International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Empirical Validation of Hierarchical Models: Assessing Performance, Interpretability, and Implications

Seth Opoku Larbi[1], Apaka Rangita[2], Joyce Otieno[3]

[1]MSc, Planning Office, Koforidua Technical University, Directorate of Quality Assurance & Academic Planning, Ghana.
[2,3]PhD, Department of Statistics & Actuarial Science, Maseno University, School of Mathematics, Statistics and Actuarial Science, Kenya.

Abstract
Hierarchical data structures emerge when observations are nested within higher-level units or clusters. Existing research often ignores the hierarchical structure of data, leading to biased estimates, suboptimal model selection, and challenges in identifying important predictors and dependencies. This study contributes to hierarchical frameworks by addressing interpretability challenges of random effects, scalability, regularization and transparency in the standard Bayesian model. This study specific aims are to use a unique hierarchical Bayesian model to improve the analysis of hierarchical data and also to test and compare the performance and interpretability of this unique hierarchical Bayesian model against the standard model empirically. The unique Bayesian hierarchical model advances the Standard Bayesian Hierarchical Model by introducing a contextual variable (X_{ijz}) and parameters to the random effects. These advancements significantly improve the accuracy, reliability, and interpretability of the model. Hierarchical Bayesian Information Criteria (HBIC) was the technique used for model selection. This reduces overfitting and improves the accuracy of estimation, particularly in accounting for heterogeneity. The findings indicate that the introduction of the contextual variable 'X_{ijz}' (Estimate = −0.0505, SE=0.0127, p<0.01) enables the model to account for cluster-specific effects, thereby improving its ability to identify region-specific phenomena. The fixed effect values for age factor (Estimate = 0.008, SE=0.0012, p<0.01) and religious factor (Estimate=0.0282, SE=0.0096, p<0.05) highlight the essential relationships captured by the unique model. The introduction of shrinkage parameters ϕ_j and ψ_j plays a crucial role in regulating parameter estimates toward a common value. The random effects demonstrated variability at the group level with county intercept variance (0.135) and age factor variance (0.00328), indicating the model's capacity to capture group-specific heterogeneity. Lastly, the study demonstrates that this unique hierarchical Bayesian model outperforms the standard model by improving accuracy, interpretability, scalability, and regularization.

Keywords: Shrinkage, Hierarchical, Parameters, Prior, Posterior, Uncertainty, HBIC, Unique Variable

Introduction
There has been a surge in modern research and an interest in learning and understanding data-driven decision-making. Having a larger dataset for analysis is no more sophisticated, as insights in the

underlying huge dataset are no more hidden which makes it exciting and appealing (Larbi et al., 2024). Despite such advancement in research methodologies, there is a limitation in hierarchical model procedure where potential researchers find it difficult to extract all issues surrounding the data. Most fascinating studies have looked into examining the predictor variables separately ignoring the aspect of interaction and the nested aspect within the dataset (Johnson, 2018). Hierarchical models and structures including the levels of data or clustered observations are common in social science, ecology, and epidemiology areas (Green & White, 2017). Disregarding such structures leads to estimation bias and incorrect conclusions demanding robust statistical tools that address the complexity of data (Rajaraman, 2020). A common approach in analysing multilevel data is through a hierarchical framework by introducing random effects to account for variability in the level of structures (James & Douglas, 2016). The Unique model adopts a hierarchical structure where the individual-level predictors fixed effects, and group-level variables are captured through the random effects (Larbi et al., 2024; Klein & Moeschberger, 2018). It incorporates interactions within the individual-level explanatory variables and group-level random effects which regulate whether a lower-level factor can moderate that of a higher-level one (Larbi et al., 2024; Gelman & Hill, 2007).

Hierarchical Bayesian modelling is a formal statistical framework that has established norms that guide its application and interpretation. Hierarchical modelling ensures that the procedure for a model is valid, reproducible, and transparent. The hierarchical model basic norm concentrates specifically on the prior distribution of parameters which should be clearly defined. The knowledge and beliefs of the parameters considered in the model ought to be chosen in the right manner (Larbi et al., 2024). Hierarchical modelling allows for the nested structures within data and the incorporation of uncertainty, prior information, and variability with the structure settings. The parameters of interest are treated as random variables in Hierarchical modelling. Prior distributions are also defined for each parameter of interest at different levels in the hierarchy to cater for the uncertainty about the parameters before the data is observed. Bayesian theorem is used to observe the posterior distribution of data. The posterior distribution is updated using the combination of likelihood data and prior information.

The hierarchical Bayesian model has gained several successes although there have been teething limitations and challenges that ought to be considered to formulate a parsimonious model. One of the biggest challenges encountered by researchers in hierarchical modelling is its intensiveness in computing large datasets and at times complex structures. This results in long computational times, high autocorrelation, and slow convergences because of the posterior inferences from the Markov Chain Monte Carlo (MCMC) algorithm. Informative priors that present accurate prior knowledge selection are essential but become a challenge when there is no or slight essential information. Wrong specification of priors results in biased estimates and unreliable conclusions (Klein & Moeschberger, 2018). If the model does not include important predictors or it is not well understood how the data was designed, it can fail to represent significant structures and patterns contained in the data. The hierarchical structure of data is many at times ignored by researchers which leads the model to bias in estimation, suboptimal model selection, and also challenges to be able to identify the essential predictors and dependencies which leads to potential suboptimal model construction and interpretation (Larbi et al., 2024).

