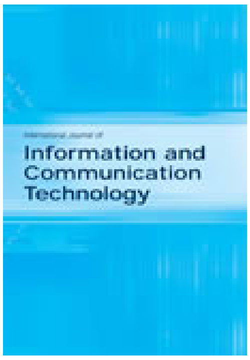


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Assessing chorus size effects on art music performance quality using hesitant bipolar fuzzy multi-criteria decision-making

Qiao Weilong

DOI: 10.1504/IJICT.2025.10071787

Article History:

Received:	18 April 2025
Last revised:	14 May 2025
Accepted:	15 May 2025
Published online:	10 July 2025

Assessing chorus size effects on art music performance quality using hesitant bipolar fuzzy multi-criteria decision-making

Qiao Weilong

Changchun Medical College,
Changchun, Jilin – 130031, China
Email: qiaoweilong123@126.com

Abstract: In this paper, a novel approach is given to evaluate the effect of chorus size on the quality of art music performance using the hesitant bipolar fuzzy multi-criteria decision-making (HBFS-MCDM) framework. The method accounts for positive and negative expert evaluations under conditions of hesitancy concerning five performance criteria: tonal balance, articulation precision, dynamic range, audience cohesion, and emotional expression. To aggregate, expert conductors and musicologists evaluated a hesitant bipolar fuzzy number, and then a weighted decision model was proposed. Results show limited evidence of balanced and consistent performance among all evaluated dimensions across various ensemble sizes. The HBFS-MCDM approach offers a robust and expressive tool for supporting decisions in professional chorals and for actionable insights on ensemble configuration.

Keywords: choral performance evaluation; chorus size; hesitant bipolar fuzzy sets; multi-criteria decision-making; MCDM; art music analysis; subjective judgment modelling.

Reference to this paper should be made as follows: Weilong, Q. (2025) 'Assessing chorus size effects on art music performance quality using hesitant bipolar fuzzy multi-criteria decision-making', *Int. J. Information and Communication Technology*, Vol. 26, No. 23, pp.114–127.

Biographical notes: Qiao Weilong is a member of Changchun Medical College, Changchun, Jilin, China. He is affiliated with medical education and research within the institution. His academic interests likely align with clinical training and healthcare sciences.

1 Introduction

Since art music performance quality is inherently multi-dimensional and subjective, evaluating performance quality in these contexts poses special challenges (Li et al., 2022). Nuanced auditory perception and interpretive judgments across criteria of tonal Balance, articulation precision, dynamic range, ensemble cohesion, and emotional expression all factor into performance assessments (Linaberry, 2021). Chorus size is one of the most fundamental and least investigated variables affecting the performance result. Conductors and music directors usually adjust ensemble size according to the demands and repertoire that can be accommodated. Still, such decisions are frequently made by experience or tradition rather than systematic evaluation (Zhang et al., 2024; Halvasi, 2018). A data-informed and structured approach to understanding how chorus size affects overall performance quality could significantly impact ensemble planning and artistic outcomes.

Most existing assessment methods in music performance are based upon qualitative scoring or experts' single-point numerical rating (Berenzweig et al., 2004). These approaches are helpful but do not allow for the expressive power necessary to capture the hesitation and conflicting impression typical of artistic evaluation (Haroutounian, 2000; Stanley et al., 2002). To take a small example, the expert may like the richness of sound of a large choir but dislike its lack of precision (Ternström, 2003; Ternström and Sundberg, 1988). Such bipolarity and uncertainty do not fit into traditional evaluation frameworks, which simplistically evaluate the expert's judgment and depth of performance analysis.

Without such a methodological gap, this paper presents a hesitant bipolar fuzzy multi-criteria decision making (HBFS-MCDM) model for evaluating the effects of chorus size on the performance quality of art music. Hesitant bipolar fuzzy sets (HBFS) provide a means to express evaluations of more than one degree of satisfaction and dissatisfaction for any criterion while considering the hesitance and ambiguity in assessments. When incorporated in a multi-criteria decision-making (MCDM) framework, this model allows the systematic aggregation of complex expert opinions to develop a coherent ranking of the alternatives. This study evaluates five performance criteria of small, medium, and large chorus size by a panel of expert conductors and musicologists across three chorus size configurations. For each evaluation, hesitant bipolar fuzzy numbers are formed, and the final decision is obtained with a weighted aggregation method, in which experts first assign importance to the criteria. The contributions of this study are as follows:

A novel HBFS-MCDM model was developed to evaluate choral performance quality that simultaneously considers satisfaction and dissatisfaction with evaluator hesitation. First application of HBFSs to multi-criteria evaluation in choral music. Linguistic transformation of the expert assessments into hesitant bipolar fuzzy linguistic structured data about five primary program performance criteria. An interpretable aggregation model has been designed based on weighted summation, where the weights are obtained from the experts' Delphi method consensus. Validate the model's robustness under changing criterion weight scenarios with a comprehensive sensitivity analysis. Guidance on practical insights necessary for optimal chorus size selection for conductors, program directors, and educational institutions.

