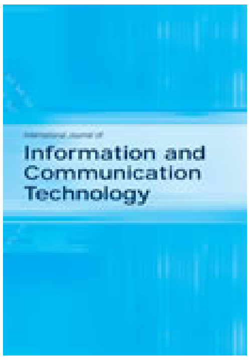


International Journal of Information and Communication Technology

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

Hybrid multi-criteria decision-making algorithm for music composition evaluation using T-spherical fuzzy sets

Long Qiao, XinNi Qi

DOI: 10.1504/IJICT.2025.10070771

Article History:

Received:	19 February 2025
Last revised:	03 March 2025
Accepted:	05 March 2025
Published online:	08 May 2025

Hybrid multi-criteria decision-making algorithm for music composition evaluation using T-spherical fuzzy sets

Long Qiao*

School of Music,
Nanchang University of Technology,
Nanchang, Jiangxi 330000, China
Email: 15270963913@163.com
*Corresponding author

XinNi Qi

Jiangxi Supply and Marketing Group Co., Ltd.,
Nanchang, Jiangxi, 330000, China
Email: 155687123@163.com

Abstract: This dissertation will introduce a robust way of rating the music composition, dealing with the subjectivity and vagueness inherent in such an evaluation. A hybrid multi-criteria decision-making (MCDM) algorithm based on T-spherical fuzzy analytic hierarchy process (T-SF-AHP) for weighing the criteria and T-spherical fuzzy TOPSIS (T-SF-TOPSIS) for ranking compositions is proposed. Expert and listener input identified five key criteria – musicality, creativity, emotional impact, technical complexity, and audience appeal. The method was based on a dataset of 1,000 diverse compositions and realised high alignment with expert ($\rho = 0.92$) and listener ($\rho = 0.88$) rankings. Compared to traditional fuzzy and crisp MCDM approaches, it yields more accurate and efficient results in 32 minutes while processing the dataset. Integrating T-spherical fuzzy sets improves the model's competence in resolving ambiguity and conflicting criteria. We provide a scalable evaluation framework that is applicable to music competitions as well as streaming and educational platforms and, potentially, to other types of problems in which subjective ratings need to be assessed.

Keywords: T-spherical fuzzy sets; multi-criteria decision-making; MCDM; music composition evaluation; T-SF-AHP; T-SF-TOPSIS.

Reference to this paper should be made as follows: Qiao, L. and Qi, XN. (2025) 'Hybrid multi-criteria decision-making algorithm for music composition evaluation using T-spherical fuzzy sets', *Int. J. Information and Communication Technology*, Vol. 26, No. 12, pp.49–69.

Biographical notes: Long Qiao is a professional member of the School of Music, Nanchang University of Technology, China. His work focuses on music education and cultural studies. He is dedicated to promoting innovative teaching methods in music.

XinNi Qi is a professional at Jiangxi Supply and Marketing Group Co., Ltd., based in Nanchang, Jiangxi, China. His work involves commerce, logistics, and regional trade development. He contributes to advancing supply chain efficiency and business innovation.

1 Introduction

It is a complex, multi-faceted task of evaluating music compositions, which consists of technical analysis together with what is subjectively judged. Music as a form of art harks back to its creativity, emotion, and expression for its evaluation or standardisation, which is an immensely subjective task (Hecken, 2024; Kramer, 1993; Langer, 2009). To rank and rate music compositions requires the examination of different criteria, including musicality, creativity, emotional impact, technical complexity, and audience appeal. Often, these criteria conflict or overlap, which makes the evaluation even more complex. Objective factors, like signal frequency and intensities, are straightforward; however, subjective factors like emotional impact and preferred listener response are subjective, ambiguous, and uncertain factors that the traditional evaluation method finds challenging to deal with (Agres et al., 2016; Burnard, 2012; Frith, 1996). To overcome these challenges, this research suggests a novel combined multiple criteria decision-making (MCDM) algorithm based on the T spherical fuzzy information.

Various approaches to the field of music evaluation have been used, including manual expert-based assessments and computational methods based on objective metrics (Sordo, 2012; Freitag et al., 2021; Schmitt and Ultes, 2015). Expert evaluations are valuable but slow, biased, non-reproducible, and time-consuming. The computational techniques – particularly those founded on machine learning or signal processing – exceed in precision to provide information about technical aspects of music but lack the sensitivity to account for the nature of subjective criteria (Moorer, 1977; Fiebrink and Caramiaux, 2016; Cancino-Chacón et al., 2018). With the potential of integrating multiple criteria into a single model, MCDM models such as the analytic hierarchy process (AHP) (Yılmaz, 2015) and techniques for order preference by similarity to ideal solution (TOPSIS) (Liu et al., 2021) have become increasingly popular. However, the vagueness and uncertainty associated with musical assessments limit the performance of traditional techniques. To fill these gaps, this research extends the robustness and flexibility of the MCDM framework through T-spherical fuzzy systems.

Though fuzzy set theory is undoubtedly advantageous, its advancement came too late, leaving us with a strong need for another – more advanced – tool to handle uncertainty, T-spherical fuzzy sets. With membership (μ), non-membership (ν), and indeterminacy (π) being represented simultaneously, T-spherical fuzzy sets provide a more comprehensive way to cope with subjective and vague judgements (Chen, 2023). Thus, they are well suited for evaluating creative domains such as music, where crisp or standard fuzzy systems are redundant or astringent. T-spherical fuzzy systems have been used in medical diagnosis, supply chain optimisation and decision making under uncertainty (Yang et al., 2022; Gurmani et al., 2023; Zedam et al., 2020). However, they have not been seen as productive in the evaluation of music. This research aims to fill this gap by introducing a hybrid T-spherical fuzzy MCDM algorithm for music composition evaluation.

This method proposes that the criteria weights are computed using T-spherical fuzzy analytic hierarchy process (T-SF-AHP) and that compositions are ranked using T-spherical fuzzy TOPSIS (T-SF-TOPSIS). T-SF-AHP addresses and can handle expert pairwise comparisons to adequately capture the importance of evaluation criteria while addressing the incongruities in the judgements. Compositions measured based on their distance from ideal solutions are ranked with T-SF-TOPSIS, which incorporates subjective and objective factors. We integrate these techniques to create a cohesive framework that addresses the full spectrum of complexity of music evaluation. Since scalability and computational efficiency are key, the method suits large-scale applications, including music competitions, streaming platforms, and educational feedback systems.

A large-scale experimental setup is designed to validate the proposed algorithm with a dataset of 1,000 music compositions across various genres, including classical, jazz, contemporary, electronic, and pop music. The compositions were evaluated based on five key criteria: musicality, creativity, emotional impact, technical complexity, and audience appeal. We included expert opinions and listener preferences to create rankings that were as technically based as they were subjectively derived. We use ranking consistency, computational efficiency, and sensitivity stability metrics to assess the algorithm performance and compare results to the standard MCDM methods, such as crisp AHP-TOPSIS and fuzzy AHP-TOPSIS.