However, the major hope for researchers in addressing the aforementioned challenges is through the use of the Hierarchical Bayesian model. Bayesian hierarchical modelling provides a reproducible framework and transparency with articulated guidelines for validating models. There are even practical challenges with model transparency, uncertainty quantification, scalability, interpretability of the random effects, and computational feasibility of which researchers have to be circumspect (Larbi et al., 2024). Failure of existing research or statistical methodologies in identifying problems to accommodate hierarchical structures in datasets often leads to suboptimal model and variable selection, and inefficient and biased estimates. Therefore, in order to address hierarchical data analysis challenges and enable accurate model interpretability, transparency, and making informed decisions, there is a need to address model comparison challenges, reproducibility, uncertainty quantification, and model transparency through a robust statistical technique.

The hierarchical structure of a data is addressed amidst the limitations encountered in hierarchical modelling methodologies by using the unique model formulated by Larbi et al., (2024) to enhance the robustness and predictive accuracy as well as accounting for heterogeneity. The use of model selection criterion Hierarchical Bayesian Information Criteria (HBIC) as a tool for the selection of best variables into the unique model improves the regularisation and also reduces overfitting. The existing research methodologies regarding model selection procedures focused on individual characteristics; though the unique model addresses the heterogeneity and dependencies within nested data. The stability of the variables selected in the model, the interpretability of model results, model fitness, and predictive accuracy is the outcome of the unique hierarchical Bayesian model. The general objective of this study is to advance on the standard Bayesian model to enhance accuracy, interpretability, and generalizability. This study has specific aims which are to use the unique hierarchical Bayesian model to improve the analysis of hierarchical data and also to test and compare the performance and interpretability of the unique hierarchical Bayesian model against the standard model empirically.

The study contributes to scientific knowledge by introducing a hierarchical Bayesian model that is based on a global test (of fixed or random effects) under the likelihood ratio test framework. Secondly, the study uses model selection techniques within the hierarchical modelling framework to improve the performance of the model. The results of this study have direct implications in a wide range of domains including education, finance, healthcare, and social sciences, where hierarchical structures and high-dimensional data are common. The study also enhances the quality and interpretability of methods used in science and other investigations. Specifically, by introducing a unique model Larbi et al., (2024) enhanced the accuracy and interpretability data in the hierarchical arena. This paper aims to make more feasible good decisions where transparency and understanding of model outputs are of utmost priority.

Literature Review

Modern data analysis has brought a surge in statistical literature on hierarchical model and variable selection techniques for addressing complex data more interpretably and transparently. These techniques have gained much attention as a result of the hierarchical structure's ability to provide insight into data patterns and address its selection limitation to approaches. The hierarchical model allows for accurate capture of complex relationships and also models the dependencies among observations within the hierarchies. Improving the model's parsimony and interpretability is ensuring that a model selection technique is employed to elect the right predictors. A valuable contribution was made to the literature on Bayesian hierarchical modelling, by introducing a new approach to improving predictive performance through model selection criteria (Zong & Bradley, 2023). The article's theoretical and empirical contributions provide a valuable resource to researchers seeking to improve the accuracy and efficiency of Bayesian hierarchical models across a diverse set of applications by incorporating an often-used criterion from model selection (Zong & Bradley, 2023). A special case of several information criteria expressions was proposed by Zong & Bradley (2023) which they labelled as Covariance Penalized Error (CPE). By the variance empirical Bayes estimator, a penalized mean value was formed. Zong & Bradley's (2023) prime aim was to obtain a small value of the CPE criterion, by truncating the joint support of the data and the parameter space using Bayesian hierarchical modelling. The authors achieved their objective of minimizing the squared error by identifying a subset parameter space that produces lower values than the Bayesian model averaging yields, provided that there is a non-zero probability value within this truncated set.

$$\mathrm{CPE} = \mathrm{E}[(\mathrm{g}(\theta) - \acute{\mathrm{g}}(\mathrm{y}))^2] + \lambda \cdot \mathrm{Tr}[\mathrm{Cov}(\acute{\mathrm{g}}(\mathrm{y}))] \qquad \textbf{[1]}$$

Where

- The 1st term of equation 1 measures the mean square error
- The 2nd term of equation 1 is the penalty term involving the trace of the Covariance matrix of the estimator
- λ is the Penalty weight
- Tr(Cov) reflects the posterior uncertainty (variable estimator)
- The Penalty helps in controlling overfitting in model selection

They introduced the idea of restricting the parameter space or the joint support of (y, θ); thus, effectively using a truncated version of the hierarchical posterior to get better predictions:

$$\theta \in \Theta^* \subset \Theta \text{ such that } \mathrm{P}(\theta \in \Theta^* \mid \mathrm{y}) > 0$$

Parameter truncation; This subset Θ^* is selected to minimize the CPE, i.e., find a "simpler" model where predictions remain good with lower uncertainty.

