Journal of Mathematics and Computing Science, ** (*): 131-140, 2025
Copyright © UiTM Press
eISSN: 0128-0767

Components Analysis in Finite Mixture Model

Athirah Nadiah Mohd Tajol Aurus[1] and Phoong Seuk Yen[2*]

[1,2]Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia

Authors' email: mastertyra98@gmail.com, phoong@fsmt.upsi.edu.my*

[*]Corresponding author

Received **; Received in revised **; Accepted ***
Available online ***
DOI: https://doi.org/10.24191/jmcs.*******

Abstract: Determining the appropriate number of components within a finite mixture model is a challenging issue in statistical modelling. This is because the inappropriate choice misinterprets the data and could lead to poor performance. For this purpose, this study will apply empirical data on energy consumption and exchange rate in Malaysia to determine the suitable number of components using the finite mixture model. This study applies Akaike Information Criterion and Bayesian Information Criterion model selection criteria for considering model adequacy. The results demonstrated that the log-likelihood increases as the number of components increases. Indeed, even Akaike Information Criterion still favoured more complex models. However, Bayesian Information Criterion reached a minimum value of three components, which resulted in choosing the most suitable model structure. These results also indicated that Akaike Information Criterion overestimates the number of components. However, Bayesian Information Criterion yielded a more parsimonious and interpretable solution. This study tries to give practical meaning to model selection in mixture modelling. This study also points out the effectiveness of Bayesian Information Criterion in identifying meaningful group structures when the data are heterogeneous.

Keywords: Component, Finite Mixture Model, Akaike Information Criterion, Bayesian Information Criterion, Energy Consumption, Exchange Rate

1 Introduction

Most of the real-world data do not come from a single homogeneous population but from the mix of various underlying subpopulations Shi et al. [1]. These heterogeneous data structures arise in many different fields, including economics, energy studies, biology and finance. The Finite Mixture Model (FMM) offers a flexible statistical modelling framework wherein it is presumed that the overall distribution is a mixture of several component distributions Phoong et al. [2]. Each one represents a distinct subgroup within the data that shares similar characteristics.

FMM can capture such heterogeneity well by estimating each subgroup with an individual distribution. The ability of FMM to estimate different data subgroups is one reason why FMM is extremely useful in areas like market segmentation, genetics and environmental science. According to Ma et al. [3], opine that the classification is stable in multi-centre cohort as finite mixture modelling and k-means clustering both reach the same number of classes.

This is an open access article under the CC BY-SA license (https://creativecommons.org/licenses/by-sa/3.0/).

Additionally, FMM perform well at classification and clustering tasks. FMM are perfect for unsupervised learning applications because of their capacity to recognize discrete subpopulations within a dataset, which enables them to classify fresh observations Wulff et al. [4]. FMM are very helpful in fields like image analysis and consumer segmentation because of their capacity to simulate the actual distribution of data which produces more precise and significant clustering findings.

One of the most important and difficult tasks in the application of the FMM is the proper selection of the number in components Manole and Khalili [5]. Too few may result in underfitting and hence loss of important structure present in the data. Many data will lead to overfitting and poor generalization. FMM increasingly exploited for modelling unknown distributional shapes and for identifying group structures McLachlan and Peel [6]. Hence, identification of the proper number of components will help to achieve good estimation, accurate interpretation and meaningful classification of data.

Several statistical criteria have been developed, include the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). In context of model selection, Akaike Information Criterion and Bayesian Information Criterion are widely used Zhang et al. [7]. AIC, introduced by Hirotugu Akaike, combines estimation with structural and dimensional determination Cavanaugh and Neath [8]. BIC, also known as the Schwarz criterion, includes a penalty term for the number of parameters, which is larger than in AIC, to prevent overfitting Dziak et al. [9]. These criteria may give different results depending on data characteristics, sample size and the presumed distribution form of components. In practice, it is often advisable to consider multiple criteria simultaneously, as the choice of criterion can significantly impact results.

In the following example, the FMM will be applied to two different datasets which are energy consumption and exchange rate. This is used to demonstrate how to identify the number of components in real data. The datasets represent typical examples of economic and energy-related variables that often show multimodal behaviour or hidden group structures. The research objectives to show how various model selection criteria perform to identify the number of mixture components for each dataset.

