Evaluation of lecturers’ performance using a novel hierarchical multi-criteria model based on an interval complex Neutrosophic set

Anh Duc Doa,e*, Minh Tam Phamb, Thi Hang Dinhc, The Chi Ngod, Quoc Dat Luue, Ngoc Thach Phamf, Dieu Linh Hagh and Hong Nhat Vuongh

aNational Economics University, Vietnam
bFaculty of Geography, VNU University of Science, Vietnam National University, Vietnam
cHanoi University of Natural Resources and Environment
dAcademy of Finance, Vietnam
eVNU University of Economics and Business, Vietnam National University, Vietnam
fHanoi University, Vietnam
gMinistry of Education and Training, Vietnam
hInstitute of Geography, Vietnam Academy of Science and Technology, Vietnam

Chronicle

Abstract

Performance assessment of teaching competency plays an important role in educational activities. Previous assessments of lecturers’ performance have failed to distinguish between potential capacity and their teaching effectiveness. To solve this problem, the integrated approach of quantitative assessment and multi-criteria decision-making models has become one of the main trends for assessing the performance of lecturers in multiple dimensions: self, peer-, manager- and student-based evaluation. This paper proposes a novel hierarchical approach, developed by the Technique for Order preference by Similarity to Ideal Solution method in an interval-valued complex neutrosophic set environment, to more accurately and comprehensively understand the evaluation process and fit it into a systematic framework. An application is given to illustrate a practical solution in lecturer’s evaluation. The accuracy of the proposed method is verified by comparing with other methods.

Keywords:
AHP
Lecturer evaluation
Interval complex neutrosophic
tOPSIS

1. Introduction

In recent years, university lecturers have become the key factors in educational goals of national development strategies (King, 2014; Wiliam et al., 2017). Therefore, performance assessments of their teaching competency were used as an effective tool for supporting personnel decision-making through rewards, punishments, employment and dismissal. The assessments can also serve as the main criteria for verifying qualification certificates in academic institutions and universities (Wu et al., 2012; Jiayi & Ling, 2012; Maltarich et al., 2017; Zhou et al., 2018). A university should not only function as a training institution but also as a scientific research center that encourages lecturers to carry out scientific research activities (Fauth et al., 2014; Cuevas et al., 2018). Moreover, this approach could enhance the creation of an equal environment that improves the cooperative strategies (Cegarra et al., 2016; Wu et al., 2018), learning spirit (Cegarra et al., 2017) and autonomy of each student (Parrish, 2016; Darling-Hammond, 2017; Fischer et al., 2018). An effective system of assessing lecturers’ performance can...
directly help to estimate educational achievements from many perspectives, such as: improving meaningful and sustainable learning (Almeida, 2017); finding and fostering young talents (Bohlmann & Weinstein, 2013); and indirectly impacting the wealth of each country (Lazarides et al., 2018), and becoming a preferred policy at the global and local levels (Steinberg & Garrett, 2016; Tuytens & Devos, 2017). Lecturer assessment has been regarded as a complex issue with several complicated factors such as personal interests and the development strategy of the education system (Schön, 2017). One of the most difficult issues at any university is to have a fair and accurate assessment of its lecturers’ activities from which to delegate their respective tasks and positions. The absence of an appropriate set of standards and tools may lead to inaccuracy and subjectivity in assessing the competence of each lecturer. There is a need to constantly evaluate lecturers from principals/managers (OECD, 2009; Marzano & Toth, 2013), through self-report (Singh & Jha, 2014), students (Kilic, 2010; Nilson, 2016; Lans et al., 2018) and peer-review from colleagues (Alias et al., 2015). Needless to say, most lecturers expect to receive good and fair reports, regardless of reality (Liu & Zhao, 2013; Nahid et al., 2016). Indeed, a multi-dimensional assessment could augment lecturers’ knowledge background, expand their teaching repertoires and develop them professionally (Malakolunthu & Vasudevan, 2012; Skedsmo & Huber, 2018). It has been argued that a multi-objective formal process can improve lecturers’ ability to make professional decisions and judgments (Bambaeeroo & Shokrpour, 2017). Furthermore, since each locality and context is unique, lecturer evaluation should consider different local characteristics and various methodology and data resources (Sonnert et al., 2018). Criteria for this assessment include standards related to research capacity, teaching capacity and service activities that act as a multi-standard decision-making process (Wu et al., 2012).