Variance Empirical Bayes Estimator (VEB) was used as a penalized mean estimator,

$$\acute{\theta}_{VEB} = \arg \underbrace{min}_{\theta} \{(y - \theta)^2 + \lambda \cdot Var\,(\theta|y)\} \qquad \textbf{[2]}$$

Penalized estimator: This combines the square error (fit) and Posterior variance (complexity penalty)

The limitation in their paper comprises complexity of implementation, interpretability and Transparency, and Dependency on Criterion Selection.

The paper by Gomez-Mendez and Chainarong (2023) examines poverty-related variables such as income, education, and others in the multiresolution governing structural data of Thailand. Bayesian hierarchical models are utilized for this analysis. The authors discuss the progression of their modelling methodology from simple to more complex models and assess each model's effectiveness based on its ability to explain relevant variables while balancing complexity considerations. The methodology in the article “Income, education, and other poverty-related variables: a journey through Bayesian hierarchical models” by Gómez-Méndez and Chainarong (2023) can be summarized in mathematical and statistical terms as follows:

The hierarchical model aims to explain income and other poverty-related variables in Thailand. It can be represented as:

$$\mathrm{Y}_{ij} = \beta_0 + \sum_{k=1}^{p} \beta_k \mathrm{X}_{ijk} + \phi_j \varphi_{0j} + \varepsilon_{ij} \qquad \textbf{[3]}$$

Where

Y_{ij} is the income or poverty − related variable for observation i in region j,

β_0 is the intercept

β_k are the coefficients for covariates X_{ijk} (such as education level, household characteristics, etc.),

ϕ_j is the regional random effect capturing unobserved heterogeneity at the regional level,

φ_{0j} represents the random effect for region j,

ε_{ij} is the error term

The journey started by the authors in designing the hierarchical models in a simple and then gradually complex form. The performance of each model was assessed based on the ability to explain the variability in income and poverty-related variables. In practice, the preference for simpler models should not only be based on the explanatory power but also the complexity by accounting for dependencies. If two models provide comparable explanatory power a simpler model is preferred to a more complex one. So, the evaluation of model fit, overfitting, and computational complexity for each model should be discussed. Thus, the regularization techniques (e.g., fixed-effects, random-effects, and mixed-effects models), model averaging, and sensitivity analysis for Bayesian hierarchical models. The limitations highlighted in G´omez-M´endez and Chainarong (2023) article as a single policy may not adequately address poverty-related issues in different areas, each with its unique challenges. Custom-made policies for each area separately are unrealistic due to resource limitations. The article points out the limitation of ignoring dependencies or relationships between different geographic areas when formulating poverty-related policies. While Bayesian hierarchical models provide a flexible and powerful framework for analyzing complex data, the authors' model has drawbacks, such as challenges in model specification, computational complexity, and potential overfitting because a large number of parameters were included. The availability, and quality, of data, may also limit the utility of the modelling approach.

The Article by Porter et al., (2023) titled “Objective Bayesian Model Selection for Spatial Hierarchical Models with Intrinsic Conditional Autoregressive Priors” addresses the problem of model selection for Gaussian hierarchical models with intrinsic conditional autoregressive (ICAR) spatial random effects. Porter et al., (2023) acknowledges the problem of selecting covariates and spatial model structure independently in spatial hierarchical models; previously, methods required one to be fixed a priori and the other to be selected, which necessitated arbitrary decisions. Methods for simultaneous selection solve this problem but there are few Bayesian methods for simultaneous selection; the Article aims to do so, using fractional Bayes factors for model selection under automatic reference priors. In developing this methodological contribution, the Article first develops Bayesian model selection by fractional Bayes factors, so that fixed effects especially are selected simultaneously with spatial model structure; the approach also uses automatic reference priors, so that hyperparameters for priors do not need to be specified. The Article then establishes the stochastic ordering of two ICAR specifications and its implication for the fractional Bayes factor for the ICAR model under the reference prior. A comparison of the methodology to traditional model selection criteria and its performance in a simulation study is also undertaken.

Porter et al., (2023) hierarchical model under review is represented as:

$$Y_i = \beta_0 + \sum_{k=1}^{p} \beta_k X_{ik} + \phi_i + \varepsilon_i \qquad \textbf{[4]}$$

Where

Y_i is the response varable for observation i

β_0 is the intercept

β_k are the coefficients for covariates X_{ik}

ε_i is the error term

Spatial Random Effects: The model includes spatial random effects represented by the ICAR prior, denoted as ϕ_i.

Fractional Bayes Factors (FBFs): FBFs are used for model selection. M_0 represents the null model with no covariates and only spatial random effects, and M_1 represents the model with covariates and spatial random effects. The Fractional Bayes Factors (FBFs) for comparing models M_0 and M_1 are as follows:

$$FBF_{01} = \frac{BF_{01}}{1+ PBF_{01}} \quad \textbf{[5]}$$

where

BF_{01} is the Bayes factor comparing models M_0 and M_1,

FBF_{01} is the fractional Bayes factor.