This paper highlighted five sections. Section 1 is on introduction. Section 2 reviews related literature with respect to finite mixture models and methods of determining the number of components. Section 3 outlines the methodology used in model estimation and model selection. The empirical results and discussions using the two datasets shows in Section 4. Finally, Section 5 concludes the study with recommendations for future research.

2 Literature Review

Finite Mixture Model is one of the most flexible statistical tools for analyzing data originating from multiple underlying sources or subpopulations Kim and Mokhtarian [10]. Rather than assuming that all observations are emanating from one homogeneous group, the FMM considers the probability that the data have been generated from a number of distinct groups Bourouis [11]. However, each represented by its own probability distribution. Such modelling offers an effective way to describe complex data structures that may exhibit multimodality, skewness, and hidden heterogeneity.

Mixture distribution was initially addressed by Pearson [12], fitting a combination of normal distributions to data that was not well fitted by a single distribution. Since then, the mixture model framework has developed into a well-established area in modern statistics. Applications range from economics and biology to engineering and social sciences. In the last few years, the method has become very popular for clustering analysis, density estimation, and pattern recognition due to its flexibility in detecting hidden structures in data Scrucca [13].

Estimating the proper number of components is one of the major challenges in using finite mixture models Wang and Yang [14]. In most practical applications, the true number of underlying groups in a dataset is unknown Miller and Harrison [15]. If there are too few components, underfitting might result, which will not do justice to significant data structures. Overfitting can lead when many components occur in mixture model Van Havre et al. [16]. The model becomes unnecessarily complicated and may not generalize well to new data. Thus, determining an appropriate number of components becomes a very important task that has a great impact on the interpretability and reliability of the results Hunter [17].

To handle this issue, various model selection criteria have been suggested by statisticians that balance model fit with the complexity of the model. Probably the two most popular are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) Shatwan et al. [18]. While both of these methods fundamentally work by penalizing over-fitting, AIC and BIC apply the penalty in a different way. AIC works out a penalty that grows linearly with the number of parameters, while BIC imposes a stronger penalty dependent on sample size Lubke et al. [19]. A consequence of this is that AIC tends to favor more complex models, while BIC favors simpler ones.

The AIC was indeed developed by Akaike [20] from the standpoint of information theory. It essentially measures a trade-off between goodness of fit and model complexity by estimating the information lost when a given model is used to represent the data-generating process. Since AIC development, the criterion has gained extensive use in many empirical applications due to its simplicity and the ease with which it allows for model selection Sutherland et al. [21]. Nevertheless, one of its known limitations is that it tends to favor models with more parameters, especially when the sample size is small or when the components in the mixture model overlap substantially Grimm et al. [22].

BIC, proposed by Schwarz [23], adopts a more conservative choice, including a heavier penalty term, increasing with the number of parameters and with the sample size. As a result of the property mentioned above, BIC is consistent in finding the true number of components as the sample size increases Gohain and Jansson [24]. In practice, BIC has been found to work well in large-sample situations or when the component distributions are well separated. Because of this, it has become a preferred criterion for many statistical software packages and applied research studies involving mixture models Nguyen and Nguyen [25].

Several studies have compared the performance of AIC and BIC in mixture modelling. Among the early works are those by McLachlan and Peel [6] and Phoong and Ismail [26], which demonstrated that both criteria perform well but that their relative performance strongly depends on data characteristics. According to Xu et al. [27], when sample size is small, the AIC value tends to outperform when components overlap, whereas BIC tends to provide more accurate results in large sample sizes also with well-separated components. These differences in performance indicate that one should take into both context of the data and the goals of the analysis when selecting a criterion.

The AIC and BIC performance using simulation studies and applications involving real data also being studied by Chakrabarti and Ghosh [28]. For instance, studies dealing with economics and environmental science indicate that AIC may define several subgroups within a highly variable dataset, while BIC tends to simplify the model by choosing fewer groups. Such a feature expresses the intrinsic difference between the two criteria. AIC is more sensitive to small variations of the data, while BIC puts greater emphasis on the interpretability and parsimony of the model Grimm et al. [22]. So, mixture modelling has been very useful in many applications and real-world datasets Zhao et al. [29].