The rest of the paper is organised as follows: a literature review in Section 2 reveals the performance evaluation models, fuzzy decision-making models, and hesitant bipolar fuzzy theory. Section 3 details the methodology, such as the definition of evaluation criteria, its formulation into the HBFS-MCDM model, and the aggregation process, as well as the results and analysis of chorus size alternatives in Section 4. The theoretical and practical implications of the findings are considered in Section 5. The study is concluded in Section 6 with a summary of some key learnings, and Section 7 outlines the directions for future research. Acknowledgments are given in Section 8, with the complete list of references.

2 Literature review

The evaluation of musical performance, particularly in ensemble contexts such as choral music, poses unique methodological challenges due to artistic interpretation's inherently subjective and multi-dimensional nature. Within this section, the review addresses the literature in choral and art music performance quality evaluation, fuzzy logic, MCDM approach in the music domain, and any other qualitative domain, as well as HBFS-based complex decision modelling. This review provides the basis for the rationale for applying an HBFS-MCDM approach to the problem of sizing a chorus to determine how chorus effects on performance quality depend upon chorus size.

2.1 Performance quality evaluation in choral and art music

Choral music evaluation work, until now, has been based on sound, pedagogical interpretive, acoustic, instrumental, and expressive factors, such as tonal blend, balance, rhythmic precision, diction, intonation, and expressiveness (Adongo, 2020; Batovska et al., 2023; Clark, 2024). Some studies that studied how the group vocal performance experience and auditory perception matter are those of Ternström et al. (2018) and Weiss and Trehub (2023). Some have empirically studied ensemble configurations, such as spatial positioning and acoustic impact of choirs; however, there is relatively scant work on the effect of chorus size on multi-dimensional performance quality (Yarnall, 2017). Ensemble size is often selected by conductors based on repertoire requirements (Fischer, 2024), but no formal evaluation methods exist to systematically compare various configurations on several qualitative criteria.

Most of the few quantitative efforts in this area have used subjective rating scales or rubrics, where evaluators often have to reduce the complex impressions into a single number per criterion (Gordon, 2002). However, few of these approaches mention evaluator hesitation, instances where a judge might hesitate between two or more ratings or conflicting sentiments between two opinions, pleasure, and displeasure (Becker, 2004). For example, an evaluator may consider a large choir to be powerful but imprecise, and the traditional scoring techniques demand that they find a middle ground to flatten out this duality (Latimer, 2009). Such limitations lend themselves to the need for evaluation models that more accurately portray the uncertainty and nuance of expert assessments in musical contexts.

2.2 Fuzzy logic and MCDM in music evaluation

With the growing existence of imprecision and subjectivity in many decision support systems, fuzzy logic technology is more widespread. It is applied in education, health care, and the performing arts (Mahfouf et al., 2001). Fuzzy MCDM techniques have been used in music evaluation for problems of instrument selection, hall acoustics assessment, and conductor ranking (Chong and Lalla, 2020). Fuzzy AHP has been used by Chong and Lalla (2020) to evaluate classical instruments regarding tonal quality and cultural value. One such application of a fuzzy TOPSIS model to assess the performance of an orchestral conductor concerning technical and interpretative criteria (Pouyakian et al., 2024). Fuzzy MCDM techniques to select the best location for concert halls according to acoustic features (Chong and Lalla, 2020).

However, these models successfully structure the evaluation process by making a few assumptions: each judgment is represented as a crisp, fuzzy number or a linguistic term. However, imitation learning does not hold in artistic domains where evaluators are uncertain or partially agree with multiple linguistic terms. In addition, traditional fuzzy MCDM models are open, only representing the degree of satisfaction or preference. In real-world musical evaluation, experts are willing to approve of some aspects and disapprove of others, for example, appreciation of a choir's dynamic control and lack of ensemble cohesion (Yang, 2022). Unipolar models cannot accurately represent the contradictory sentiments (from a representative system point of view) we observe.

2.3 HBFSs: theory and applications

They introduced the HBFS, later extended by Asaad et al. (2024) and Mahmood et al. (2024), which forms a robust mathematical model to deal with fuzzy, complex, ambiguous, and multidirectional judgments. An HBFS generally deals with multiple positive membership degrees (satisfaction) and multiple negative ones (dissatisfaction) for each decision alternative for a given criterion (Mandal and Ranadive, 2019; Hightower, 2014; Rodler, 2020). HBFS is especially powerful in its dual structure, where decision-makers tend to be hesitant (tend not to commit to a single judgment) and can have bipolarity (there are pros and cons simultaneously).