Porter et al., (2023) approach offers a unique solution to model selection in spatial hierarchical models. Porter et al., (2023) model in the hierarchical framework has challenges including computational complexity, potential challenges in applying the method to large-scale spatial datasets, and sensitivity to prior specifications that the researchers encountered. Additionally, the approach's performance may vary based on the unique characteristics of each dataset and model specifications. Porter et al., (2023) limitation could be addressed using the Larbi et al., (2024) model which could be relevant to the article in these aspects.

Methodology

The methodology section of this paper outlines the systematic approach employed to achieve the objectives of adding a unique variable to the Standard Bayesian model postulated by Larbi et al., (2024) using empirical data. The paper advances the Standard Bayesian model with a unique variable to the fixed effect and parameters to the random effects respectively. The paper is an extension of Larbi et al., (2024) which adopted a theoretical design approach; highlighting how the introduction of a unique variable and shrinkage parameters improve the model's theoretical grounding and possibly align it more closely. The theoretical design approach focuses on the conceptual justification that explores model structures conceptually without needing empirical validation thereby accounting for unobserved heterogeneity or the group-level effects. Larbi et al., (2024) emphasizes on the theoretical significance of the introduction of parameters ϕ_j, ψ_j, and X_{ijz} as a contextual factor to be observed in the model. This paper validates the theoretical design, the performance of the Standard Bayesian model and the unique model formulated by Larbi et. Al, (2024) using real-world data from PMA 2022. The dependent variable considered was Maternal Health Utilisation (General Health of Mothers) and the explanatory variables were age and religious factors after study employed variable and model selections techniques for the selection of such predictors into the model. This was to prevent the model from suffering from multicollinearity, overfitting and select the best variables for parsimonious hierarchical structure. The PMA dataset contained 50, 938 observations out of which 11,053 were eligible for the study. The predictive Mean Marching approach was employed to handle missing values under the Multiple Imputation by Chained Equation (MICE) framework. The MICE use an iterative approach based on Gibbs sampling where an initial guess is made using the simple technique of the mean. The variables were treated based on their characteristics(type) as Linear Regression for Continuous and Logistic Regression for Categorical Variables under the umbrella of Gibbs for the clean-up.

Model Selection

The Hierarchical Bayesian Information Criterion (HBIC) was employed for the model selection. Inclusively targeting all levels, it merges information standards while integrating adaptive penalizations corresponding to hierarchy gradation that combats overfitting via weighting level contribution based on intrinsic hierarchal intricacy.

$$\mathrm{HBIC} = -2\sum_{i=1}^{N}\log\{P(y_i|\theta_i)\} + k\sum_{j=1}^{J} w_j(\log N_j) \qquad [6]$$

Where:

N is the total number of observations, $P(y_i|\theta_i)$ is the likelihood for observation i given parameters θ_i,

k is a penalty factor for the model complexity,

J is the number of levels in the hierarchy

w_j is the weight associated with level j,

N_j is the number of observation at level j.

Original /Standard Hierarchical Bayesian Model

Below is the Standard Bayesian hierarchical model:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij1} + \beta_2 X_{ij2} + u_{0j} + u_{1j}X_{ij1} + \varepsilon_{ij} \qquad [7]$$

Where:

Y_{ij} is the dependent variable for observatjion i in group j

X_{ij1}, X_{ij2} are the predictor variables,

u_{0j} , u_{1j} are group − specific random effects,

$\beta_0, \beta_1, \beta_2$ are the fixed effects,

ε_{ij} is the individual − level error term.

Introduction of a Unique Variable to the Model

The introduction of a unique variable, X_{ijz}, where z represents the innovative variable capturing a unique aspect within each cluster. The new model then becomes:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \beta_3 X_{ijz} + u_{0j} + u_{1j}X_{ij1} + \varepsilon_{ij} \qquad [8]$$

Where:

Y_{ij} is the dependent variable for observatjion i in group j

X_{ij1}, X_{ij2} are the predictor variables,

u_{0j} , u_{1j} are group − specific random effects,

$\beta_0, \beta_1, \beta_2$ β_3 are the fixed effects,

the introduction of X_{ijz} allows the model to capture cluster-specific effects associated with the innovative (z) aspect,

ε_{ij} is the individual − level error term.

New Methodology and Combine Hierarchical Model

The unique variable formulated by Larbi et al., (2024) that brings contribution to the field of Hierarchical Bayesian model is found below:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \beta_3 X_{ijz} + \phi_j u_{0j} + \psi_j u_{1j}X_{ij1} + \varepsilon_{ij} \qquad [9]$$

Where:

Y_{ij} is the dependent variable for observatjion i in group j

X_{ij1}, X_{ij2} are the predictor variables,

u_{0j} , u_{1j} are group − specific random effects,

$\beta_0, \beta_1, \beta_2$ β_3 are the fixed effects,

ϕ_j, ψ_j are cluster-specific shrinkage parameters

the introduction of X_{ijz} allows the model to capture cluster-specific effects associated with the novel (z) aspect,

ε_{ij} is the individual − level error term.