In this study, the energy consumption and exchange rate data are used to find the latent group structures or distributional patterns without homogeneity assumptions. These datasets are applied due to the reason that both variables usually have a high level of variability, multimodality and possible structural breaks in time series. The conditions under which mixture modelling analysis may be ideal. Energy consumption data will typically reflect various behavioral such as seasonal changes, industrial demand shifts and changes in efficiency levels. Similarly, exchange rate data often contain multiple underlying processes. For example, market interventions, economic shocks and policy adjustments that leading to heterogeneous distributions. These two kinds of data can thus shed light on clearly different subpopulations or hidden regimes that are not visible under traditional single-distribution assumptions by using the finite mixture model. Therefore, these datasets serve to provide practical illustrations of how component analysis can capture the real-world underlying complexity of economic and energy related phenomena.

Component analysis is a necessary procedure. This due to understand the underlying heterogeneity of complex datasets and to find distinct subpopulations. Furthermore, latent structures that cannot be modeled using traditional single-distribution models. Finite mixture models allow researchers to break down data into meaningful groups. This can be show that each representing a unique statistical behavior and characteristic pattern. Indeed, this process can help reveal otherwise hidden variability. Moreover, enhance the interpretability of a model by relating the statistical components to real-world phenomena. Accurate component analysis has a positive effect on final decisions. This because since can recognize different regimes and segments within overall outcomes. For example, energy consumption patterns or an exchange rate behavior. Without adequate component analysis, data might turn out to be homogeneous and could lead to oversimplified interpretation with the risk of misrepresenting the real structure. Therefore, component analysis is important for extracting useful information, guiding policy formulation, and enhancing model accuracy in general applied statistical research.

The AIC and BIC are the most widely used and practical tools that used for component analysis. Both criteria have different theoretical grounds and operational features. Thus, may lead to different conclusions depending on data characteristics. This invites in-depth assessment of their performance using real datasets, such as energy consumption and exchange rate data, which illustrates how each performs in actuality. This also will provide empirical guidance for future applications. As for the section 3 will explain details about finite mixture model, model selection criteria for AIC and BIC also with data description.

3 Methodology

A Finite Mixture Model

In a finite mixture model, it is assumed that the data are obtained from many mixture distributions. Instead of fitting a single distribution to the general observation, FMM allow each data point belong to one of several latent groups or components with their respective parameters Moore et al. [30]. This provides a flexible way to model heterogeneous data with multiple modes or varying variances.

FMM is useful in econometrics area and also used to estimate unobserved heterogeneity which plays important roles in industrial organization, labor economics, exchange rates and other fields Hao and Kasahara [31]. This means that a FMM with function valued unobserved heterogeneity can be identified in a cross-section setting, without assuming the pattern of dependence between the regressor and unobserved heterogeneity Kitamura and Laage [32].
The general equation of a finite mixture model can be expressed as:

$$f(x|\theta) = \sum_{k=1}^{K} \pi_k f_k(x|\theta_k) \tag{1}$$

Where $f(x|\theta)$ is the probability density function for overall of the data. π_k of k-th component is the mixing proportion, with $\sum_{k=1}^{K} \pi_k = 1$ and $0 \leq \pi_k \leq 1$. Additionally, $f_k(x|\theta_k)$ is the probability density function for k-th component that parameterized by θ_k. The number of components (subpopulations) in the mixture is K. The observed data points represent x. This equation shows that the likelihood of any observed data point x is a weighted sum of the likelihoods under the different components $f_k(x|\theta_k)$, where the weights are the mixing proportions π_k.

In this study, the finite mixture model framework is applied to model the distribution of energy consumption and exchange rate data. Each dataset is analyzed independently to determine how many components best describe its distribution. The assumption is that the data may arise from several hidden regimes or states, such as different economic conditions or consumption behaviors.

By estimating mixture models with various numbers of components, the study aims to identify the value of K that offers the best balance between model fit and complexity. The selection of the optimal number of components using information criteria such as AIC and BIC.

B Model Selection Criteria

The step of determining the number of components K in a mixture model is crucial. Poor model performance can lead a wrong choice. For example, selecting a smaller number of components can result in underfitting while too many components can cause overfitting. Therefore, the model selection process is necessary to employ and guide statistical criteria.