HBFS has been used in applications with nuanced tradeoffs, e.g., in policy decision-making, risk assessment, and human resource management. This method nicely supports decision-makers in representing arguments for and against, and their subjective nature in a single evaluation space, suitable for such problems with high subjectivity and interpretive complexity (Asaad et al., 2024; Lotfi et al., 2023; Waqas et al., 2024). Yet HBFS has not seen much actual implementation in the performing arts, and to date, there has been no prior practice of applying it to ensemble performance in music.

Within this study, HBFS is an ideal tool to develop a model of the hesitancy and polarity observed in expert evaluation of chorus size. For example, the balanced tonal expert gives {0.6, 0.7} positive and {–0.3, –0.4} negative ratings for cohesion and hesitant ratings for both, while the cohesion expert gives {0.8, 0.9} positive and {–0.1, 0} negative ratings for tonal balance and {–0.6, –0.7} negative and {0.6, 0.7} hesitant ratings for cohesion due to ensemble size. It maintains the evaluative richness of traditional fuzzy or crisp MCDM methods that are neglected.

2.4 Methodological integration and research gap

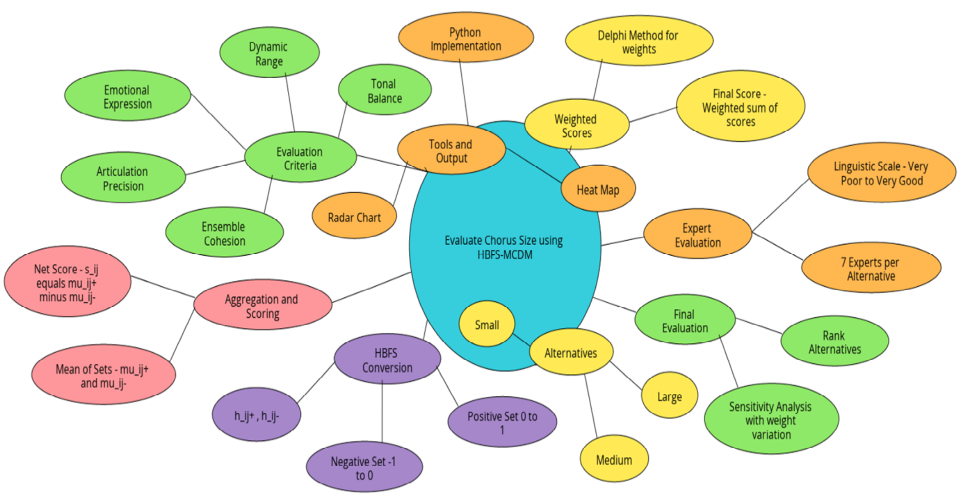
This paper presents a novel methodological contribution to music performance evaluation that integrates HBFS with a MCDM model. Considering complex and uncertain expert opinions enables aggregating such views into interpretable and robust performance rankings. Implicit in its application to evaluating five key performance criteria across small, medium, and large choirs, this study fills a significant research gap between musicology and decision science.

Although the previous models of fuzzy AHP and TOPSIS proved useful in structuring evaluation, they are not enough to handle dual and hesitant musical judgments by providing insufficient expressive depth. In addition to capturing these complexities, the proposed HBFS-MCDM framework provides a transparent and computationally sound way of aggregating expert opinions using the criterion weights obtained via the Delphi method. It allows comparisons that consider the complexity of human evaluation and provides valuable guidance to ensemble configuration in practical and educational environments.

3 Methodology

In this study, the methodological framework is developed to assess the effect of the chorus size on art music performance quality based on the hesitant bipolar fuzzy multi-criteria decision-making (HBFS-MCDM) model. Using this methodology in Figure 2, we can systematically aggregate hesitant expert opinions with dual (positive/negative) sentiments common in aesthetic performance evaluation. In this work, HBFS theory has been proposed to integrate into a weighted criteria decision-making structure to evaluate and rank three chorus size alternatives: small, medium, and large, on five expert-defined criteria for performance.

Figure 1 Methodological framework for evaluating chorus size in art music performance (see online version for colours)



3.1 Problem definition

Assume the set of alternatives is denoted by $A = A_1, A_2, A_3$, where A_1, A_2, and A_3, correspond to small, medium, and large choral ensembles, respectively. A set of five criteria evaluates each alternative.

$$C = \{c_1, c_2, c_3, c_4, c_5\} \tag{1}$$

where

- C_1: tonal balance
- C_2: articulation precision

- C_3: dynamic range
- C_4: ensemble cohesion
- C_5: emotional expression.

There are seven expert evaluators per alternative to assess each criterion based on the seven-point linguistic scale of very poor to very good. They are thereby converted to hesitant bipolar fuzzy elements with two degrees of satisfaction and dissatisfaction.