Assumptions of Hierarchical Bayesian Shrinkage

Several assumptions are necessary for the estimates of the model to be reliable and valid. Assumptions of Hierarchical Bayesian Shrinkage include:

1. The model assumes a hierarchical structure, where parameters are organized into different levels or clusters (e.g., individual-level, individual, and cluster-level parameters). By allowing for the borrowing of information across groups, this assumption facilitates the generation of more consistent and standardized estimates - an essential valuable feature when there are insufficient observations within particular units or groups.
2. The clusters and levels in the hierarchy are assumed to be exchangeable, which implies that the statistical properties of one cluster are assumed to be identical to any other cluster. The exchangeability assumption is foundational for pooling information across clusters. Information sharing makes it possible to estimate group-specific parameters more accurately by borrowing strength from other clusters when data within the specific cluster is/are limited.
3. The data distributions of the random effects at the cluster level are assumed to be normally distributed; that is the assumption is necessary when using shrinkage methods (Gaussian shrinkage priors). Normality makes the model so easy to control and allows for closed-form solutions. It also facilitates the prediction and interpretation of the hierarchical shrinkage process.
4. The variance of the random effects is homogeneous across all clusters. The assumption of homogeneity is a simplifying device in the model and prevents the shrinkage procedures from overly favouring one particular cluster. This assumption can often be relaxed for more flexible models.
5. Each cluster has enough observations to inform the estimation of group-level parameters adequately. When the number of observations is too small in a cluster or unit, hierarchical shrinkage may not result in much improvement and regularisation; the estimates might be dominated by the prior.
6. Assuming independence among observations within each cluster and across the hierarchies is necessary for maintaining the meaningfulness of the hierarchical structure and accurate differentiation between variability within- versus between groups.

Assumptions Underlying Hierarchical Modelling

For hierarchical Bayesian modelling to perform well, it is essential not to ignore the underlying assumption which provides a concrete foundation. Avoiding the violation of the underlying assumption is by employing diagnostic tools to verify whether the model is devoid of bias estimates or inefficient parameter estimates. The following assumptions underlying hierarchical modelling need to be considered. The observations within each level of the hierarchy are assumed to be independent. This assumption is necessary for the model to accurately estimate the within-level and between-level variation by the researchers doing empirical validation. The random effects at each hierarchy level are assumed to be normally distributed; this assumption allows the model to estimate the mean and variance of the random effects. The relationships between the dependent and explanatory variables, at both the individual and the group levels, are assumed to be linear. This assumption means that the effects of the predictors of the outcome are additive. The variance of the dependent variable is assumed to be constant across different levels of the hierarchy. This assumption ensures that the model captures the variation in the outcome variable. The variance of the random effects at each hierarchy level is assumed to be constant. This assumption implies that the variation between groups is consistent across the hierarchy levels. The random effects at each level are assumed to be normally distributed with a mean of zero. The assumption allows the model to estimate the variation in the outcome variable attributable to the different levels of the hierarchy.

Performance and Interpretability

Table 1 shows the model selection output of which the best two predictors were selected to test the performance of both the standard and unique Bayesian models. A composite variable was formed labelled Contextual Variable Factor (CVF) which is represented as 'X_{ijz}' in the unique model which is one of the variables introduced to the model with insurance, education, household income, marital status, religion and environmental factor as the inclusion variables. The Least Absolute Shrinkage and Selection Operator (Lasso) as a variable selection technique was employed to select the variables that formed the CVF(z) and formation of other variables to contest for the model selection based on their posterior mean values. It is worthy to note that, all the aforementioned variables were used in the formation of z because the variables indicated consistent inclusion across the iterations processes which makes them significant for inclusivity. On model selection criteria, the Hierarchical Bayesian Information Criterion (HBIC) was employed to select Age and Religious Factor from among the factors Socio-economic Factor (SEF), Age, Environmental Factor, and Religious Factor. The HBIC reassesses the accuracy versus complexity trade-off to tackle the centralized hierarchical learning issue and Age and Religious Factor were found to have the lowest HIC Scores respectively.

Table 1; HBIC Model Selection of Maternal Health Indicators

Predictors	HIC
Age	22688.21
Religious Factor	22688.21
Socio-economic Factor (SEF)	22855.84
Environmental Factor	22866.72

Source: R Software output from PMA Data, 2025

Table 2 illustrates the test performance of the standard and the unique models based on their statistical indexes. The test performance is to determine whether the unique model outperforms the standard model using a real-world dataset considering the age and religious factor of mothers influences on the general health of mothers using Performance Monitoring for Action (PMA) dataset, 2022.