Akaike Information Criterion (AIC) together with Bayesian Information Criterion (BIC) are two most commonly used criteria. Both criteria are based on the maximum log-likelihood value. AIC and BIC adjusted by a penalty term for model complexity. The most appropriate model of AIC and BIC are when in the smallest value

The Akaike Information Criterion (AIC) is defined as:

$$AIC = 2k - 2 \ln (L) \tag{2}$$

Where L is the maximum likelihood of the model. For k is the total number of estimated parameters. AIC aims to select the model with best approximates for true data-generating. This can be by the process of balancing model fit and parsimony.

The Bayesian Information Criterion (BIC) is defined as:

$$BIC = \ln(n)\, k - 2 \ln (L) \tag{3}$$

Where n is the number of observations. As the sample size increases, BIC imposes a stronger penalty for model complexity compared to AIC. However, BIC tends to favor simpler models. While AIC may select models with more components.

In this study, both AIC and BIC are computed for mixture models fitted with varying numbers of components. The optimal model is selected based on the lowest value of each criterion. The results are then compared to evaluate whether AIC and BIC agree or differ in their suggested number of components for each dataset.

C Data Description

The datasets used in this study consist of two time series which are energy consumption and exchange rate. These data are selected to show how finite mixture model can determine underlying components in real-world data. Each data set is explored for its distributional features, rather than investigating any causal relationship among these data sets.

These are energy consumption data which, depending on the time period considered, may reflect variations due to industrial demand, changes in policy, or seasonal effects. Many such data sets contain nonlinear trends and multiple modes, and hence can be modeled with mixtures. The aim is to identify whether the energy consumption data can be better represented by multiple sub-distributions.

The exchange rate data reflect the value of one currency relative to another and are known to exhibit high volatility and regime changes. Periods of stability and sudden fluctuations may lead to multimodal distributions, which can be effectively modeled using a finite mixture framework. Identifying the number of regimes or underlying components can provide insights into the behavior of exchange rate movements.

In summary, the datasets serve as practical examples for applying the finite mixture model and for comparing AIC and BIC in identifying the optimal number of components. The next section presents the empirical results, including parameter estimates, model selection outcomes, and interpretation of the findings for both energy consumption and exchange rate data.

4 Results and Discussions

A Identifying the Number of Components

The main goal of this analysis is to determine the best number of components that describes the underlying data distribution. Different models were fitted in each with a different number of components ranging from one to five. For each fitted model, the log-likelihood, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were computed to evaluate model fit yet parsimony using Python software.

The most important results of the model fitting are summarized in Table 1. Firstly, the number of components in this study is from 1 to 5. Then, the log-likelihood value from 6770.255 to 8934.673. The value shows it increases correspondingly. That is expected, as increasing the number of components allows the model to better approximate data. At the same time, a higher log-likelihood does not necessarily mean the model is superior because it does not account for model complexity. Due to this fact, both AIC and BIC were used to penalize the increase in parameters in order to identify the most appropriate model structure.

Table 1 Information based Model Selection Criteria

Number of Components	Number of Parameters	Log-likelihood	AIC	BIC
k = 1	6	6770.255	-13528.511	-13498.205
k = 2	13	8814.665	-17603.330	-17537.668
k = 3	20	8912.034	-17784.067	-17683.047
k = 4	27	8926.019	-17798.037	-17661.661
k = 5	34	8934.673	-17801.345	-17629.612

The values of AIC show improvement as the number of components increases. Indeed, AIC decreases from –13,528.511 at 1 to –17,801.345 at 5. It means that the model fits better when more components are added, according to AIC. As AIC tends to overestimate the number of components, since it favors complex models when extra parameters yield marginal improvements in likelihood, further verification of this through BIC is needed.

On the contrary, BIC has a different tendency. BIC quickly decreases from –13,498.205 at 1 to –17,683.047 at 3 while reaching a minimum at this point. Then, increases slightly for models with four and five components. That means even though the fit improves marginally by increasing the number of components beyond three, this gain does not justify additional model complexity given the BIC criterion. Therefore, BIC selects the three-component model as being the best representation of the data.