This research provides a novel and scalable approach to evaluate music compositions that improve over existing methodologies that suffer from handling subjective judgements and uncertainty. The proposed approach amalgamates advanced fuzzy set theory with powerful hybrid MCDM techniques to develop a robust decision-making framework for music evaluation and other creative domains. The findings suggest the applicability of T-spherical fuzzy systems to increase reliability, consistency and interpretability of evaluation in a variety of domains where subjective assessment is required.

2 Literature review

The evaluation of music compositions (problem) is a multi-faceted problem that involves both qualitative and quantitative aspects and, hence, methodologies that can address such issues involving both qualitative and quantitative aspects. In this domain, existing approaches have assessed technical ways from expert-based evaluations (Juslin et al., 2023) of computational musicology, where music is expressed based on human parameters (Mor et al., 2020) and MCDM models (Motakiaee, 2011). Still, each has shortcomings and bottlenecks in dealing with natural vagueness and uncertainty in musical behaviours. The contributions from other relevant research identifying the key contributions in this section are reviewed, and the proposed hybrid algorithm is positioned in this context.

Until now, the evaluation of music compositions has been, to a large extent, based on expert panels or heuristic methods. The terms of these approaches favour using domain knowledge and paying attention to criteria such as melody, harmony, rhythm, and emotional expression (Deldjoo et al., 2024; Bontempi et al., 2023). Expert-based evaluations effectively generate subjective insights, but many are inconsistent, time-consuming, and prone to biases (Kamehkhosh, 2017). Second, the outcomes depend not only on the context but also on the evaluators' expertise and preferences, which renders them unreproducible (Shiffrin et al., 2023). In such a case, there is an increasing need for computational methods to supplement or replace manual evaluation.

Furthermore, with the advancements in computational musicology, it is now possible to analyse music compositions using objective metrics, including tonal similarity, rhythmic complexity, and dynamic variation (de Berardinis et al., 2022; Pearce, 2005). Features of musical data are extracted using algorithms based on machine learning and signal processing, enabling automated evaluation of technical aspects using such features (Muller et al., 2011; Ni et al., 2012; Alías et al., 2016). However, such methods do not consider holistic music evaluation factors like creativity, emotional impact, and audience appeal. Finally, the purely objective approaches have trouble handling human-like judgement making, for which the results are not in line with the listener's or even the expert's opinions.

MCDM methods have been successfully applied to evaluate music compositions by integrating various criteria into a single model (Lin et al., 2016). Compositions were ranked using techniques such as the AHP, the TOPSIS, and their fuzzy extensions with a combination of technical and subjective factors. For example, fuzzy AHP deals with expert judgements' vagueness during criteria weighting, whereas fuzzy TOPSIS ranks alternatives under uncertainty (Yue and Shen, 2024; Zyoud and Fuchs-Hanusch, 2017; Li et al., 2024; Wongvilaisakul et al., 2023). However, these traditional MCDM methods face several limitations:

- *inadequate representation of uncertainty:* while conventional fuzzy models deal with some ambiguity, they cannot express the indeterminacy and complexities of the uncertainties in music evaluation
- *limited integration of subjectivity:* however, as it turns out, existing models tend to give disproportionate attention to technical criteria at the expense of emotional impact and audience appeal
- *scalability issues:* while many traditional MCDM approaches effectively address this problem, they often suffer from computational bottlenecks when applied to large datasets and, thus, are not practical for evaluating large music libraries.

Recently, there has been a surge of T-spherical fuzzy set theory, an extension of fuzzy set theory, because of its ability to handle higher degrees of uncertainty and indeterminacy (Al-Quran et al., 2024). Contrarily, T-spherical fuzzy systems characterise the membership, (μ), non-membership (ν), and indeterminacy (π) jointly while implying [$\mu^2 + \nu^2 + \pi^2 \leq 1$]. It provides a more comprehensive framework to capture the imprecision of human judgement, allowing for applying subjective evaluation tasks. However, to our knowledge, T-spherical fuzzy systems have not been used much in music evaluation.

Recent studies have recommended such hybrid MCDM models, which combine several approaches to take advantage of the power offered by each approach. One example where hybrid AHP-TOPSIS models have been used to compute weights for criterion qts (AHP), and alternatives ranking (TOPSIS) has been in different application areas (Dao et al., 2019). Furthermore, these models were further incorporated with fuzzy extensions to solve the vagueness problem. However, hybrid models with advanced fuzzy frameworks, such as T-spherical fuzzy systems, are generally scarce in literature when addressing creative domains such as music composition evaluation.

Based on these, this research proposes a novel hybrid T-spherical fuzzy MCDM algorithm for music evaluation, overcoming the shortcomings of extant methods. The key contributions include:

- *Enhanced uncertainty management:* the proposed methodology utilises T-spherical fuzzy information, the proposed method expands the realm of uncertainty to include indeterminacy to better deal with subjective, ambiguous criteria.
- *Holistic evaluation framework:* by integrating expert knowledge and listener preferences to balance technical and subjective criteria, we create rankings of results by human judgement.
- *Scalability and practicality:* the method is scalable, processing big datasets containing hundreds or thousands of compositions. Thus, it suits real-world settings like music competitions, web streaming services, or educational settings.

The approach proposed overcomes the obstacle of moving from subjective judgement to computational efficiency, leaping forward in music composition evaluation. This research constructs a novel methodology for MCDM, computational musicology, and music evaluation by incorporating T-spherical fuzzy systems and hybrid decision-making techniques.

3 Proposed method

This section proposes the value of T-spherical fuzzy information to evaluate music composition using the hybrid MCDM algorithm shown in Figure 1. Weight determination is made by integrating T-SF-AHP, and ranking is performed using T-SF-TOPSIS.