Table 2; Test of Performance of the Standard and Unique Model

Indicator	Standard Model	Unique Model	p-value
	Estimate (SE)	**Estimate (SE)**	
REML Criterion	22117.3	22112.8	
Fixed Effects			
Intercept	1.8286 (0.0534)	1.60 (0.0500)	
T - statistic	34.240	32.1	
CVF or ‘X_{ijz}’	Not Present	-0.0505 (0.0127)	< 0.01
Age	0.0874 (0.0112)	0.008 (0.0012)	< 0.01
T - statistic	7.808	6.58	
Religious Factor	0.0311 (0.0096)	0.0282 (0.0096)	<0.05
T - statistic	3.251	2.94	
Correlation	0.57	-0.00514	
Random Effects			
County	*V = 0.0270* (0.1643)	*V = 0.135* (0.57)	
Age	*V = 0.0009* (0.0307)	*V = 0.00328* (0.057)	
Residual	*V = 0.4303* (0.656)	*V = 0.656* (0.810)	

Source: R Software output from PMA Data

It is realised in terms of model fitness in table 2 that, both models have very similar Restricted (Residual) Maximum Likelihood (REML) criteria (22117.3 vs. 22112.8), suggesting comparable fit to the data. The difference of 4.5 magnitude fitness over the standard model portrays how the unique model outperforms the standard one. The rule of thumb stipulates that a REML point of 2 indicates a significant difference. REML fitness of 2 points lower of comparable hierarchical models portrays a significant impact especially with nested models (West et al., (2015); Zuur et al. (2009)). The study revealed that the unique model's Contextual Variable Factor (CVF) is significant with a negative effect unlike the standard Bayesian hierarchical model where the Standard Bayesian model only included maternal age and religion as significant predictors. The CVF provides additional insights regarding how contextual factors affect general health of mothers beyond just age and religion.

Fixed Effects

The slightly higher intercept in the standard model (1.8286) suggests it accounts for higher baseline general health of mothers compared to the model with the contextual variable (i.e. unique model). The Age factor's influence is smaller in the unique variable model (0.008) compared to the standard model (0.0874) although they are all significant. This demonstrates that the inclusion of the contextual variable has absorbed much variability that was initially attributed to the Age factor. The religious factor's effect is slightly reduced in the unique variable model, indicating it has less explanatory power when the contextual variable is included. The contextual variable (CVF) has a statistically significant negative effect, suggesting it plays a pivotal role in shaping general health conditions of mothers, potentially competing with other variables in explaining variance.

Random Effects

The unique Bayesian model (0.135) postulated in table 2 shows greater variability in general health of mothers across counties as against the standard model (0.027). This indicates that the introduction of contextual variable amplifies differences at the county level. The Age Factor exhibits slightly greater variability in the unique variable model (0.00328), suggesting the Contextual Variable Factor interacts with Age factors across counties. Residual variance is larger in the unique variable model (0.656), suggesting that the contextual variable factor introduces complexity, leaving more unexplained variance in the data.

The correlation between the intercept and the Age factor in the unique model is -0.00514, indicating a near-zero relationship. This contrasts with the standard model, where no correlation is explicitly mentioned, implying less interaction between random effects in the standard model.

Comparison and Implications of Performance

The study found that the standard Bayesian model shows smaller variance and residuals values, suggesting that it is more stable and parsimonious. This implies that the standard model provides a more straightforward interpretation (simple model), which is suitable for general applications where contextual variables are not considered and complexity. However, the unique Bayesian hierarchical model introduces higher variability and greater complexity, indicating that the contextual variable explains new dimensions of variance that are not captured by the standard Bayesian model. The contextual variable (X_{ijz}) has a significant negative effect on general health conditions of mothers. Its inclusion shifts the contribution of other factors, Age and religious factors, implying that the 'X_{ijz}' captures some context-specific determinants of general health of mothers that were previously unexplained in the standard model. The unique model highlights the importance of contextual factors, such as marital status, household income, environmental or societal influences, which can be critical for policy-making and targeted interventions at the local level or communities. The results suggest that the unique Bayesian hierarchical model introduced in this study outperformed the standard Bayesian hierarchical model regarding accuracy, interpretability, scalability, and regularisation through shrinkage parameters and contextual variable (X_{ijz}). The negative value from 'X_{ijz}' (-0.0505) even further demonstrates the model's ability to mitigate unobserved heterogeneity, thus demonstrating the model's robustness and reliability**.** The adaptive and Laplace priors activeness in the random intercept's variance (0.135) and Age factor (0.00328) suggest enhancement in estimating group-level variations, proving the model's versatility. The inclusion of both fixed and random effects demonstrates that the unique model can be utilized for analyses on different scales from national to community, overcoming the limitations of standard Bayesian models. The unique hierarchical model characterized in this study mitigates the concerns stated in Zong & Bradley (2023) which pertains to the underspecified model's lack of simplicity and broad applicability to varying datasets. The unique Bayesian hierarchical model introduces a contextual variable (X_{ijz}), providing a mechanism to account for unobserved heterogeneity. This feature simplifies model application by addressing community-specific effects, as shown in the significant and negative 'X_{ijz}' estimate (-0.0505, $p<0.01$) reinforcing the model's purpose. The inclusion of 'X_{ijz}' to the study allows to analyze context-specific phenomena (i.e. healthcare and other socioeconomic interventions), making the model adaptable for different scenarios (i.e. national, urban and rural). Also, Gómez-Méndez & Chainarong (2023) limitation which was the inability of previous models to account for group-specific heterogeneity using cluster-specific random effects and shrinkage parameters. The random effects for county intercepts (0.135) and Age factor (0.00328) demonstrate the model's ability to capture heterogeneity. It was discovered that, the correlation between intercepts and Age factor (−0.00514) highlights the significant relationships observed across groups. This approach enables better regularization of group-specific effects in sparsely populated clusters (Counties), allowing for more precise identification of unique characteristics (i.e. differences between rural and urban counties). Lastly, Porter et al. (2022) limitations which was separately modelling fixed and spatial effects limited the robustness of spatial dependencies in previous models are addressed by integrating both fixed and random effects, the unique model accounts for spatial dependencies and covariate influences simultaneously. The fixed effects (Age and religious factors) capture most essential influences, while random effects account for region-specific effects. This integration mitigates the problem of spatial confounding and enhances the reliability of regional policy interventions, most especially for contextually influenced factors like socioeconomic, healthcare or education access.