The sharp difference between AIC and BIC in the present analysis is revealing of the theoretical features of the two criteria discussed earlier. While AIC rewards goodness of fit and only lightly penalizes model complexity, it might turn out to be too sensitive to minor changes in data. BIC imposes a more robust penalty, whose value increases with sample size and number of parameters. For this reason, selects simpler and more parsimonious models. While AIC selects the five-component model as the best, BIC identifies the three-component model as more interpretable and more efficient.

Further support for the BIC choice is given by a closer look at the log-likelihood improvement. The log-likelihood only increases from 8912.034 at 3 to 8934.673 at 5. This also increase about 22.639 units for 14 added parameters. Such a slight improvement indicates that the extra components add little to explain the underlying structure of the data. In this respect, from both a practical and an interpretative point of view, the three-component model seems to provide a good balance between fit and simplicity.

From the point of view in modelling, the three-component solution might correspond to distinct behavioral regimes within the series. For example, in energy consumption these could be low, moderate, and high usage periods, determined by economic or seasonal factors. Likewise, for exchange rate data, the three components might capture conditions of stability, moderate volatility and high volatility in the markets. Such interpretations are intuitive and consistent with the idea of mixture models finding latent subpopulations or regimes in complex economic and environmental data.

When the number of components increased, the convergence took more iterations, as there were more parameters. Therefore, all models reached a stable log-likelihood value. This can guarantee that the estimation process was numerically stable. This consistency adds confidence that the observed differences between AIC and BIC are meaningful yet not due to computational instability.

Comparing these two datasets leads to the following interpretation which mixture models represent a flexible tool in discovering hidden structures in heterogenous data. While energy consumption probably undergoes gradual transitions between levels. However, exchange rates might show abrupt regime changes. The three-component model represents such distinctions nicely and shows that FMM can account for smooth as well as discrete forms of heterogeneity within one consistent framework.

In other words, the conclusion is that whereas the AIC criterion favored a five-component model, the BIC criterion did favor the three-component model as the best fit to describe data. Now, considering the trade-offs underlined above between fit and parsimony also with interpretability in applied contexts, the three-component model is preferred. These findings were indeed in agreement with the theoretical considerations of the previous sections. This study underlined the belief that BIC provides generally a more reliable criterion for estimating the true number of components in finite mixture models.

5 Conclusions

Component analysis is crucial in making sure that the finite mixture model fits the real underlying structure data. Determining the right number of components also particularly relevant since it can affect the reliability, interpretability and performance of the model directly. A good component structure provides a reasonable balance between model bias and overfitting while capturing meaningful heterogeneity without unnecessary model complexity. The results of this study have shown that selecting the appropriate number of components can lead to a greater degree of precision in statistical analysis. Moreover, realistic patterns of real-world data such as energy consumption and exchange rates. The appropriate analysis of components provides more robust decision-making, enhances model validity and makes the finite mixture model particularly strong for practical applications in numerous empirical research areas.

The framework of a FMM was adopted because it offers a flexible statistical representation of heterogeneity in data. This means FMM cannot be appropriately modeled using a single distribution.

The analysis showed how information criteria, especially AIC and BIC can be applied in choosing the best number of components in a mixture model framework. Therefore, beyond the three components, the increase in log-likelihood became marginal whereas the complexity of the model increased significantly. AIC favored the five-component model but BIC identified the three-component model. This is because it penalized complexity more.

This can reflect back with the underlying theory and the actual behavior of AIC and BIC. AIC has a theoretical tendency to favor model fit for moderate sample sizes. Hence, tending to select models with more parameters. In contrast, BIC involves a more substantial penalty increasing with the number of parameters and observations. Thus, it promotes simpler and more interpretable models. This study find that BIC gave a better balance in this study by selecting the three-component model. This because, that suits well with the underlying structure in the variation of the data.

The three-component model provided an interpretable representation of the latent structure in the data. For energy consumption, the components might be considered as representing the low, medium, and high consumption regimes that would be determined by seasonal or economic factors. For exchange rates, the components may represent distinct market conditions, such as stable, moderately volatile and highly volatile periods. This finding illustrates the ability of finite mixture models to capture complex data behavior and reveal hidden subgroups within observed phenomena.