3.2 HBFS representation

In this framework, each evaluation of an alternative, A_i, under criterion, c_j, is represented by a hesitant bipolar fuzzy value:

$$h_{ij} = \left(h_{ij}^{+}, h_{ij}^{-}\right) \tag{2}$$

- $h_{ij}^{+} \subseteq [0, 1]$
- $h_{ij}^{-} \subseteq [-1, 0]$.

Such sets consist of several or one evaluator's ratings that indicate uncertainty at various levels. Consider that two experts give ratings for a given criterion as 'good' and 'very good'. The corresponding h_{ij}, might be {0.6, 0.8}, while h_{ij}, might be {–0.2, –0.1}, reflecting slight dissatisfaction in secondary dimensions.

3.3 Linguistic scale to HBFS mapping

The same positive and negative membership degrees are mapped to each linguistic term used in the evaluations. Below is an example of the mapping in Table 1.

Table 1 Linguistic terms and their corresponding HBFSs representing positive and negative membership degrees for performance evaluation

Linguistic term	h^{+} *(positive set)*	h^{-} *(negative set)*
Very poor	{0.0, 0.1}	{–0.9, –1.0}
Poor	{0.2, 0.3}	{–0.8, –0.7}
Fair	{0.4, 0.5}	{–0.6, –0.5}
Good	{0.6, 0.7}	{–0.4, –0.3}
Very good	{0.8, 0.9}	{–0.2, –0.1}

These sets are flexible and can be adjusted based on the expert panel's calibration and the evaluation context.

3.4 Aggregation of evaluations

The positive and negative components for each evaluation, h_{ij}, are aggregated by calculating the arithmetic mean of the respective sets:

$$\mu_{ij}^{+} = \frac{1}{\left|h_{ij}^{+}\right|}\sum\nolimits_{x\in h_{ij}^{+}} x,\ \mu_{ij}^{-} = \frac{1}{\left|h_{ij}^{-}\right|}\sum\nolimits_{y\in h_{ij}^{-}} |y| \tag{3}$$

The net performance score for each alternative under each criterion is then computed as follows:

$$s_{ij} = \mu_{ij}^{+} - \mu_{ij}^{-} \tag{4}$$

It also captures this overall evaluative position by balancing positive and negative satisfaction and subjective dissatisfaction in such a way as to punish alternatives that generate strong negative responses despite high positive assessments.

3.5 Criterion weighting and aggregated score computation

The relative importance of each criterion is captured through a set of weights $W = \{w_1, w_2, w_3, w_4, w_5\}$, determined using the Delphi method. Experts were asked to assign and adjust weights until consensus was achieved iteratively, ensuring that $\Sigma_{j=1}^{5}$ and each $w_j \in [0, 1]$. The final performance score for each alternative A_i is computed through a weighted summation:

$$S(A_i) = \sum\nolimits_{j=1}^{5} w_j \cdot s_{ij} \tag{5}$$

This score represents the overall desirability of the chorus size configuration based on expert evaluation across all criteria.

3.6 Ranking and sensitivity analysis

Finally, the three alternatives A_1, A_2, A_3 are ranked in descending order according to their final scores $S(A_i)$, in which the alternative with the highest score is considered optimal. To show the robustness of the decision model, each weight w_j is varied in a [−10%, 10%] range, and the ranked order is observed with the perturbation of each weight. This step guarantees that a slight change in evaluator preferences does not dramatically change the model's outcome.

3.7 Model implementation

All computations were scripted in Python to process linguistic input, Hbfs transformation, and net score computation, as well as perform ranking and sensitivity analysis. Visualising tools like radar charts and a performance heat map represented the evaluative outcomes.

4 Results and analysis

The results of the proposed HBFS-MCDM model are presented in this section and have been applied to decide the effect of chorus size on art music performance quality. The assessments are based on evaluations of seven expert evaluators across five criteria. The methodology described in Section 3 was used to compare three configurations of chorus size (small, medium, and large). The results are provided in four subsections: evaluation matrix of the HBFS, visual analysis of performance overall criteria, final ranking and aggregate scores, and sensitivity analysis.

4.1 HBFS evaluation matrix

First, the hesitant bipolar fuzzy evaluations obtained from the experts were transformed into positive and negative sets for each criterion and alternative. Net scores were then computed from the mean values of these sets. The satisfaction and dissatisfaction under each criterion are captured by $s_{ij} = \mu_{ij}^{+} - \mu_{ij}^{-}$. These values are summarised in Table 2 for all three alternatives.

