.41	0.019	0.109	0.000	0.000	0.000	0.000	0.000	0.000

Table 3.5: Means and Stds of the Posterior Probability of True Model 2 using the modified prior

n	50		100		1000		10000		100000	
Prior	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
$1/n^3$	0.8929	0.0941	0.9564	0.0523	0.9993	0.0021	0.999	0.000	1.000	0.000
$1/n^4$	0.9819	0.0285	0.9941	0.0237	0.9999	0.0000	1.0000	0.000	1.000	0.000
$1/n^5$	0.9980	0.0048	0.9995	0.0028	1.0000	0.0000	1.0000	0.000	1.000	0.000

Table 3.6 Quartiles of ratio of Posterior Model (2) Probability; True model vs Next best Model using the modified prior

n	**50**			**100**			**1000**			**10000**			**100000**		
Prior	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
$1/n^3$	3.2	6.8	15.1	9.3	15.6	29	19.1	84.3	289	29.1	104.3	3602	191	843	26358
$1/n^4$	3.5	7.2	16.5	9.3	15.9	33	19.1	84.3	302	29.1	114.3	3675	191	845	28402
$1/n^5$	4.0	8.1	18	9.7	17.3	38	19.1	84.3	321	29.1	124.3	3876	197	847	29973

Table 3.7 depicts the means and standard deviations of the number of visited models sing the null model. The modified g- prior shows that the number of models visited decreases as sample size increases. The prior performs best when r = 5.

Table 3.7: Means and Stds of Model 2 visited using the modified priors

N	**50**		**100**		**1000**		**10000**		**100000**	
Prior	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
$1/n^3$	1077.69	950.804	438.58	493.241	25.81	24.510	13	1.907	12.67	1.441
$1/n^4$	190.51	260.135	78.73	214.628	12.95	1.708	12.72	1.583	12.65	1.441
$1/n^5$	39.39	46.36	19.64	24.684	12.59	1.436	12.7	1.541	12.65	1.441

3.2. Predictive Inference

In statistical inference, predictive inference is the prediction of future observations based on past observations. For this study, we conditioned our predictions on the values of the regressors by choosing 19 different vectors for the S = 15 regressors and focus especially on the vectors that lead to the minimum for the model 1. Table 3.8 below presents the median of the Log Predictive Score, computed across 100 samples for different sample sizes. The LPS is only a Monte Carlo approximation based on 100 draws, therefore, the lower bound is not always strictly adhered to. The modified prior performs best at n =1000, while as the sample size increases the prior is already over predicting at all levels.

Table 3.8: Conditional Medians of LPS point prediction using the modified priors

n	50	100	1000	10000	100000
Prior	X^*_{min}	X^*_{min}	X^*_{min}	X^*_{min}	X^*_{min}
$1/n^3$	2.357	2.401	2.249	2.395	2.402
$1/n^4$	2.357	2.401	2.388	2.396	2.402
$1/n^5$	2.357	2.401	2.388	2.401	2.402

To compare the differences between the predicted density and the sampling density of Model 1, we plot both densities for different sample sizes as shown in Figure 3.1. The figures shows the comparison for different values of n and the predictives for 25 of the 100 generated samples. The dark line corresponds to the actual sampling density while the green lines represent the predictive densities. It is observed from the figures that as the sample size increases, the predictive densities merge closer together in the direction of the actual sampling density.

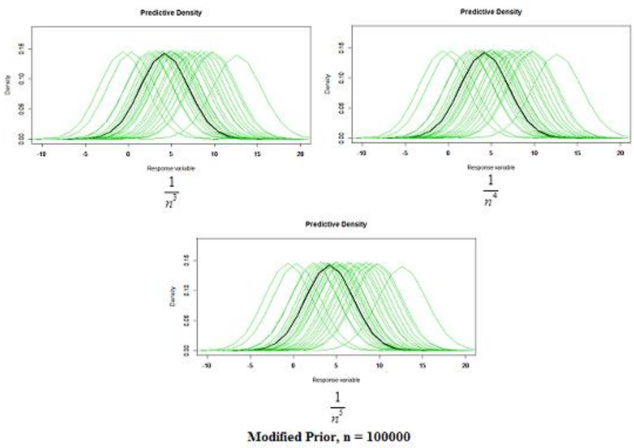

Figure 3.1: The modified g prior when n = 100000

Comparing the overall predictive performance from the Log Predictive Score for the 19 different values of the regressors and the 100 samples of replications. This is displayed in the Table 3.9 below, where the medians are recorded for the modified prior and different sample sizes. At n = 50, the modified prior predicts well at r = 5 because it is close to the threshold of 2.335.