Altogether, these results point to a practical importance of model selection criteria when performing mixture modelling. The comparison between AIC versus BIC provided meaningful insights into how each criterion behaves under real-world data applications. This can show that, AIC is able to detect subtle variations in the data. For BIC guarantees that the final model remains parsimonious and interpretable. Therefore, the use of both criteria jointly is recommended in interpreting the results by researchers with regard to the research context and characteristics of data.

From a methodological perspective, this paper confirms that the FMM and information criteria, are excellent analytical technique for identifying data heterogeneity. The applied procedure is particularly valuable for research areas like energy economics, finance and environmental modelling, where multiple data structures are common. The framework developed in this paper can be explored for other variables and domains of interest where there may be latent subpopulations or dynamic regimes present.

In summary, the present study contributes to a better understanding of how to identify the appropriate number of components in FMM. The results confirm once again that information-based criteria, such as AIC and BIC, provide critical help in model selection. However, BIC yielded a more consistent and interpretable choice by pointing out a three-component model as the best representation of the data. Future studies could further investigate other selection criteria or Bayesian approaches that would allow validation and enhancement of the robustness of mixture modelling in diverse applications.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or private sectors.

Conflict of Interest Statement

The authors agree that this research was conducted in the absence of any self-benefits, commercial or financial conflicts and declare the absence of conflicting interests with the funders.

References

[1] X. Shi, Z. Pan, and W. Miao, "Data integration in causal inference, "*Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 15, no. 1, 2022, doi: 10.1002/wics.1581.

[2] S. Y. Phoong, S. L. Khek, and S. W. Phoong, "The Bibliometric Analysis on Finite Mixture model," SAGE Open, vol. 12, no. 2, 2022, doi: 10.1177/21582440221101039.

[3] P. Ma et al., "Individualized resuscitation strategy for septic shock formalized by finite mixture modelling and dynamic treatment regimen, "*Critical Care*, vol. 25, no. 1, p. 243, 2021, doi: 10.1186/s13054-021-03682-7.

[4] P. Wulff, M. Kubsch, and C. Krist, "Basics of Machine Learning, "in *Applying Machine Learning in Science Education Research*, Springer Texts in Education, 2025, doi: 10.1007/978-3-031-74227-9_2.

[5] T. Manole and A. Khalili, "Estimating the number of components in finite mixture models via the Group-Sort-Fuse procedure," *The Annals of Statistics*, vol. 49, no. 6, 2021, doi: 10.1214/21-aos2072.

[6] G. McLachlan and D. Peel, *Finite mixture models*, Wiley series in probability and statistics, vol. 6, no. 1, pp. 355–378, 2000, doi: 10.1002/0471721182.

[7] J. Zhang, Y. Yang, and J. Ding, "Information criteria for model selection," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 15, no. 5, 2023, doi: 10.1002/wics.1607.

[8] J. E. Cavanaugh and A. A. Neath, "The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements, "*Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 11, no. 3, 2019, doi: 10.1002/wics.1460.

[9] J. J. Dziak, D. L. Coffman, S. T. Lanza, R. Li, and L. S. Jermiin, "Sensitivity and specificity of information criteria," *Briefings in Bioinformatics*, vol. 21, no. 2, pp. 553–565, 2020.

[10] S. H. Kim and P. L. Mokhtarian, "Finite mixture (or latent class) modelling in transportation: Trends, usage, potential, and future directions," *Transportation Research Part B: Methodological*, vol. 172, pp. 134–173, 2023, doi: 10.1016/j.trb.2023.03.001.

[11] S. Bourouis, "Recent Advances in Statistical Mixture Models: Challenges and Applications," in *Proc. 12th Int. Conf. Pattern Recognition Applications and Methods (ICPRAM 2023),* 2023, pp. 312–319, doi: 10.5220/0011660900003411.

[12] K. Pearson, "Contributions to the mathematical theory of evolution," *Philosophical Transactions of the Royal Society of London A*, vol. 185, pp. 71–110, 1894, doi: 10.1098/rsta.1894.0003.

[13] L. Scrucca, "A model-based clustering approach for bounded data using transformation-based Gaussian mixture models," arXiv:2412.13572, 2024, doi: 10.48550/arxiv.2412.13572.