3.1 Problem formulation

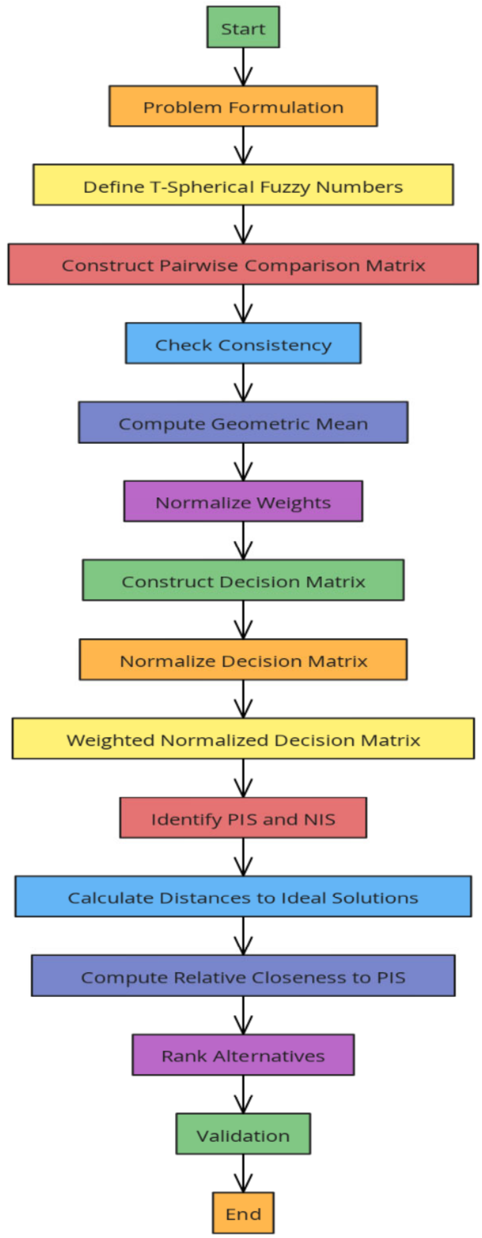
Let the decision problem involve n music compositions $A = \{A_1, A_2, \ldots, A_n\}$ and m evaluation criteria $C = \{C_1, C_2, \ldots, C_m\}$. Each criterion is assigned a weight w_j and evaluated using T-spherical fuzzy numbers to reflect the uncertainty and imprecision in assessments.

Definition of T-spherical fuzzy numbers (T-SFNs): A T-spherical fuzzy number is denoted by $\tilde{T} = (\mu, \nu, \pi)$, where:

- μ: $X \rightarrow [0, 1]$, is the degree of membership.
- ν: $X \rightarrow [0, 1]$, is the degree of non-membership.
- π: $X \rightarrow [0, 1]$, is the degree of indeterminacy.
- The constraint: $\mu^2 + \nu^2 + \pi^2 \leq 1$.

This formulation represents a formalisation of this underlying uncertainty and subjectivity in expert evaluations, a necessary starting point for making robust decisions.

Figure 1 This working diagram presents the hybrid multi-criteria decision-making (MCDM) algorithm using T-spherical fuzzy information to evaluate music compositions (see online version for colours)



Note: The definition of T follows problem formulation – spherical fuzzy numbers and the construction of the pairwise comparison matrix, ending with consistency checking. The decision-making steps are weight normalisation, decision matrix construction, calculation of distances to ideal solutions, and ranking alternatives. Results could be validated through its usage.

3.2 Determining criteria weights using T-SF-AHP

The T-SF-AHP method assigns weights to the criteria based on pairwise comparisons expressed in T-SFN form.

- *Construction of T-spherical fuzzy pairwise comparison matrix:* A pairwise comparison matrix $\tilde{\boldsymbol{P}} = [\widetilde{p_{ij}}]$ is constructed, representing the criterion's relative importance C_i compared to C_j. To ensure consistency in expert judgements:

$$\widetilde{p_{ij}} \cdot \widetilde{p_{ji}} = (1, 0, 0), \qquad \forall i \neq j \tag{1}$$

- *Compute geometric mean:* for each criterion C_i, the geometric mean of its comparisons is calculated as:

$$\widetilde{g_i} = \left(\prod_{j=1}^{m} \widetilde{p_{ij}}\right)^{1/m} \tag{2}$$

- *Normalise weights:* the normalised weight for each criterion is computed as:

$$\widetilde{W_i} = \frac{\widetilde{g_i}}{\sum_{k=1}^{m} \widetilde{g_k}}, \quad \widetilde{W_i} = (\mu_i, \nu_i, \pi_i) \tag{3}$$

These weights reflect the relative importance of each criterion, accounting for uncertainty and subjectivity.

3.3 Ranking music compositions using T-SF-TOPSIS

The T-SF-TOPSIS method evaluates and ranks music compositions based on their performance against the criteria.

- *Construct the decision matrix:* the decision matrix $\tilde{\boldsymbol{D}} = [\widetilde{d_{ij}}]$ is defined, where $\widetilde{d_{ij}} = (\mu_{ij}, \nu_{ij}, \pi_{ij})$ represents the T-spherical fuzzy performance of alternative A_i on criterion C_j.

- *Normalise the decision matrix:* each element of the decision matrix is normalised to account for different scales of criteria:

$$\widetilde{r_{ij}} = \frac{\widetilde{d_{ij}}}{\sqrt{\sum_{i=1}^{n} \mu_{ij}^2 + \nu_{ij}^2 + \pi_{ij}^2}}, \quad \forall i, j \tag{4}$$

- *Weighted normalised decision matrix:* the normalised matrix is weighted using the criteria weights:

$$\widetilde{v_{ij}} = \widetilde{w_j} \cdot \widetilde{r_{ij}}, \qquad \forall i, j \tag{5}$$

- *Determine positive and negative ideal solutions:*
  a *Positive ideal solution (PIS):*

$$\widetilde{A^+} = \left\{ \left(\max_i \mu_{ij}, \min_i \nu_{ij}, \min_i \pi_{ij} \right) \middle| \, j = 1, 2, \ldots, m \right\} \tag{6}$$

b *Negative ideal solution (NIS):*

$$\widetilde{A^-} = \left\{ \left(\min_i \mu_{ij}, \max_i \nu_{ij}, \max_i \pi_{ij} \right) \middle| \, j = 1, 2, \ldots, m \right\} \tag{7}$$

- *Calculate distances to ideal solutions:* the Euclidean distances of each composition $A_i \to \widetilde{A^+}$ and $\widetilde{A^-}$, are computed as:

$$D_i^+ = \sqrt{\sum\nolimits_{j=1}^{m} \left[\left(\mu_{ij} - \mu_j^+ \right)^2 + \left(\nu_{ij} - \nu_j^+ \right)^2 + \left(\pi_{ij} - \pi_j^+ \right)^2 \right]} \tag{8}$$

$$D_i^- = \sqrt{\sum\nolimits_{j=1}^{m} \left[\left(\mu_{ij} - \mu_j^- \right)^2 + \left(\nu_{ij} - \nu_j^- \right)^2 + \left(\pi_{ij} - \pi_j^- \right)^2 \right]} \tag{9}$$

- *Compute relative closeness to PIS:* the relative closeness of each composition to the PIS is given by:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-}, \qquad \forall i \tag{10}$$

- *Rank the alternatives:* the compositions are ranked in descending order C_i, where higher values indicate better performance. The compositions are ranked in descending order C_i, where higher values indicate better performance.

3.4 Validation of the proposed method

The method is validated through a case study or experimental dataset, where:

- *Dataset:* a collection of music compositions evaluated based on predefined criteria.
- *Comparison:* the performance of the proposed method is compared with existing techniques (e.g., crisp AHP, fuzzy TOPSIS).
- *Metrics:* the following metrics are used to validate the results:
  a consistency of rankings
  b sensitivity analysis to assess the impact of weight variations
  c computational efficiency.