Currently, multi-criteria decision-making (MCDM) is used to navigate real-world problems and the uncertainty of human thinking at large (Li et al., 2015; Yang & Pang, 2018). The Analytical Hierarchy Process (AHP) and Technique for Order preference by Similarity to Ideal Solution (TOPSIS) are the most popular MCDM models that allow identifying and selecting the best solutions that use heterogeneous data (Torkabadi et al., 2018). The AHP model can be used to analyze complex problems by separating branching structure systems to calculate the weight of each criterion (Saaty, 2008). Although this model has some weaknesses regarding the number of criteria for quantitative analysis, it does not require clear information (Ishizaka & Labib, 2009; Karthikeyan et al., 2016). In contrast, TOPSIS allows determining the ranking through the use of many criteria (Hwang & Yoon, 1981; Chi & Liu, 2013; Wang & Chan, 2013). The basic principle of TOPSIS technique is that the most preferred alternative should simultaneously have the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution. It also reflects the rationale of human choice (Baykasoğlu et al., 2013). This method requires using the input data to find the weight of each criterion. Thus, integrating these popular MCDM techniques could effectively improve quantitative assessments for determining the performance and relative importance of lecturers.

Smarandache (1998) proposed the neutrosophic set, which is independently characterized by a truth-membership degree (T), an indeterminacy-membership degree (I) and a falsity-membership degree (F), all of which are within the real standard or nonstandard unit interval (Chi & Liu, 2013). If a range is restrained within this interval, the neutrosophic set can be easily applied to problems in education (Akram et al., 2018). In this aspect, Wang et al. (2010) introduced the concept of the single-valued neutrosophic set as a sub-class of the neutrosophic set. Wang also proposed the use of interval-valued neutrosophic set as a subclass of neutrosophic sets with the values of truth-membership, indeterminacy-membership, and falsity-membership. This set has been applied in different fields, such as decision-making sciences, social sciences, and humanities to solve problems involving imprecise, indeterminate and inconsistent information (Zhang et al., 2014). Later, Ye (2014) introduced another concept, the interval neutrosophic linguistic set, that involves new aggregation operators for interval neutrosophic linguistic information. In the same effort, Said et al. (2015) proposed another decision-making method that extends the TOPSIS method to deal with uncertain linguistic information in interval neutrosophic sets. However, to the best of our knowledge, there has been no research on integrating hierarchical TOPSIS in interval neutrosophic complex sets, especially for lecturer evaluation. Therefore, the
combination of two useful techniques of MCDM such as AHP and TOPSIS and the interval-valued complex set in neutrosophic environment can reduce the shortcomings of traditional approach for lecturer evaluation (Biggs & Collis, 2014; Gormally et al., 2014).

In this paper, lecturer evaluation is the particular case study of the MCDM models. However, the complexity and uncertainty of this approach mean that it is necessary to integrate the hierarchical neutrosophic TOPSIS and the interval-valued complex set. Thus, this study presents the results of weighting performance evaluation criteria to rank five different lecturers of the University of Economics and Business - Vietnam National University, Hanoi. The rest of the study is organized as follows: Section 2 displays a review of the principal characteristics of lecturer evaluation. Section 3 presents the methodology of using the hierarchical neutrosophic TOPSIS to rank alternatives. An illustrative application is then presented in Section 4 to describe how the model works. Finally, conclusions and discussion are given in Section 5.

2. Literature review on lecturer evaluation methods and criteria

2.1. Lecturer evaluation methods

In recent decades, lecturer evaluation has received much attention from researchers seeking to enhance professional teaching (King, 2014). According to Colby et al. (2002), lecturer evaluation concerns competency, professionalism, advancement and student achievement. Buttram and Wilson (1987) suggested that the best evaluation is identifying the effective approaches used in teaching and knowledge at the university level. Doing this can improve the quality of students in the future. In another study (Davey, 1991), a lecturer was evaluated based on the dimensions of effective job performance, comprehensive excises and the use of multiple objects to eliminate bias. This process required frequent assessments and appropriate development strategies from the relevant institution. It has been argued that, assessment is primarily an organizational problem, not a technical problem (Schön, 2017). However, ineffective efforts are typically diagnosed in terms of a useless assessment instrument, prompting the search for better instruments (Lans et al., 2018). Evaluation experiences have long been considered influential in organizational behavior as sources of support for feedback, need satisfaction, feelings of competence and psychological success. Moreover, a lecturer evaluation system should include different components: vocational morality, attendance rate in school meetings and events, teaching and researching ability and student performance (based on student tests and report scores) (Chi & Liu, 2013; Reddy et al., 2018).