Table 3.9: Medians of LPS overall prediction using the modified priors

N	50	100	1000	10000	100000
Prior	X^*_{min}	X^*_{min}	X^*_{min}	X^*_{min}	X^*_{min}
$1/n^3$	2.301	2.401	2.249	2.395	2.402
$1/n^4$	2.313	2.402	2.295	2.397	2.402
$1/n^5$	2.326	2.426	2.331	2.402	2.402

4. Application of BMA with the Best Modified Prior to CO_2 Emissions in Nigeria.

One of the main greenhouse gases in the atmosphere is CO_2. It is emitted to the atmosphere through many ways, but the larger emissions of the gas in the atmosphere leads to higher concentration in the atmosphere thereby altering the global carbon cycle and causing global warming of the earth planet. Emissions from a number of growing economies have been increasing rapidly over the last few decades. Fast-forwarding to annual emissions in 2014, we can see that a number of low to middle income nations are now within the top global emitters. In Nigeria, CO_2 gas is emitted from a lot of sources. To this end, BMA approach is applied to CO_2 emissions to account for the uncertainties embedded in both the parameters and the model itself. The data used are yearly which spans through 1975 to 2015 (41 years) and were sourced from World Bank website and EDGAR database reported by UN. The study variable is Annual CO_2 emissions while the predictors are the sectorial emissions groupings which are International Bunkers (INTBK), Waste (WST), Resident and Commercial (RESCOM), Industry (INDST), Transport (TRPT), Agriculture (AGR), Forestry (FRST), Land Use (LDUS), Energy (EGY), Fossil Fuel (FOSFUEL), Gaseous Fuel Consumption (GFC), Liquid Fuel Consumption (LFC), Solid Fuel Consumption (SFC) and Other Sources (OTS). The MC^3 sampler uses 50,000 draws after the burn-ins of 20,000 with uniform distribution as the prior model and the modified g-prior g = $1/n^5$ for parameters. Therefore, the CO_2 emission model is given as:

$$CO_2 emissions = \theta_0 + \theta_1 INTBK + \theta_2 WST + \theta_3 RESCOM + \theta_4 INDST + \theta_5 TRPT + \theta_6 AGR$$

$$+\theta_7 FRST + \theta_8 LDUS + \theta_9 EGY + \theta_{10} FOSFUEL + \theta_{11} GFC + \theta_{12} LFC + \theta_{13} SFC + \theta_{14} OTS + \varepsilon$$

Where, ε is a stochastic error term, independent and identically distributed as $N(0,\sigma^2)$.

Table 4.1 represents the means and standard deviations of the Posterior Inclusion Probabilities (PIP) of each regressors in the CO_2 emission model. Post Mean displays the coefficients averaged over all models, including the models wherein the variable was not contained (implying that the coefficient is zero in this case). The covariate Industrial Emissions with PIP of 99% has a comparatively large coefficient band seems to be the most important. Other covariates Residential and Commercial Emissions with PIP of 78.7% and Agricultural Emissions with PIP of 75.06% are also important variables in modelling Nigeria CO_2 emissions. This shows that for any CO_2 emission model selection, Industrial, Residential & Commercial and Agricultural sector plays a crucial role. The posterior mean of the important variables are negative which conforms to the theory. The negative sign is clearly explained by the conditional posterior sign in the 5th column of Table 4.1. In Table 4.1, all other PIPs are less than 50% and also, the posterior standard deviations are all greater than the posterior means.