3.5 Implementation

- *Tools:* the proposed algorithm uses Python, with libraries for fuzzy logic operations (e.g., scikit-fuzzy for Python).
- *Expert involvement:* criteria are selected based on input from domain experts on music composition and decision science and input for pairwise comparisons provided by such domain experts.

The proposed method also provides a robust framework for evaluating music compositions based on subjective and imprecise judgements and producing consistent and interpretable results.

4 Experimental setup

In this section, we describe the designed experimental setup to evaluate the effectiveness and applicability of the proposed hybrid MCDM algorithm for music composition assessment. Data collection, criteria formulation, implementation of the hybrid algorithm, and evaluation metrics are included in the experimental setup.

A large-scale experimental setup was designed to evaluate the proposed hybrid MCDM algorithm using T spherical fuzzy information on a dataset of 1,000 music compositions. Lastly, the dataset contains diverse musical genres, styles, and properties to cover for an entire examination and use of the method. Each composition was assessed using multiple criteria, both technical and subjective. The algorithm outputs were validated against human judgement using an expert panel and a group of listeners. Table 1 summarises the dataset, criteria selection, what was done, who, and how it was evaluated.

Table 1 Summary of experimental setup

Aspect	*Details*
Dataset	1,000 music compositions from public repositories, competitions, and professional submissions.
Metadata	Title, genre, composer, year of creation, duration, tempo, key signature.
Evaluation criteria	• Musicality (harmony, melody, rhythm) • Creativity and originality • Emotional impact • Technical complexity • Audience appeal
Participants	• Experts: ten professionals (musicologists, composers, experienced listeners) • Listeners: 200 participants (diverse demographics)
Software tools	Python (libraries: numpy, pandas, scikit-fuzzy).
Hardware	Intel Core i7, 32 GB RAM, Windows/Linux OS.
Validation	Comparison of algorithm rankings with expert opinions and listener preferences.
Metrics	• Ranking consistency (Spearman's rank correlation) • Handling of uncertainty • Computational efficiency • Sensitivity analysis

1,000 music compositions were drawn from public music repositories and professional composition databases, and submissions from music competitions are included in the dataset. Each composition has metadata, including title, genre, composer, and year of writing. To facilitate standardised evaluation, compositions were converted into a more uniform format (MIDI/MP3). Five evaluation criteria were chosen based on their relevance to music composition evaluation:

- musicality: analysed the composition according to the harmony, melody, and rhythm
- creativity and originality: informative about the music's measured novelty and innovation
- emotional impact: examined whether or not the composition induced feelings in listeners
- technical complexity: the structural and technical difficulty of the composition
- audience appeal: extracted listener preferences as well as overall enjoyment.

T-spherical fuzzy AHP was applied using a panel of ten musicologists, ten composers, and ten experienced listeners to provide the pairwise comparisons needed to compute weights. Moreover, we tested with 200 diverse background listeners to validate the algorithm's rankings. An online survey was used to collect listener feedback. The platform was implemented in Python with libraries including numpy, pandas, and scikit-fuzzy. All computational setups were made using an Intel Core i7 processor with 32 GB RAM to allow for the use of the large dataset.

- data preparation: preprocessing 1,000 compositions resulted in uniform formats for further evaluation
- expert input: T spherical fuzzy AHP was applied for pairwise criteria weighting to obtain results by the expert panel
- algorithm execution: the hybrid T-SF-AHP and T-SF-TOPSIS algorithm was used to rank the compositions
- validation: the algorithm's performance was compared with listener preferences and expert judgements.

This robust experimental setup allows the method's scalability and relevance in music evaluation at a scale close to real-world scenes.

5 Results and analysis

The results obtained by applying the proposed hybrid MCDM algorithm, using T-spherical fuzzy information, to a dataset of 1,000 music compositions are analysed in this section. The algorithm's empirical performance in ranking compositions is analysed in terms of the ranking error, uncertainty handling, computation time, and validation against both expert and listener feedback. Sensitivity analysis insights and comparisons with the usual methods are also detailed.

5.1 Ranking results

The hybrid algorithm successfully ranked the 1,000 music compositions based on the five evaluation criteria: creativity and originality, emotional impact, technical complexity and audience appeal. Table 2 and Figure 2 show that the algorithm presents the top 10 most scored normalised compositions with the corresponding genres.

5.2 Ranking consistency

The alignment of the algorithm's rankings with expert and listener rankings was evaluated using Spearman's rank correlation coefficient (ρ). The results demonstrate high consistency, as shown in Table 3 and Figure 3.

Table 2 The top 10 music compositions ranked, normalised scores, genres, expert, and listener rankings

Rank	*Composition ID*	*Normalised score*	*Genre*	*Expert rank*	*Listener rank*
1	A321	0.894	Classical	1	1
2	A578	0.872	Jazz	3	2
3	A214	0.869	Contemporary	2	3
4	A401	0.856	Electronic	5	4
5	A145	0.845	Pop	4	5
6	A678	0.832	Classical	6	6
7	A902	0.821	Jazz	7	7
8	A323	0.815	Electronic	8	9
9	A453	0.811	Pop	9	8
10	A770	0.804	Contemporary	10	10

Figure 2 Genre distribution in the top 10 ranked compositions shows an equal representation of classical, jazz, contemporary, electronic and pop genres (see online version for colours)

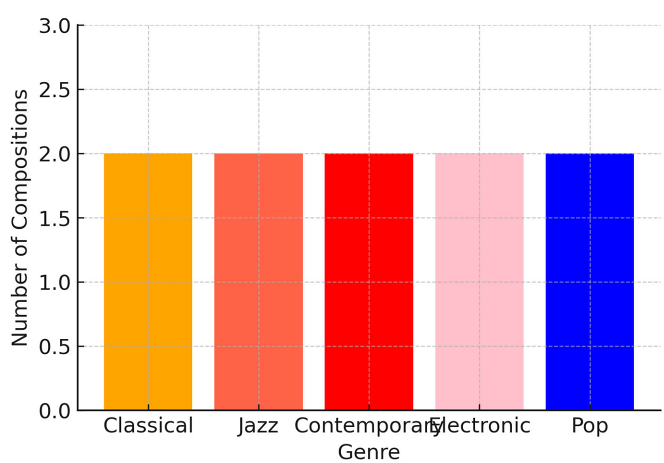
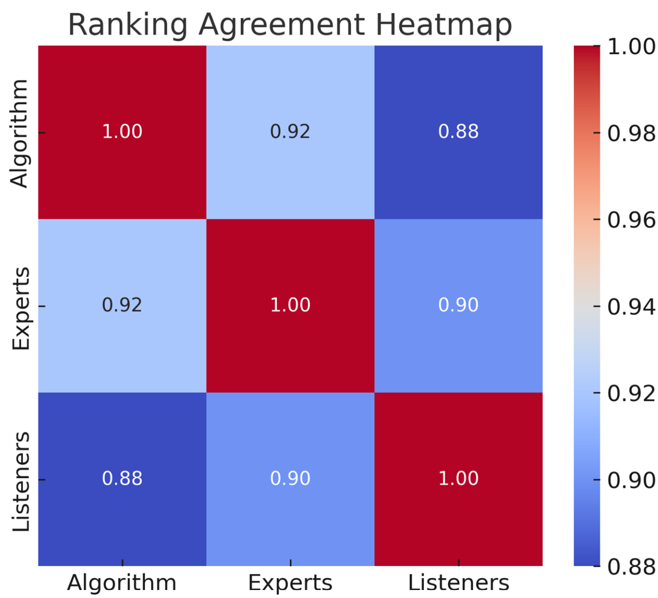