Table 2 HBFS-derived scores for each alternative under each criterion

Criterion	*A_1: small choir*	*A_2: medium choir*	*A_3: large choir*
Tonal balance (c_1)	0.61 − 0.09 = 0.52	0.75 − 0.04 = 0.71	0.74 − 0.06 = 0.68
Articulation precision (c_2)	0.70 − 0.10 = 0.60	0.78 − 0.05 = 0.73	0.64 − 0.06 = 0.58
Dynamic range (c_3)	0.65 − 0.10 = 0.55	0.84 − 0.05 = 0.79	0.89 − 0.06 = 0.83
Ensemble cohesion (c_4)	0.74 − 0.10 = 0.64	0.86 − 0.06 = 0.80	0.75 − 0.08 = 0.67
Emotional expression (c_5)	0.68 − 0.09 = 0.59	0.84 − 0.07 = 0.77	0.80 − 0.08 = 0.72

Results show that medium-sized choirs (A_2) fared consistently well in all five criteria, and small choirs (A_1) scored the lowest in the net of all five criteria regarding tonal balance and dynamic range. The dynamic range of these large choirs (A_3) was best. Still, scores were slightly below in articulation and cohesion, presumably due to the experience of not being able to synchronise these large choirs.

4.2 Visual analysis: radar chart of criterion scores

To visually compare the performance of each chorus size across the five criteria, a radar plot in Figure 2 was constructed using normalised net scores.

The radar chart confirms that the medium choir gives the best-balanced performance profile. It has high evaluations in all criteria, indicating that it efficiently integrates sonic richness with control and ensemble clarity.

4.3 Final aggregate scores and ranking

The final performance scores for each alternative were computed using the weighted summation model:

$$S(A_i) = \sum_{j=1}^{5} w_j \cdot s_{ij} \tag{6}$$

Expert consensus using the Delphi method yielded equal weights for each criterion:

$$w_j = 0.20 \text{ for } j = 1, \ldots, 5 \tag{7}$$

Using the net scores s_{ij}, from Table 2, the aggregate scores for each alternative were calculated:

- $S(A_1) = 0.20(0.52 + 0.60 + 0.55 + 0.64 + 0.59) = 0.58$
- $S(A_2) = 0.20(0.71 + 0.73 + 0.79 + 0.80 + 0.77) = 0.76$
- $S(A_3) = 0.20(0.68 + 0.58 + 0.83 + 0.67 + 0.72) = 0.70$.

Figure 2 Radar plot of normalised net scores across all criteria (see online version for colours)

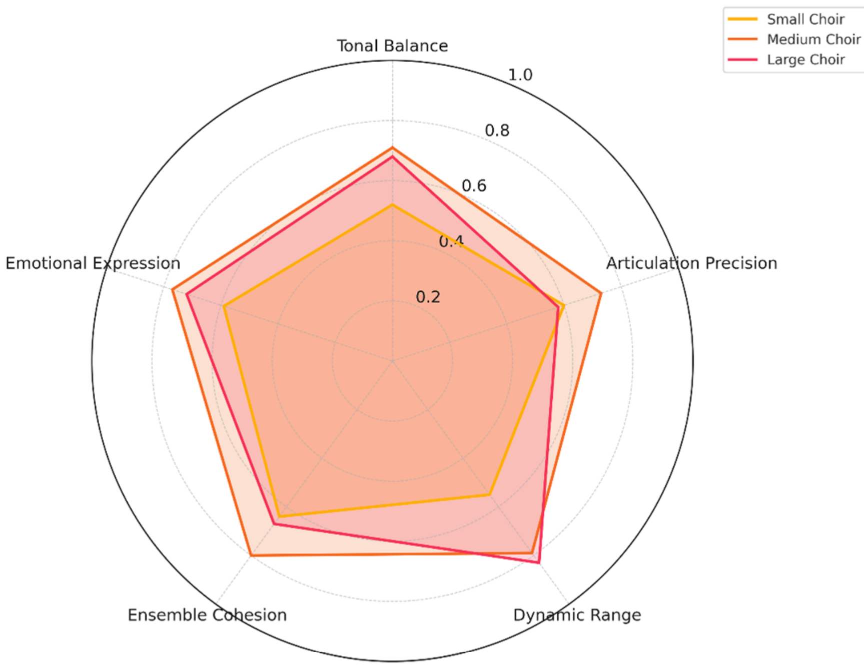


Table 3 Final aggregated scores and ranking

Alternative	*Final score*	*Rank*
Medium choir (A_2)	0.76	1
Large choir (A_3)	0.70	2
Small choir (A_1)	0.58	3