Table 4.1: Posterior Probabilities of including the Regressors in the CO_2 emission model

Regressors	PIP	Post Mean	Post SD	Cond.Pos.Sign	Index
INDST	0.999	-0.59	0.102	0.000	4
RESCOM	0.787	-0.25	0.143	0.000	3
AGR	0.751	-0.27	0.162	0.000	6
LDUS	0.171	0.06	0.142	1.000	8
SFC	0.111	-0.03	0.078	0.000	14
GFC	0.014	-0.00	0.025	0.021	12
FRST	0.012	-0.00	0.024	0.000	7
FOSFUEL	0.011	-0.00	0.017	0.000	10
LFC	0.011	-0.00	0.017	0.000	13
INTBK	0.006	0.00	0.011	1.000	1
TRPT	0.005	-0.00	0.009	0.000	5
OTS	0.004	-0.00	0.006	0.000	11
EGY	0.004	0.00	0.006	0.136	9
WST	0.004	0.00	0.009	0.458	2

It is seen that the number of explanatory variables is 14 and observations of interest is 41 years. The number of model space is 16384 and the number of models visited is 12672 which indicates 76% of the models was visited. The best modified g-prior $\frac{1}{n^5}$ used for the application to CO_2 emissions in Nigeria established that the shrinkage factor close 1 which denotes over fitting. Table 4.2 shows the posterior probabilities of the best five models among the 2672 models visited for both the MCMC (MC^3) and the exact samplers. This table indicates that the best model with probability of 53% represents CO_2 emissions in Nigeria which includes the Industry, Resident & Commercial and Agriculture as the predictor for the model. The table indicates that the true CO_2 emission model (0d00) is always favored compared to any other model.

Table 4.2: Best 5 models of 12672 models visited

	PMP (Exact)	PMP (MCMC)	Predictors
0d00	0.53179	0.55026	RESCOM, INDST and AGR
0d01	0.10311	0.10508	RESCOM, INDST and AGR
0440	0.08914	0.08104	INDST and LDUS
0500	0.06664	0.06240	INDUST and AGR
0c40	0.06096	0.06170	RESCOM, INDST and LDUS

Figure 4.1 is a visualization of the mixture marginal posterior density for the important regression coefficients. (Industry, Residential & Commercial, Agriculture and Land Use Sector). The dotted vertical lines in red shows the corresponding standard deviation bounds from the MCMC approach, the red and green vertical line shows the

conditional expected value and median respectively. It is observed from the plots that the values of the conditional expected value and the median are close to each other. The density in each of the graphs describes the posterior distribution of the regression coefficient given that the corresponding variable is included in the regression.

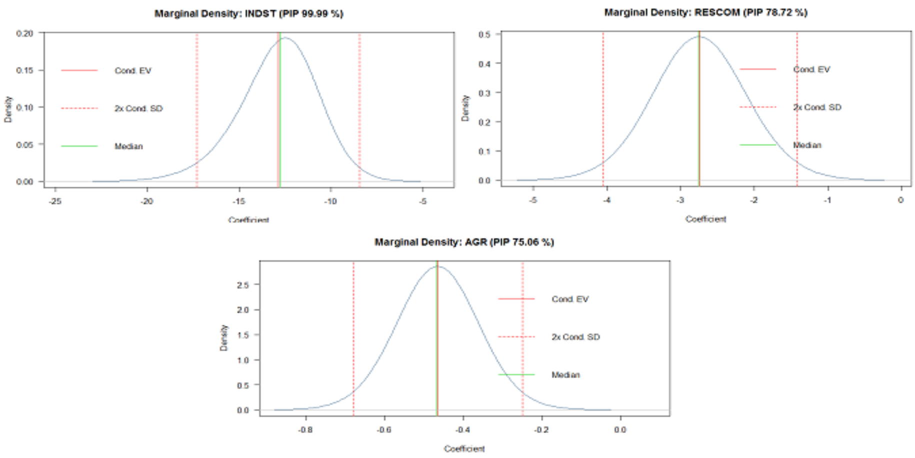

Figure 4.1: Posterior Density of Important Coefficients

Figure 4.2 shows the cumulative model inclusion probabilities based on the best 50 models. It also depicts the inclusion of a regressor with its sign in the model selection process. The red color corresponds to a negative coefficient, the blue color corresponds to a positive coefficient and the white to a non-inclusion of the respective variable. The horizontal axis is scaled by the models' posterior model probabilities.