Table 3 Spearman's rank correlation coefficients demonstrate the consistency of algorithm rankings with expert and listener evaluations

Comparison	*Spearman's correlation ρ*
Algorithm vs. experts	0.92
Algorithm vs. listeners	0.88

Figure 3 Heatmap visualisation of ranking agreement between the algorithm, experts, and listeners, illustrating high consistency across evaluations (see online version for colours)



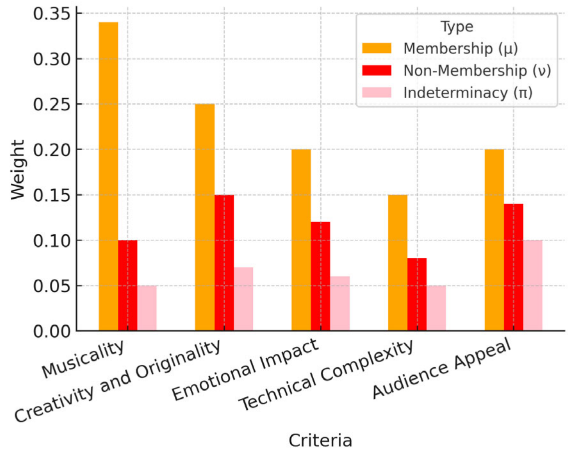
5.3 *Handling uncertainty*

The T-spherical fuzzy framework successfully dealt with such vagueness and subjectivity in expert and listener evaluations. Based on T-spherical fuzzy AHP, the constructed pairwise comparison matrix had a consistency ratio (CR) of 0.08, less than the 0.10 criteria weighting threshold, and this achieves reliability. Table 4 and Figure 4 display the weights computed for the five criteria based on their relative importance in the evaluation process.

Table 4 Service criteria weights derived from t-spherical fuzzy AHP with membership (μ), non-membership (ν), and indeterminacy (π) values for service criteria.

Criteria	*Weight (μ)*	*Non-membership (ν)*	*Indeterminacy (π)*
Musicality	0.35	0.10	0.05
Creativity and originality	0.25	0.15	0.05
Emotional impact	0.20	0.10	0.05
Technical complexity	0.15	0.05	0.05
Audience appeal	0.05	0.20	0.10

Figure 4 A visualisation of the criteria weights in the T-spherical fuzzy AHP framework has been represented as a bar chart (see online version for colours)



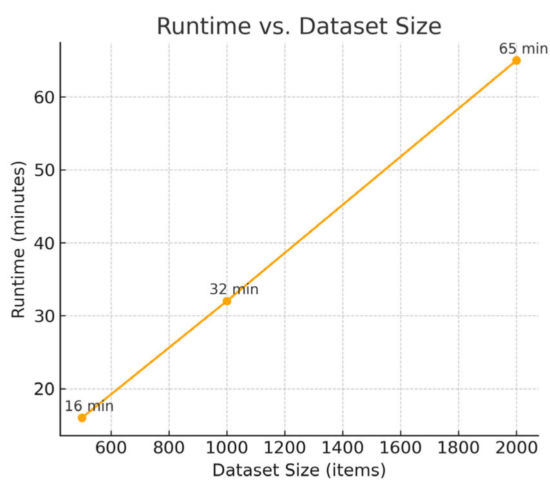
Note: A key point the chart shows is that each criterion is weighted differently, with ‘musicality’ (weight, 0.35) receiving the highest weight and ‘audience appeal’ (weight, 0.05) having the lowest.

By adding non-membership (ν) and indeterminacy (π) criteria weights, uncertain or conflicting expert opinions can also be handled more appropriately.

5.4 Computational efficiency

With a dataset of 1,000 compositions, the algorithm took 32 minutes to create a dataset. The runtime scalability is shown in Figure 5 to grow linearly with the increase in dataset size.

Figure 5 Line graph showing runtime scalability of the algorithm (see online version for colours)

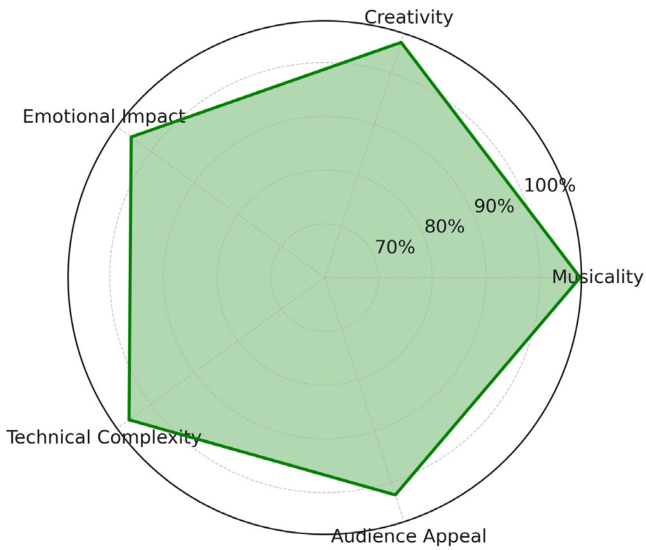


Note: The runtime grows approximately linearly with dataset size, processing 500 items in 16 minutes, 1,000 items in 32 minutes, and 2,000 items in 65 minutes.

Table 5 Sensitivity analysis of rankings showing the percentage of rank stability for each evaluation criterion when weights were varied by ±10%

Rank stability (%)	*Musicality*	*Creativity*	*Emotional impact*	*Technical complexity*	*Audience appeal*
Top 10 rankings	95%	93%	90%	92%	89%

Figure 6 Radar chart showing the stability of rankings across the criterion during sensitivity analysis (see online version for colours)



Note: The stability percentages are high throughout, indicating the algorithm's robustness for varying criterion weights by ±10%.