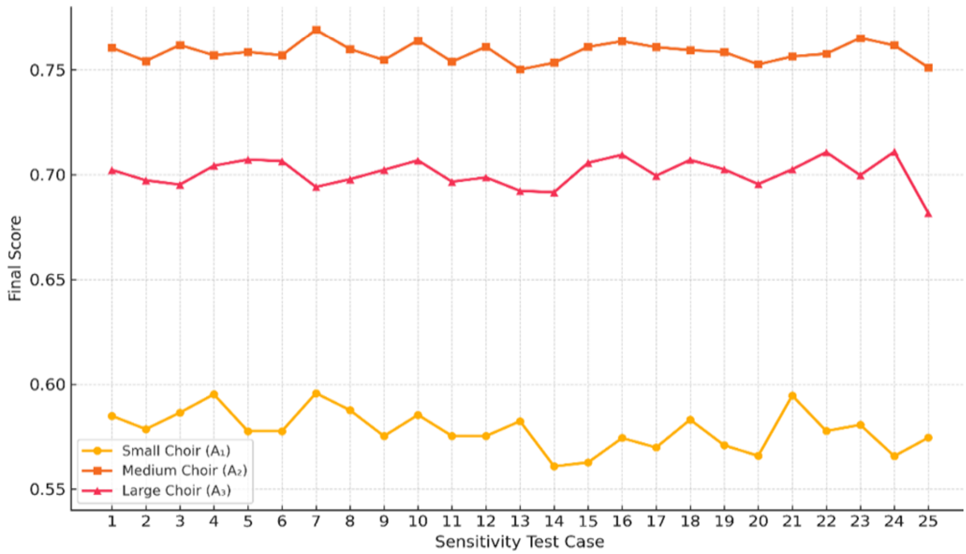
These results prove the hypothesis that there is an optimal point between multiple optimisation criteria in performance and that larger and smaller ensembles do not achieve this balance.

4.4 *Sensitivity analysis*

A sensitivity analysis was performed to assess the robustness of the final rankings. Each weight w_j, was perturbed by ±10% (within the constraint $\sum w_j = 1$, the resulting final scores and rankings changes were recorded across 25 test scenarios in Figure 3.

All scenarios, however, confirmed that the ranking order was stable and that medium choirs performed better than the alternatives in all cases. When the dynamic range weight increased, the difference in the scored value between A_2 and A_3 became slightly less; however, A_2 continued to be the best choice, which bore witness to the reliability and stability of the HBFS-MCDM model.

Figure 3 Line chart of alternative scores under weight variations (see online version for colours)



5 Discussion

This study applies the HBFS-MCDM model to provide novel insights into how chorus size influences performance quality in art music and offers methodological contributions in computational music evaluation. Section 4 presents results that validate the HBFS framework as a helpful way of describing subtle content-based subjective information and producing a systematic basis for optimising ensemble configurations for artistic and educational purposes.

It confirms practitioners' long-held but previously unsubstantiated beliefs that medium-sized choirs obtain the highest aggregate performance overall across five critical criteria. These ensemble members seem to have reached an ideal balance between clarity and richness. The tonal balance scores indicate that medium ensembles are adequately layered vocally to create a full harmonic texture without masking effects in larger ensembles. Their superior ratings in articulation precision and ensemble cohesion mirror this more manageable rehearsal coordination and better conductor control, which are more often a hardship in large choruses.

On the flip side, the large choir configuration is perfect in dynamic range because of sheer vocal mass, but weak articulation and cohesion. Such circumstances support the hypothesis that increased ensemble size leads to loss of micro-level control in passages that require rhythmic precision and strict vocal synchronisation. Although solid in cohesion and articulation, the performance of the small choirs failed to reach some minimum tonal depth and dynamic power, implying that they might be insufficiently forward in handling large-scale repertoire or emotionally extensive works.

It further reinforces the multifaceted impact chorus size has on artistic performance. The HBFS-MCDM model impressively reflects particular tradeoffs in each configuration through its bipolar and hesitant structure. Unlike other MCDM methods that distinguish between unipolar and deterministic judgments, the complexity of aesthetic evaluation is truly captured: a performance can inspire and disappoint a given evaluator simultaneously. For example, this way of balancing competing assessments is mathematically interpretable, precisely what the model demonstrates by integrating conflicting expert opinions, such as praising dynamic control while criticising cohesion.

Moreover, the results also point out the practical implications of using such a model for real-world choral programming. Logical and artistic constraints are regular burdens that creative directors and conductors bear in determining ensemble size. The model offers stakeholders a structured, multi-expert input that gives them an evidence-based way to weigh qualitative performance aspects. In particular, this has particular relevance in the academic, competitive, and professional choral contexts in terms of performance outcomes being determined, for instance, by the venue acoustics, choice of repertoire, and availability of singers.

From the methodological perspective, its success in this application indicates a broad application of the HBFS-MCDM model beyond choral music. Because its ability to manage ambiguity, hesitation, and contradiction is particularly well suited to evaluation problems involving other performing art forms (and subsequent judgments), its use as a framework for evaluating these problems (in dance, theatre, and instrumental performance) is a compelling development. Further, the model is flexible in treating linguistic inputs and is, therefore, suitable for evaluators without a technical background in digital performance assessment systems or conductor training systems.