Lecturer evaluation requires the establishment of reference standards and evaluation criteria (OECD, 2009). Traditionally, this approach depended on classroom observations conducted by managers of a university (Danielson, 2000). This approach gave powerful tools for human resources, but the effects of this system are mixed (Zerem, 2017). Indeed, manager-based evaluation has many disadvantages regarding transparency and promoting the image of the university. Furthermore, the traditional approach used the test scores of students to determine lecturer performance (Tondeur et al., 2017). It was based on an image of the lecturer and beliefs about teaching that is inconsistent (Zare et al. 2016). Consequently, it had negative impacts on professional development and failed to improve the quality of teaching (Chappuis et al., 2016). Lecturers tended to try to impress managers and compete with their peers all of cost (Liu & Teddlie, 2007). Current curriculum reforms focus on the participation of managers, peers, students and the lecturers themselves for self-evaluation (Ovando, 2001; Muijs & Reynolds, 2017). Even if given feedback from such an evaluation system, lecturers might not be inclined to reflect on their practices (Cheng et al., 2009; Kurtz et al., 2017). In the best practices, the evaluation of teaching should provide an opportunity for dialogue between lecturers and evaluators based on a shared understanding of good teaching (Nilson 2016).

2.2. Main criteria used in the lecturer evaluation framework

In this study, a multi-criteria evaluation process was introduced to assess the efficiency and capacity of lecturers at the university level. Based on previous studies, the following criteria were divided into four
main groups: self-, manager-, peer- and student-based evaluation - and 13 sub-criteria (Wu et al., 2012) as shown in Fig. 1. These four aspects can be used to clearly evaluate and improve lecturer performance (Odden, 2014).

2.2.1. Scientific publication (C₁₁)

Scientific publications are an effective criterion for academics, methodical evaluation and human recruitment. Published articles treat complex problems (Zare et al., 2016). Also, having journal articles published important in the academic community in developing countries (Jaramillo et al., 2017). Writing an article for publication is difficult, so publications can be used to determine and classify a lecturer’s academic ability (Wu et al., 2012). Thus, in this study, this criterion was assessed as the ratio between the number of articles over two per year and the total in a year. Another important aspect is the duration of research: a good lecturer usually requires less time to publish an article (Zerem, 2017).

2.2.2. Supervising postgraduate students (C₁₂)

Students expect their supervisors to have sufficient professional ability and knowledge to provide advice on research (Wu et al., 2012). Thus, lecturers must be up-to-date and pursue research activities in numerous aspects and multiple fields (Sharp et al., 2017). They must supervise many students and act as learning mentors (Wisker, 2012). In addition, publications made during the supervision of trainees have certain scientific value. When a lecturer’s trainees require less time to publish, it reflects well on the lecturer.

2.2.3. The journal peer-review process (C₁₃)

Researchers usually seek various opportunities to indicate their knowledge and skills. One of the key steps in the climb to academic success is becoming a peer reviewer (Wolf, 2016). Content and methodology experts review papers and create recommendations to increase the value of the publications for a specific journal (Thomas et al., 2014). They supply feedback for articles and research, suggest improvements and make recommendations to the editors about whether to accept, reject or request changes to articles (Iantorno et al., 2016). Thus, to become a journal peer reviewer, researchers must spend a great deal of time accumulating professional experience (Wu et al., 2012). As one of the aspects of evaluation, this study used the length of time before becoming a journal reviewer.

2.2.4. Lecturing activities (C₂₁)

This criterion involves preparation time and statutory teaching time. It is the number of hours spent teaching a group or class of students according to the formal policy in a country. At universities, lecturing time is counted by the number of lessons. The duration of each lesson is regulated at 45 or 50 minutes, and this was used to determine the time spent teaching. The ratio between the number of lessons and subjects per year and the total in a year was used to evaluate standard lecturing time. Additionally, the number of scientific publications can be compared with lecturing time (Zare et al., 2016).

2.2.5. Language of instruction (C₂₂)

A search of the relevant literature revealed a lack of research on pre-service English lecturer teaching programs. There is a concern for the standards of teaching and learning in a non-native language (Wu et al., 2012). Consequently, lecturers are compelled to constantly adapt their lecture, which affects the standards and amount of content taught during a semester. Thus, limitations in lecturers’ linguistic competencies have negative effects on program quality (Bradford, 2015).
2.2.6. Lecturing attitude and spirit (C23)

Lecturing attitude and spirit are a sum of several behaviors. However, a gradual decline in attitude over a lecturer’s career may flatten variations between these behaviors (Frunză, 2014). Lecturing attitude is a common concern in psychology pedagogy research (Zyad, 2016). For example, some lecturers come to class late, which has negative impacts on student acceptance.

2.2.7. Evaluation and scoring system (C24)

The procedures used for training lecturers to score students’ work dependably are consistent across colleges (Wu et al. 2012). Student work samples represent completely different levels of performance on rubrics. To score fairly, lecturers should give examples and clarify the distinctions between score levels. Lecturers rated a pre-selected example to evaluate scoring (Tondeur et al., 2017).