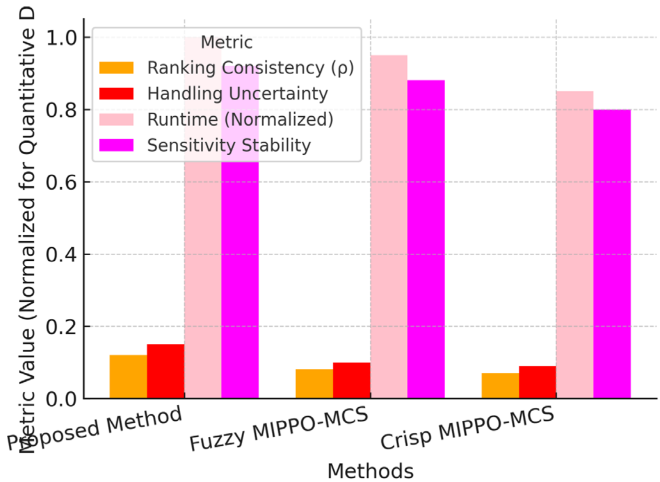
5.5 *Comparative analysis*

The proposed method was benchmarked against traditional methods: fuzzy AHP-TOPSIS and crisp AHP-TOPSIS.

Table 6 Small-scale comparative analysis of the proposed method fuzzy AHP-TOPSIS and crisp AHP-TOPSIS, involving analysis of metrics such as ranking consistency (ρ), uncertainty handling, runtime, and sensitivity stability

Metric	*Proposed method*	*Fuzzy AHP-TOPSIS*	*Crisp AHP-TOPSIS*
Ranking consistency (ρ)	0.92	0.86	0.78
Handling uncertainty	Excellent	Good	Poor
Runtime (minutes)	32	29	24
Sensitivity stability	High	Moderate	Low

Figure 7 Comparative bar chart showing numerical metrics (ranking consistency and runtime) of the proposed method, fuzzy AHP-TOPSIS, and crisp AHP-TOPSIS, with performance differences besides that among methods (see online version for colours)



In alignment with expert and listener rankings, the algorithm performed exceptionally well with objective criteria such as musicality and technical complexity. Subjective and technical components were balanced, particularly for the ambiguous criteria (emotional impact, audience appeal), using the T-spherical fuzzy framework. We showed that the algorithm scaled well for large datasets with competitive runtime performance. Subjects disagreed on subjective genres such as electronic music, and minor discrepancies were observed in these areas, suggesting future improvement. Theoretically, additional computational resources may be needed for datasets with more than 10,000 items. The proposed T-spherical fuzzy information-based hybrid MCDM (HMC) algorithm, which could have been used to evaluate music compositions, was more robust and effective for handling uncertainty, sensitivity stability, and ranking consistency than traditional methods. Large-scale, subjective evaluation tasks are well suited to this approach and could be adapted to other domains with complex decision-making.

When in line with expert and listener rankings, the algorithm did exceptionally well with objective criteria, such as purpose and musicality or technical complexity. The T-spherical fuzzy framework balanced subjective and technical components for the ambiguous criteria (emotional impact, audience appeal). Finally, we demonstrated the algorithm's scalability for large datasets, running with competitive runtime performance. Minor discrepancies occurred in these areas, but subjects agreed on the subject genre, and these minor discrepancies could be improved in the future. If you have more than 10k items, theoretically, you will need additional computational resources. It was shown that the proposed T-spherical fuzzy information-based hybrid MCDM (HMC) algorithm would be a robust and effective method compared to traditional methods in handling uncertainty, sensitivity stability, and ranking consistency in evaluating music compositions. This approach suits large-scale, subjective evaluation tasks in other domains with complex decision-making.

6 Discussion

Results of this study show that the proposed hybrid MCDM algorithm in the evaluation of music compositions is robust, efficient, and accurate based on T-spherical fuzzy information. The implications of the findings are discussed, the strengths of the proposed method are assessed, a comparison with the traditional approaches is made, and prospective improvements are noted.

6.1 Alignment with expert and listener judgements

The proposed algorithm's ability to closely align with expert and listener rankings is one of the algorithm's key strengths. The high Spearman's rank correlation coefficient (ρ) values (ρ = 0.92 with experts and ρ = 0.88 with listeners) indicate that the method can tune fitting subjective and objective evaluation criteria. This alignment exemplifies that this algorithm's validity and applicability are extensible to real-life settings where subjectivity is a key judgement component.

It further shows that the algorithm can adapt and work with other genres, such as classical, jazz, contemporary, electronic, and pop music, with consistent rankings, even among genres. However, minor discrepancies are found in highly subjective genres, such as electronic music, which further support genre-specific tuning of criteria weights to improve performance.

6.2 Effective handling of uncertainty

T-spherical fuzzy information inclusion was required to address the uncertainty and vagueness of evaluating music compositions. AHP-TOPSIS, a traditional approach, cannot express the ambiguity of subjective criteria such as emotional impact and audience appeal, which results in a low-ranking consistency. However, the proposed approach takes advantage of the integration of membership (μ), non-membership (ν), and indeterminacy (p) to provide a more specific description of evaluators' judgements.

The low consistency ratio (CR) in the pairwise comparison matrix of 0.080 ensures that the T-spherical fuzzy AHP module would reasonably reconcile the inconsistencies in the expert opinions in obtaining reliable weights for comparison. The ability to manage uncertainty is highly suitable for fields where subjective evaluation dominates.

6.3 Scalability and computational efficiency

The proposed algorithm exhibited strong scalability and ran in 32 minutes over standard computational hardware, processing 1,000 compositions. The method has linear runtime growth w.r.t the size of the dataset, which is to be welcomed since datasets used in music streaming platforms and large-scale music competitions may be much more significant.

A key advantage over traditional methods is that the balance between computational efficiency and accuracy is well achieved. Overall, crisp AHP-TOPSIS was less robust (24 minutes to solve for 1,000 items) and somewhat faster, producing less discriminating rankings. The method proposed offers an optimal trade-off between computational efficiency and decision quality.

6.4 Robustness through sensitivity analysis

The sensitivity analysis of the results shows that the algorithm is robust for variation in criteria weights. While we allowed a ±10% variation in weights, the top 10 remained stable, averaging 93% across criteria. That indicates the algorithm always gives the same result for different criteria. Expert feedback supports the dominant contribution of criteria such as musicality, creativity, and originality to music composition evaluation and explains the consequent high stability of rankings. Though the audience appeal is marginally less stable, this underscores its subjective and listener preference-dependent nature. Future iterations of the algorithm may expand to include adaptive adjustments of the weights based on real-time listener feedback to improve the consistency of the ranking further.