However, the study also has its limitations. Equal weights, though methodologically neutral, do not necessarily reflect the evaluator's priorities or particular repertoire requirements. In future model versions, dynamic weighting schemes or more context-aware versions that change criterion importance depending on genre, period, or performance setting could be used. Additionally, the HBFS structure is robust but relies on the quality of the linguistic to numerical mappings. To further improve consistency and accuracy, closer granularity could be achieved in linguistic scale design and evaluator training.

Overall, the findings support the adequacy of large ensembles in attaining holistically satisfying performance quality, as well as the HBFS-MCDM framework's ability to describe evaluation abilities in complex environments with fidelity and realism. At the same time, this approach posits an essential contribution to research in computational aesthetics and performance analysis and provides actionable guidance for artistic decision-making in the choral domain.

6 Conclusions

Within evaluating the effects of chorus size on art music performance quality, this study proposed and implemented an HBFS MCDM framework. To address the dual sentiments (satisfaction and dissatisfaction) and the hesitancy expressed by an expert evaluator, the methodology was built on the concept of the two sentiments and an expert's hesitancy, which was incorporated into the HBFSs along with weighted decision aggregation. The study provided a systematic, data-informed approach to a question typically answered anecdotally or intuitively over five critical performance criteria through the structured assessment of small, medium, and large choral ensembles. The findings suggest that medium-sized choirs induce the highest and least variable balance and consistently high performance on tonal, technical, and expressive dimensions. As this suggests, large choirs are best at producing dynamic power, small choirs at cohesion, but medium ensembles seem best at combining clarity, resonance, broadcast, and expressive control. This work provides practical value for the ensemble directors and program planners, helping them pinpoint configurations that conform to the artistic goals and logistical constraints. The HBFS-MCDM model methodologically proved its ability to loophole the complexity and conflict of evaluations transparently and logically mathematically. It combined the richness of expert opinion while permitting aggregation and ranking amenable to performance assessment in artistic domains. Some future work might investigate adaptive weightings, real-time evaluation integration, or other performance settings. Contours of ensemble performance optimisation in clotting can be better understood in the context of fuzzy decision-making. This research contributes to the theoretical landscape of fuzzy decision-making and a practical understanding of ensemble performance optimisation in art music.

7 Future work

Future research can incorporate real-time acoustic measurements and audience feedback data to supplement the expert effort in extensions to the HBFS-MCDM model and create a hybrid model that includes subjective and objective performance indicators. Finally, the model can be applied to different musical genres, cultural contexts, and ensemble types (chamber groups and orchestras) to increase generalisability. Future responsiveness could be enhanced through adaptive or dynamic weighting mechanisms, either within these mechanisms, by utilising repertoire characteristics, or even conductor preferences. Other promising directions of integration with machine learning techniques in linguistic term calibration and fuzzy set tuning would lead to better accuracy and scalability of decisions.

Declarations

The authors declare that they have no conflict of interest.

References

Adongo, E. (2020) *Pedagogical Strategies for Developing Musician's Understanding of Styles and Conceptual Knowledge of Contemporary African Choral Music*, University of Florida.

Asaad, B.A., Musa, S.Y. and Ameen, Z.A. (2024) 'Fuzzy bipolar hypersoft sets: a novel approach for decision-making applications', *Mathematical and Computational Applications*, Vol. 29, No. 4, p.50.

Batovska, O., Byelik-Zolotaryova, N. and Ivanova, J. (2023) 'Modern global trends in the development of choral performance', *Studia Universitatis Babes-Bolyai, Musica*, Vol. 68, No. 7, pp.37–55.

Becker, J. (2004) *Deep Listeners: Music, Emotion, and Trancing*, Indiana University Press, USA.

Berenzweig, A., Logan, B., Ellis, D.P. and Whitman, B. (2004) 'A large-scale evaluation of acoustic and subjective music-similarity measures', *Computer Music Journal*, Vol. 5, No. 1, pp.63–76.

Chong, P.P. and Lalla, T.R. (2020) 'Applying fuzzy QFD MCDM to evaluate musical instruments', *Book of Abstracts*, Vol. 7, No. 3, p.106.

Clark, T.J. (2024) *Harmonizing Voices: Vocal Pedagogy in 21st Century Music Education*, Doctoral Dissertations and Projects, Liberty University, USA [online] https://digitalcommons.liberty.edu/doctoral/6320.

Fischer, M.J. (2024) *Selecting Appropriate Repertoire for the Intermediate Wind Band/Ensemble Towards an Understanding of Selecting Repertoire that is Engaging, Educational, and Inspiring*, Shenandoah University.

Gordon, E. (2002) *Rating Scales and their Uses for Measuring and Evaluating Achievement in Music Performance*, GIA Publications, Mason Ave, Chicago, USA.