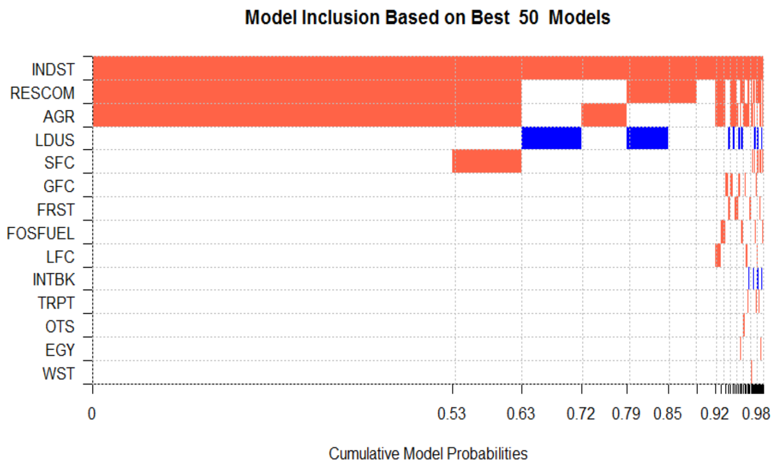

Figure 4.2: Cumulative Model Probabilities.

5. Conclusion

For this study, a g-parameter prior was elicited which is partly non informative in structure related to a Normal conjugate g-prior. The asymptotic properties (parameter prior distribution, marginal likelihood of the model, Bayes factor, posterior parameter distribution, posterior model probability, predictive distribution, relationship to an information criterion) for the modified prior was derived. The empirical results on both the posterior model and predictive inferences seem to indicate that the modified prior $1/n^r$ is a sensible choice parameter prior in the BMA technique. The modified g-priors $(g_j = 1/n^r, r = 3,4,5)$ performed better in model selection whenever informative prior is unavailable compared with the priors of FLS (2001a) and Olubusoye and Akanbi (2015). Applying the best modified prior $1/n^5$ to CO_2 emissions in Nigeria shows the sectors that contribute more to CO_2 emissions in Nigeria. The research established that the larger the r of the g-parameter prior, the more the posterior probability values and the better the performance of the model.

- Industrial Sector emissions with a PIP of 99.98% is important in modelling CO_2 emissions in Nigeria.
- Resident and Commercial Sector emissions with a PIP of 78.73 % is important in modeling CO_2 emissions in Nigeria.
- Agricultural Sector emissions with a PIP of 75.06 % is important in modelling CO_2 emissions in Nigeria.

As important as CO_2 is, in sustaining a habitable temperature, continuous increase in the emissions can disrupt the global cycle and thereby lead to a planetary warming impact. Since CO_2 emissions is rapidly increasing in Nigeria, and through this study we have been able to discover that the Industry, Agriculture, Resident and Commercial sector plays the most important role in the emission to the environment, there is need for the concerned authorities to restructure these sectors and provide necessary adjustment to reduce carbon emission being released to the environment or to provide ways by which the carbon emitted will be properly sequestered.

References

Amini, S.M. and Parmeter, C.F. (2011). Bayesian Model Averaging in R.

Eicher, T.S., Papageorgiou, C. and Raftery, A.E. (2011). Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants. *Journal of Applied Econometrics* 26(1): 30- 55.

Faust, J., Gilichrist, S., Wright, J.H. and Zakrajssek, E. (2013). Credit Spreads as Predictors of Real Time Economic Activity: A Bayesian Model Averaging Approach. *Rev. Econ. Stat.*, 95, 1501-1519.

Fernandez, C., Ley, E. and Steel, M.F.J. (2001a). Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics*, 100, 381-427.