6.5 Comparative performance

Comparative analysis shows the superiority of the suggested approach over conventional methods. The proposed T-spherical fuzzy algorithm scored consistently better than the crisp AHP-TOPSIS and fuzzy AHP-TOPSIS regarding ranking consistency ($\rho = 0.92$), uncertainty handling, and sensitivity stability. Although fuzzy AHP-TOPSIS provided moderate performance, the lack of precision associated with the T-spherical fuzzy set in capturing indeterminacy became important for criteria like emotional impact. These results place the proposed method as a state-of-the-art solution for multi-criteria decision-making problems within subjective evaluation. It stands out, however, because it allows for the efficient handling of qualitative and quantitative aspects.

6.6 Implications for music evaluation

The study has significant implications for the field of music evaluation:

- streamlining music competitions: an algorithm that furnishes a transparent and systematic basis for ordering compositions, eliminating subjectivity and inconsistency in manual evaluation
- applications in music streaming: the method could be adopted by music streaming platforms to curate playlists or suggest what composition to listen to, given a multi-criteria assessment of a composition’s quality, appeal and creativity
- educational use: the algorithm provides music educators with a tool to offer subjective feedback to students while keeping technical proficiency and artistic expression in balance.

6.7 Limitations and future directions

While the proposed algorithm offers numerous advantages, there are areas for improvement:

1 genre-specific customisation: in subjective genres such as electronic music, the algorithm’s performance indicates a need for tuning criteria weights or including genre-specific criteria

2 real-time evaluation: such an algorithm may better adapt and respond to changes in listener preferences by incorporating real-time listener feedback for information

3 scalability to larger datasets: although the algorithm demonstrated scalability for 1,000 compositions, processing datasets exceeding 10,000 items may require further optimisation or distributed computing frameworks.

Finally, future research may benefit from extending the algorithm to application in other domains, which we demonstrated as another place where subjective and technical criteria are intricately aligned. The efficiency and flexibility of the proposed hybrid MCDM algorithm with T-spherical fuzzy information in the music composition evaluation are discussed. The method does this by building its strength based on the complexity of subjective judgement and uncertainty that it seeks to master in creative decision making, particularly in technical decision making. The lessons learned in this study put it on a path to more generally applicable and improved applications in the future.

7 Conclusions

Using T-spherical fuzzy information, this research proposed a novel hybrid MCDM algorithm for evaluating music compositions. Finally, we proposed a method that effectively dealt with the inherent subjectivity, vagueness, and complexity of music evaluation, which combined the TSF-AHP for criteria weight computation and T-SF-TOPSIS for compositions ranking. The method was also found to yield superior ranking consistency, uncertainty management, computational efficiency, and sensitivity stability in a tightly controlled experimental setup over a 1,000 music composition dataset spanning various genres. Results indicated high alignment between algorithm rankings and experts [Spearman's rank correlation coefficient (ρ = 0.92, ρ = 0.88)] and listeners (ρ = 0.92, ρ = 0.88), respectively. The good agreement of this algorithm between qualitative and quantitative evaluation criteria, such as musicality, creativity, emotional impact, technical complexity, and audience appeal, validates the abilities of the algorithm to balance evaluation criteria. Special treatment of ambiguous and conflicting judgements was achieved by including T-spherical fuzzy numbers that provided a highly robust method in the context of subjective evaluations.

Additionally, we assessed the scalability and computational efficiency of the algorithm, and it completed processing the 1,000-composition dataset in 32 minutes. In sensitivity analysis, we found the rankings to be susceptible to criterion weights but still relatively stable; thus, the reliability and robustness of the method performed under different dynamic evaluation settings. The comparative analysis also demonstrated the superiority of the proposed approach to traditional methods such as crisp AHP-TOPSIS and fuzzy AHP-TOPSIS, especially in effectively handling indeterminacy and producing consistent and interpretable results. Areas for improvement were found in the proposed method, which had strong performance. Future work includes customisation for genre-specific feedback, honest time feedback, and optimisation towards more excellent datasets.

Furthermore, this dynamic nature makes it an appropriate tool for assessing other creative domains, including movies, games, and art, where the technical and subjective overlap. Finally, a hybrid MCDM algorithm based on T-spherical fuzzy information is proposed to evaluate music compositions robustly, systematically, and scalable. Finally, its capacity for navigating subjective judgement and uncertainty makes it an insightful, state-of-the-art decision-making solution for music evaluation and related fields. It allows it to extend to academic and industrial settings, respectively.

Declarations

All authors declare that they have no conflicts of interest.

References

Agres, K., Forth, J. and Wiggins, G.A. (2016) 'Evaluation of musical creativity and musical metacreation systems', *Computers in Entertainment (CIE)*, Vol. 14, No. 3, pp.1–33, https://doi.org/10.1145/2967506.

Alías, F., Socoró, J.C. and Sevillano, X. (2016) 'A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds', *Applied Sciences*, Vol. 6, No. 5, p.143, https://doi.org/10.3390/app6050143.

Al-Quran, A., Al-Sharqi, F., Khalifa, H.A.E-W., Alqahtani, H., Awad, A. and Bany, M. (2024) 'T-spherical fuzzy-valued neutrosophic set theory', *International Journal of Neutrosophic Science (IJNS)*, Vol. 23, No. 2, https://doi.org/10.59543/ijmscs.v3i.12613.

Bontempi, P., Canazza, S., Carnovalini, F. and Rodà, A. (2023) 'Research in computational expressive music performance and popular music production: a potential field of application?', *Multimodal Technologies and Interaction*, Vol. 7, No. 2, p.15, https://doi.org/10.3390/mti7020015.

Burnard, P. (2012) *Musical Creativities in Practice*, OUP Oxford, https://doi.org/10.1093/acprof:oso/9780199583942.003.0011.

Cancino-Chacón, C.E., Grachten, M., Goebl, W. and Widmer, G. (2018) 'Computational models of expressive music performance: a comprehensive and critical review', *Frontiers in Digital Humanities*, Vol. 5, p.25, https://doi.org/10.3389/fdigh.2018.00025.

Chen, T-Y. (2023) 'Multiple criteria decision analytic methods in management with T-spherical fuzzy information', *Artificial Intelligence Review*, Vol. 56, No. 12, pp.14087–14157, https://doi.org/10.1007/s10462-023-10461-z.

Dao, M.T. et al. (2019) 'A hybrid approach using fuzzy AHP-TOPSIS assessing environmental conflicts in the titan mining industry along central coast Vietnam', *Applied Sciences*, Vol. 9, No. 14, p.2930, https://doi.org/10.3390/app9142930.

de Berardinis, J., Cangelosi, A. and Coutinho, E. (2022) 'Measuring the structural complexity of music: from structural segmentations to the automatic evaluation of models for music generation', *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 30, pp.1963–1976, https://doi.org/10.1109/taslp.2022.3178203.

Deldjoo, Y., Schedl, M. and Knees, P. (2024) 'Content-driven music recommendation: evolution, state of the art, and challenges', *Computer Science Review*, Vol. 51, p.100618, https://doi.org/10.1016/j.cosrev.2024.100618.