[14] C. Wang and Y. Yang, "Estimating the number of components in finite mixture models via variational approximation," arXiv:2404.16746, 2024, doi: 10.48550/arxiv.2404.16746.

[15] J. W. Miller and M. T. Harrison, "Mixture models with a prior on the number of components," *Journal of the American Statistical Association*, vol. 113, no. 521, pp. 340–356, 2017, doi: 10.1080/01621459.2016.1255636.

[16] Z. Van Havre, N. White, J. Rousseau, and K. Mengersen, "Overfitting Bayesian mixture models with an unknown number of components," *PLoS ONE*, vol. 10, no. 7, e0131739, 2015, doi: 10.1371/journal.pone.0131739.

[17] D. R. Hunter, "Unsupervised clustering using nonparametric finite mixture models," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 16, no. 1, 2023, doi: 10.1002/wics.1632.

[18] A. B. Shatwan, A. Abdalla, H. Mohammed, E. B. Ismaeil, and A. M. Mami, "On model selection criterion for finite Gaussian mixture models," *International Journal of Sciences: Basic and Applied Research (IJSBAR),* 2024, [Online]. Available: https://www.gssrr.org/JournalOfBasicAndApplied/article/view/16792.

[19] G. H. Lubke et al., "Assessing model selection uncertainty using a bootstrap approach: An update," *Structural Equation Modelling: A Multidisciplinary Journal*, vol. 24, no. 2, pp. 230–245, 2017, doi: 10.1080/10705511.2016.1252265.

[20] H. Akaike, "A new look at statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–722, 1974.

[21] C. Sutherland et al., "Practical advice on variable selection and reporting using Akaike information criterion," *Proceedings of the Royal Society B*, vol. 290, no. 2007, p. 20231261, 2023, doi: 10.1098/rspb.2023.1261.

[22] K. J. Grimm, R. Houpt, and D. Rodgers, "Model fit and comparison in finite mixture models: A review and a novel approach," *Frontiers in Education*, vol. 6, 2021, doi: 10.3389/feduc.2021.613645.

[23] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, pp. 461–464, 1978.
[24] P. B. Gohain and M. Jansson, "Scale-invariant and consistent Bayesian information criterion for order selection in linear regression models," *Signal Processing*, vol. 196, p. 108499, 2022, doi: 10.1016/j.sigpro.2022.108499.
[25] H. D. Nguyen and T. Nguyen, "Modifications of the BIC for order selection in finite mixture models," arXiv:2506.20124, 2025, doi: 10.48550/arxiv.2506.20124.
[26] S. Y. Phoong and M. T. Ismail, "A comparison between Bayesian and maximum likelihood estimations in estimating finite mixture model for financial data," Sains Malaysiana, vol. 44, no. 7, pp. 1033–1039, 2015, doi: 10.17576/jsm-2015-4407-16.
[27] S. Xu, M. A. R. Ferreira, and A. N. Tegge, "What is in the model? A comparison of variable selection criteria and model search approaches," arXiv:2510.02628, 2025. [Online]. Available: https://arxiv.org/abs/2510.02628.
[28] A. Chakrabarti and J. K. Ghosh, "AIC, BIC and recent advances in model selection," in *Elsevier eBooks*, 2011, pp. 583–605, doi: 10.1016/b978-0-444-51862-0.50018-6.
[29] H. Zhao et al., "Probabilistic mixture model driven interpretable modelling, clustering, and predicting for physical system data," *Engineering Applications of Artificial Intelligence*, vol. 160, p. 112069, 2025, doi: 10.1016/j.engappai.2025.112069.
[30] E. W. G. Moore, A. Quartiroli, and T. D. Little, "Introduction to the best practice recommendations for longitudinal latent transition analysis," *International Journal of Psychology*, vol. 60, no. 2, p. e70021, 2025, doi: 10.1002/ijop.70021.
[31] Y. Hao and H. Kasahara, "Estimating the number of components in panel data finite mixture regression models with an application to production function heterogeneity," arXiv:2506.09666, 2025, doi: 10.48550/arxiv.2506.09666.
[32] Y. Kitamura and L. Laage, "Nonparametric analysis of finite mixtures," arXiv:1811.02727, 2018, doi: 10.48550/arxiv.1811.02727.