Halvasi, B. (2018) 'Basic properties of chorus and fundamental approaches to improve', *Educational Policy Analysis and Strategic Research*, Vol. 13, No. 5, pp.113–126.

Haroutounian, J. (2000) 'Perspectives of musical talent: a study of identification criteria and procedures', *High Ability Studies*, Vol. 11, No. 6, pp.137–160.

Hightower, M.W. (2014) *A Comparative Study of Participant Satisfaction with Group Decision-Making between Consensus and Analytic Hierarchy Process Techniques*, Northcentral University.

Latimer Jr., M.E. (2009) 'Choral performance assessment: an overview of research to date', *The Phenomenon of Singing*, Vol. 7, No. 2, pp.81–91.

Li, Q., Li, Z., Han, J. and Ma, H. (2022) 'Quality assurance for performing arts education: a multi-dimensional analysis approach', *Applied Sciences*, Vol. 12, No. 8, p.4813.

Linaberry, R. (2021) 'On the sound: tone quality, blend, dynamics, ensemble balance, and articulations', *Strategies, Tips, and Activities for the Effective Band Director*, Routledge, Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK.

Lotfi, F.H., Allahviranloo, T., Pedrycz, W., Shahriari, M., Sharafi, H. and Razipour-Ghalehjough, S. (2023) *Fuzzy Decision Analysis: Multi Attribute Decision Making Approach*, Springer, Princeton, New Jersey, USA.

Mahfouf, M., Abbod, M.F. and Linkens, D.A. (2001) 'A survey of fuzzy logic monitoring and control utilisation in medicine', *Artificial Intelligence in Medicine*, Vol. 21, No. 4, pp.27–42.

Mahmood, T., Waqas, H.M. and Rehman, U.U. (2024) 'Selection of optimal encryption algorithm based on hesitant bipolar complex fuzzy frank aggregation operators', *Journal of Applied Mathematics and Computing*, Vol. 12, No. 6, pp.1–35.

Mandal, P. and Ranadive, A. (2019) 'Hesitant bipolar-valued fuzzy sets and bipolar-valued hesitant fuzzy sets and their applications in multi-attribute group decision making', *Granular Computing*, Vol. 4, No. 1, pp.559–583.

Pouyakian, M., Azimi, H.R., Patriarca, R., Keighobadi, E., Fardafshari, M. and Hanifi, S.M. (2024) 'Application of functional resonance analysis and fuzzy topsis to identify and prioritize factors affecting newly emerging risks', *Journal of Loss Prevention in the Process Industries*, Vol. 91, No. 14, p.105400.

Rodler, P. (2020) *RBF-HS: Recursive Best-First Hitting Set Search*, arXiv preprint arXiv:2010. 04282.

Stanley, M., Brooker, R. and Gilbert, R. (2002) 'Examiner perceptions of using criteria in music performance assessment', *Research Studies in Music Education*, Vol. 18, No. 5, pp.46–56.

Ternström, S. (2003) 'Choir acoustics: an overview of scientific research published to date', *International Journal of Research in Choral Singing*, Vol. 1, No. 5, pp.3–12.

Ternström, S. and Sundberg, J. (1988) 'Intonation precision of choir singers', *The Journal of the Acoustical Society of America*, Vol. 84, No. 7, pp.59–69.

Ternström, S., Jers, H. and Nix, J. (2018) 'Group and ensemble vocal music', *Vocal, Instrumental, and Ensemble Learning and Teaching: An Oxford Handbook of Music Education*, Vol. 3, No. 3, p.31.

Waqas, H.M., Emam, W., Mahmood, T., Rehman, U.U. and Yin, S. (2024) 'Selection of cloud security by employing MABAC technique in the environment of hesitant bipolar complex fuzzy information', *IEEE Access*, Vol. 12, No. 3, pp.123127–123148, DOI: 10.1109/access.2024.3436687.

Weiss, M.W. and Trehub, S.E. (2023) 'Detection of pitch errors in well-known songs', *Psychology of Music*, Vol. 51, No. 21, pp.172–187.

Yang, W. (2022) *Developing Musical Skills within Ensemble Environments: A Qualitative Case Study of Two Community Instrumental Ensembles in Christchurch*, University of Canterbury, New Zealand.

Yarnall, S. (2017) 'The 'real' me: practical application of research into the perception of vocal timbre', *Contemporary Music Review*, Vol. 36, No. 7, pp.562–579.

Zhang, G., Siripala, W., Fauziyah, N., Anwar, K. and Suwarno, S. (2024) 'Investigating the development chorus education: challenges and potential strategies in enhancing uniqueness', *International Journal of Learning, Teaching and Educational Research*, Vol. 23, No. 5, pp.287–307.