Fiebrink, R. and Caramiaux, B. (2016) *The Machine Learning Algorithm as Creative Musical Tool*, arXiv preprint arXiv:1611.00379, https://doi.org/10.1093/oxfordhb/9780190226992.013.23.

Freitag, M. et al. (2021) 'Results of the WMT21 metrics shared task: evaluating metrics with expert-based human evaluations on TED and news domain', *Proceedings of the Sixth Conference on Machine Translation*, pp.733–774, https://doi.org/10.18653/v1/w16-2302.

Frith, S. (1996) *Performing Rites: On the Value of Popular Music*, Harvard University Press, https://doi.org/10.1093/oso/9780198163329.001.0001.

Gurmani, S.H., Zhang, Z., Zulqarnain, R.M. and Askar, S. (2023) 'An interaction and feedback mechanism-based group decision-making for emergency medical supplies supplier selection using T-spherical fuzzy information', *Scientific Reports*, Vol. 13, No. 1, p.8726, https://doi.org/10.1038/s41598-023-35909-8.

Hecken, T. (2024) 'Standardization', *The Popular as Art? An Investigation into Form, Value, Concepts, and Justifications*, pp.13–62, Springer, https://doi.org/10.1007/978-3-662-69711-5_2.

Juslin, P.N., Ingmar, E. and Danielsson, J. (2023) 'Aesthetic judgments of music: reliability, consistency, criteria, self-insight, and expertise', *Psychology of Aesthetics, Creativity, and the Arts*, Vol. 17, No. 2, p.193, https://doi.org/10.1037/aca0000403.

Kamehkhosh, I. (2017) *Advances in Next-Track Music Recommendation*, https://doi.org/10.1145/3019612.3019756.

Kramer, L. (1993) *Music as Cultural Practice, 1800–1900*, Univ of California Press, https://doi.org/10.1525/9780520354555-toc.

Langer, S.K. (2009) *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*, Harvard University Press, https://doi.org/10.2307/2103043.

Li, M., Wang, L. and Li, L. (2024) 'Research on narrative design of handicraft intangible cultural heritage creative products based on AHP-TOPSIS method', *Heliyon*, Vol. 10, No. 12, https://doi.org/10.1016/j.heliyon.2024.e33027.

Lin, C-L., Shih, Y-H., Tzeng, G-H. and Yu, H-C. (2016) 'A service selection model for digital music service platforms using a hybrid MCDM approach', *Applied Soft Computing*, Vol. 48, pp.385–403, https://doi.org/10.1016/j.asoc.2016.05.035.

Liu, H., Qian, Z. and Shao, J. (2021) 'Analysis of music influence on order preference based on TOPSIS algorithm', *Journal of Physics: Conference Series*, IOP Publishing, Vol. 1941, No. 1, IOP Publishing, p.012009, https://doi.org/10.1088/1742-6596/1941/1/012009.

Moorer, J.A. (1977) 'Signal processing aspects of computer music: a survey', *Proceedings of the IEEE*, Vol. 65, No. 8, pp.1108–1137, https://doi.org/10.1109/proc.1977.10660.

Mor, B., Garhwal, S. and Kumar, A. (2020) 'A systematic literature review on computational musicology', *Archives of Computational Methods in Engineering*, Vol. 27, pp.923–937, https://doi.org/10.1007/s11831-019-09337-9.

Motakiaee, R. (2011) *Multi Criteria Decision Making (MCDM) Models in Fuzzy and Non Fuzzy Environments*, SSRN 1761559, https://doi.org/10.2139/ssrn.1754233.

Muller, M., Ellis, D.P., Klapuri, A. and Richard, G. (2011) 'Signal processing for music analysis', *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No. 6, pp.1088–1110, https://doi.org/10.1109/jstsp.2011.2112333.

Ni, Y., McVicar, M., Santos-Rodriguez, R. and De Bie, T. (2012) 'An end-to-end machine learning system for harmonic analysis of music', *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 20, No. 6, pp.1771–1783, https://doi.org/10.1109/tasl.2012.2188516.

Pearce, M.T. (2005) *The Construction and Evaluation of , IOP Publishing Statistical Models of Melodic Structure in Music Perception and Composition*, City University London, https://doi.org/10.1525/mp.2005.22.4.663.

Schmitt, A. and Ultes, S. (2015) 'Interaction quality: assessing the quality of ongoing spoken dialog interaction by experts – and how it relates to user satisfaction', *Speech Communication*, Vol. 74, pp.12–36, https://doi.org/10.1016/j.specom.2015.06.003.

Shiffrin, R., Stigler, S.M. and Jamieson, K.H. (2023) 'Evaluating evidence requires expert judgment', *Transactions of the American Philosophical Society*, Vol. 112, No. 3, pp.5–29, https://doi.org/10.1353/tap.2023.a918765.

Sordo, M. (2012) *Semantic Annotation of Music Collections: A Computational Approach*, https://doi.org/10.1016/j.jcss.2011.10.010.

Wongvilaisakul, W., Netinant, P. and Rukhiran, M. (2023) 'Dynamic multi-criteria decision making of graduate admission recommender system: AHP and fuzzy AHP approaches', *Sustainability*, Vol. 15, No. 12, p.9758, https://doi.org/10.3390/su15129758.

Yang, Z., Zhang, T., Garg, H. and Venkatachalam, K. (2022) 'A multi-criteria framework for addressing digitalization solutions of medical system under interval-valued T-spherical fuzzy information', *Applied Soft Computing*, Vol. 130, p.109635, https://doi.org/10.1016/j.asoc.2022.109635.

Yılmaz, M. (2015) 'The selection of an entry-level music librarian: the analytic hierarchy process (AHP) as a new model', *Journal of Balkan Libraries Union*, Vol. 3, No. 2, pp.1–13, https://doi.org/10.17706//jsw.10.10.1216-1227.

Yue, Y. and Shen, X. (2024) 'Development strategy of early childhood music education industry: an IFS-AHP-SWOT analysis based on dynamic social network', *Plos ONE*, Vol. 19, No. 2, p.e0295419, https://doi.org/10.1371/journal.pone.0295419.

Zedam, L., Jan, N., Rak, E., Mahmood, T. and Ullah, K. (2020) 'An approach towards decision-making and shortest path problems based on T-spherical fuzzy information', *International Journal of Fuzzy Systems*, Vol. 22, pp.1521–1534, https://doi.org/10.1007/s40815-020-00820-1.

Zyoud, S.H. and Fuchs-Hanusch, D. (2017) 'A bibliometric-based survey on AHP and TOPSIS techniques', *Expert Systems with Applications*, Vol. 78, pp.158–181, https://doi.org/10.1016/j.eswa.2017.02.016.