Conclusion

The findings confirm that the Unique Bayesian Hierarchical Model is a significant advancement over the Standard Model. By the introduction of shrinkage parameters ϕ_j and ψ_j and a unique contextual variable (X_{ijz}), the Unique Model demonstrated superior performance in terms of model fit, interpretability, and handling heterogeneity. These advancements make the Unique Bayesian hierarchical Model a powerful tool for analyzing hierarchical data across diverse settings. Also, the innovative model (unique model) outperforms the standard model by improving accuracy, interpretability, scalability, and regularization. The study demonstrated that both models generally meet the assumptions for Bayesian hierarchical modelling and shrinkage. The unique Bayesian hierarchical model postulated has the ability to address limitations in prior literature by incorporating contextual variables (X_{ijz}), adaptive shrinkage priors, and a flexible hierarchical structure. These features make the model a powerful tool for analyzing complex relationships and dependencies, particularly in spatial and group-specific contexts. Lastly, combining fixed and random effects improves the model's predictive capacity for health outcomes, making it ideal for counties or regions with sparse data or significant variability.

5.3 Recommendations

Based on the study objectives and findings, the following key interventions are recommended:

- The significant residual variance suggests further research is needed to explore individual factors affecting the health of women in Kenya, such as education, lifestyle choices, or cultural practices that influence maternal health.
- Opinion leaders in the communities and counties must help address contextual factors that influence maternal health and pose challenges such as poverty, housing, etc... which have negative effect on the well being of mothers.
- Encourage community and religious institutions to promote maternal health, particularly in areas with limited healthcare access as this will reduce mortality on the mothers and inborn.
- The model's flexibility ensures it can be applied across diverse datasets, from national surveys to localized studies. This adaptability is crucial for researchers and policymakers aiming to address context-specific health disparities.
- Insights into the contextual variable X_{ijz} allow policymakers to identify and target region-specific needs, such as designing tailored healthcare programs in counties with unique challenges.
- Researchers and practitioners should include contextual variables like z to capture group- or region-specific effects in hierarchical data because of geographical differences. This can improve the identification of critical factors influencing maternal health outcomes in different areas (counties).

- Shrinkage parameters ϕ_j and ψ_j should be utilized in hierarchical models to enhance regularization, reduce overfitting, and improve scalability in large or complex datasets which improve the transparency of a model forecast.

References

1. Aczel, B., Hoekstra, R., Gelman, A., Wagenmakers, E.-J., Kluglist, I. G., Rouder, J. N., et al. (2018). Expert opinions on how to conduct and report Bayesian inference. https://doi.org/10.31234/osf.io/23m7f
2. Agyemang, S., & Asibey, B. O. (2018). Effect of education on health care utilization in rural Ghana: The case of selected communities in the Bekwai Municipality. KNUST Journal of Geography and Development, 114.
3. Ahmed T, Roberton T, Vergeer P, Hansen PM, Peters MA, Ofosu AA, et al. (2022) Healthcare utilization and maternal and child mortality during the COVID-19 pandemic in 18 low- and middleincome countries: An interrupted time-series
4. analysis with mathematical modeling of administrative data. PLoS Med 19(8): e1004070. https://doi.org/10.1371/journal.pmed.1004070
5. Arijit Chakrabarti, Jayanta K. Ghosh, (2011): Philosophy of Statistics. Volume 7 in Handbook of the Philosophy of Science.
6. Auguie, B. (2017). gridExtra: Miscellaneous functions for "grid" graphics. Retrieved from https://CRAN.R-project.org/packa ge=gridExtra
7. Aust, F., & Barth, M. (2018). Papaja: Create APA manuscripts with R markdown. Retrieved from https://github.com/crsh/papaja
8. Bates, D., & Maechler, M. (2019). Matrix: Sparse and dense matrix classes and methods. Retrieved from https://CRAN.R-project. org/package=Matrix
9. Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-efects models using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01
10. Betancourt, M. (2018). A conceptual introduction to Hamiltonian Monte Carlo. arXiv Preprint arXiv:1701.02434.
11. Chatters, L. M., Taylor, R. J., Jackson, J. S., & Lincoln, K. D. (2008). Religious coping among African Americans, Caribbean Blacks and non-Hispanic Whites. *Journal of Community Psychology*, 36(3), 371–386. https://doi.org/10.1002/jcop.20202
12. Cheru Atsmegiorgis Kitabo and Ehit Tesfu Damtie (2020). Bayesian Multilevel Analysis of Utilization of Antenatal Care Services in Ethiopia. Computational and Mathematical Methods in Medicine Volume 2020, Article ID 8749753, 11 pages https://doi.org/10.1155/2020/8749753
13. Danaher, T. S., Sweeney, J. C., & Johnson, L. W. (2007). A Hierarchical Model of Health Service Quality: Scale Development and Investigation of an Integrated Model. Journal of Service Research, 10(2), 123-142. DOI: 10.1177/1094670507309594
14. D'Errigo, P., Tosti, M. E., Fusco, D., Perucci, C. A., & Seccareccia, F. (2007). Use of hierarchical models to evaluate performance of cardiac surgery centres in the Italian CABG outcome study. BMC Medical Research Methodology, 7(29). doi:10.1186/1471-2288-7-29
15. Erica M. Porter, Christopher T. Franck, and Marco A. R. Ferreira (2023). Objective Bayesian Model Selection for Spatial Hierarchical Models with Intrinsic Conditional