Fernandez, C., Ley, E. and Steel, M.F.J. (2001b). Model Uncertainty in Cross Country.

Hoeting, J.A. Madigan, D., Raftery, A.E. and Volinsky, C.T. (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science*, 14: 382-401.

Leon-Gonzalez, R. and Montolio, D. (2004). Growth, Convergence and Public Investment: A BMA Approach. *Applied Economics*, 36, 1925-1936.

Ley, E. and Steel, M.F.J. (2008). On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression. *Munich Personal RePEc Archive, 6773*.

Liu, J. and Xie, Z. (2014). BMA Probabilistic Quantitative Precipitation Forecasting Over the Huahie Basin Using TIGGE Multi Model Ensemble Forecasts. *Mon. Wea. Rev.*, 142, 1542-1555.

Olubusoye, O.E. and Akanbi, O.B. (2014). Bayesian and Classical Estimation Methods in Linear Regression Model: A Monte Carlo Simulation Analysis. *Journal of Science Research*, 13: 73-83.

Olubusoye, O.E. and Akanbi, O.B. (2015). On g-Prior Elicitation in Bayesian Model Averaging Approach to Normal Linear Regression Model. *Conference Proceedings of Mathematics Department, University of Ibadan, Nigeria. In honour of S.A. Ilori's 70th Birthday*.

Olubusoye, O.E., Akanbi, O.B. and Kazeem, B.A. (2015). Modelling the Nigerian Inflation Process using the Bayesian Model Averaging Approach. *Proceedings of the 19th Annual Conference of the Africa Region Chapter of the Econometric Society, Ethiopia*, III: 117-140.

Raftery, A.E., Madigan, D. and Hoeting, J.A. (1997). Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association*, 92, 179-191.

Raftery, A.E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Mon. Wea. Rev.* 133, 1155-1174.

Raftery, A.E., Painter, I.S. and Volinsky, C.T. (2005). BMA: An R Package for Bayesian Model Averaging. *R News*, 5, 2-8.

Sala-i-Martin, X.X., Doppelhofer, G. and Miller, R.I. (2004). Determinants of Long Term Growth: A Bayesian Averaging of Classical Estimates (BACE) approach. *American Economic Review*, 94, 813-835.

Sloughter, J.M. Raftery, A.E., Gneiting, T. and Fraley, C. (2007). Probabilistic Quantitative Precipitation Forecasting using Bayesian Model Averaging. *Mon. Wea. Rev.*, 135, 3209-3220.

Dr. O.B. Akanbi: He lives in Ibadan, Nigeria. He is a lecturer in the Department of Statistics, University of Ibadan, Nigeria. His areas of research are Econometrics, Bayesian Inference or Modelling (Bayesian Model Averaging) and Operations Research. He has Professional Diploma (Statistics), B.Sc. (Statistics), M.Sc. (Statistics) and Ph.D. (Statistics) from the University of Ibadan, Nigeria. Dr. Akanbi has to his credit a number of referred journal articles and conference papers within and outside Nigeria. He is a member of some statistical societies; International Statistical Institute (ISI); International Biometric Society (IBS) Group Nigeria (GNi); Statistics and Probability African Society (SPAS); Nigerian Statistical Association (NSA); and Nigerian Young Statisticians Group (NYSG).

Oladapo M. Oladoja: He has Professional Diploma (Statistics) with a credit grade, B.Sc. (Statistics) with a First Class grade and M.Sc. (Statistics) with a Ph.D. grade all from the University of Ibadan in Nigeria. He had experiences in academics having been a Graduate Assistant at the Department of Actuarial Science, University of Jos, Nigeria before becoming a Research Scholar at the Department of Statistics, University of Ibadan. He is currently an Assistant Lecturer at the Department of Mathematics and Statistics, First Technical University, Ibadan, Nigeria. He is a member of International Biometric Society (IBS GNi) Nigeria Group. He has publications in different journals. His area of interest includes Environmental Statistics, Biometrics, Bayesian Inference, Actuarial Science, Operations Research, Mathematical Statistics and Computational Statistics.