Autoregressive Priors. Department of Statistics, Virginia Tech, Blacksburg, VA 24061. International Society for Bayesian Analysis https://doi.org/10.1214/23-BA1375

16. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis (3rd ed.). CRC Press.
17. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer
18. Irving G´omez-M´endez1and Chainarong Amornbunchornvej (2023). Income, education, and other poverty-related variables: a journey through Bayesian hierarchical models. National Electronics and Computer Technology Center (NECTEC). arXiv:2308.16578v1 [stat.ME] 31 Aug 2023.
19. Jiangshan Lai1, Yi Zou, Jinlong Zhang and Pedro R. Peres-Neto (2021). Generalizing hierarchical and variation partitioning in multiple regression and canonical analyses using the rdacca.hp R package. DOI: 10.1111/2041-210X.13800
20. Koenig, H. G. (2012). *Religion, spirituality, and health: The research and clinical implications.* ISRN Psychiatry, 2012, Article ID 278730. https://doi.org/10.5402/2012/278730
21. Larbi, Seth Opoku, Apaka Rangita, and Joyce Otieno. (2024) "Advancing Hierarchical Model: Evaluating Performance, Interpretability and Implications." *International Journal for Multidisciplinary Research (IJFMR)*, vol. 6, no. 6, 2024. www.ijfmr.com.
22. Liang, F., Song, Q., & Yu, K. (2020). Bayesian model selection with applications to hierarchical models. Journal of Computational and Graphical Statistics, 29(2), 326–337.
23. Myrthe Veenman, Angelika M. Stefan, & Julia M. Haaf (2023). Bayesian hierarchical modeling: an introduction and reassessment. Behavior Research Methods https://doi.org/10.3758/s13428-023-02204-3
24. Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, 103(482), 681–686.
25. Qingying Zong and Jonathan R. Bradley (n.d.). Criterion Constrained Bayesian Hierarchical Models. Department of Statistics, Florida State University, 117 N. Woodward Ave, Tallahassee, Fl 32306.
26. Smith, J and Griffith, D and White, A and Baker, P and Watkins, D and Drummond, M and Semlow, A (2020) COVID-19, Equity and Men's Health. International Journal of Mens Social and Community Health, 3 (1). e48-e64. ISSN 2561-9179 DOI: https://doi.org/10.22374/ijmsch.v3i1.42
27. Srivastava, A., Avan, B. I., Rajbangshi, P., & Bhattacharyya, S. (2015). Determinants of women's satisfaction with maternal health care: A review of literature from developing countries. *BMC Pregnancy and Childbirth, 15*(97). https://doi.org/10.1186/s12884-015-0525-0
28. Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267–288.
29. Tredennick, A. T., G. Hooker, S. P. Ellner, and P. B. Adler. 2021. A practical guide to selecting models for exploration, inference, and prediction in ecology. Ecology 102(6):e03336. 10.1002/ecy.3336

West, B. T., Welch, K. B., & Galecki, A. T. (2015). *Linear mixed models: A practical guide using statistical software* (2nd ed.). CRC Press.

30. Xiaokang Zhou , Member, IEEE, Xiaozhou Ye , Kevin I-Kai Wang , Wei Liang , Nirmal Kumar C. Nair, Shohei Shimizu , Zheng Yan , and Qun Jin , (2023). Hierarchical Federated Learning with Social Context Clustering-Based Participant Selection for Internet of Medical Things Applications. IEEE Transactions on Computational Social Systems, Vol. 10, No. 4, August 2023
31. Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society: Series B, 68(1), 49–67.
32. Zhou, H., & Guan, Y. (2018). A hierarchical Bayesian information criterion for model selection in multilevel models. *Bayesian Analysis*, 13(1), 1–25.
33. Zong, L., & Bradley, J. R. (2023). Covariance Penalized Error and model selection for hierarchical Bayesian models. *Statistical Science*, 38(4), 567–588
34. Zou, H. (2006). The adaptive Lasso and its oracle properties. Journal of the American Statistical Association, 101(476), 1418–1429.
35. Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer Science & Business Media.