2.2.8. Cooperation in research projects (C31)

Research projects require cooperation between researchers and a good working environment (Hein et al., 2015). Management is accountable for the conduct of editors, for safeguarding research records and for ensuring the reliability of published research. It is important for researchers to communicate and collaborate effectively on cases related to research integrity (Wager & Kleinert, 2012). Cooperation involves not only reducing the time spent on projects but also advancing and exchanging knowledge. This study used for this criterion the number of co-worker cooperation projects over two projects and the duration of these projects.

2.2.9. Teamwork in scientific and teaching activities (C32)

The advantages of cooperation and teamwork for researchers include assistance for testing and measuring, access to vast amounts of knowledge and assistance in developing new initiatives (Johnson et al., 2012). Furthermore, different researchers contribute different types and amounts of resources, which increases the number of publications of all involved in the cooperation (Wardil & Hauert, 2015). Research teamwork refers to a broad variety of activities, from simple opinion exchanges to side-by-side work in the laboratory. Thus, it is important to the evaluate lecturers’ cooperation and teamwork.

2.2.10. Participation in school meetings and events (C33)

Lecturers build good relationships with their co-workers and students when they take part in school meetings and events (Wang & Hsieh 2017). It is widely acknowledged that school meetings and events are important for guaranteeing cooperation, ensuring that lecturers are professionally ready for work and identifying basic problems related to their work (Frunză, 2014; Zyad, 2016). Lecturers can be evaluated at a high level for behavior that demonstrates professional responsibility (Frunză, 2014). Thus, the ratio between the number of attended school meetings and events and total compulsory school meetings and events was used to evaluate lecturers.

2.2.11. The content of the lessons (C41)

Regarding teaching and learning, students especially evaluate lecturers based on the quality of teaching content (Nilson, 2016). Alongside these two players (lecturer and students), this approach evaluates school factors that are expected to influence teaching and learning (Shingphachanh, 2018). Moreover, lecturers should offer real-world examples to create interest for students (Brookfield, 2017). This research used the number of students who understood lessons and the theoretical learning duration to finish the subject to a satisfactory level for this criterion.
2.2.12. Lecturer-student interaction (C42)

Lecturer-student relationships are associated with both attrition and the general mental and physical health of lecturers (Kupers et al., 2015). These relationships are usually characterized by respect, warmth, and trust, as well as low levels of social conflict. Likewise, lecturers have more experience, education, and skills than their students, and thus they have a unique set of responsibilities to students (Aldrup et al., 2018). They are expected and trained to act in the best interests of their students. Therefore, they should be motivated to act appropriately and responsibly toward students.

2.2.13. The irrelevance of the subjects (C43)

This criterion involves the essential and traditional method whereby practitioners request information on whether their teaching has impacts (Nilson, 2016). Subjects should be sporadically assessed and reviewed. The issues include content and objectives, teaching plans, assessment procedures, the behaviors of students in the class and the experience of the lecturers (Brookfield, 2017). This includes the expectations for students’ educational outcomes in a subject matter, as well as the appropriateness of the objectives and content in achieving these outcomes. Thus, any irrelevance in teaching can have negative impacts on the behaviors and outcomes of students and trainees. The list criteria used in evaluating a lecturer performance can be exhaustive. However, they can be summarized in Figure 1, which shows that lecturer evaluation depends on four main groups for assessment. Each of these patterns consists of three sub-criteria, except for the second group of manager-based evaluation, which includes four sub-criteria. This study provides an integrated approach to find the best alternative. It presents a hierarchical structure and provides the most appropriate approach to evaluate lecturer.

**Fig. 1.** The criteria used in lecturer evaluation

Table 1 also summarizes and explains the selected criteria based on the literature review. Particularly, each criterion was identified to have three corresponding aspects with the complex neutrosophic set (truth, indeterminacy, and falsity) and its real and imaginary parts. Two features of each criterion describe the amplitude and phase terms in this set, which are represented by intervals. These are the background for determining the input values based on the available data in the educational system. Consequently, experts can make changes to the levels of these parameters for a given year based on three patterns, as shown in Table 1.
Table 1
The hierarchical model for lecturer evaluation and calculation methods

<table>
<thead>
<tr>
<th>Description/Unit</th>
<th>TRUTH (T)</th>
<th>INDETERMINAC Y (I)</th>
<th>FALSITY (F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-evaluation (C1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1. The scientific publication activities (C11) (+) (Wu et al. 2012; Zare et al. 2010; Jaramillo et al. 2017; Zeren 2017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of articles at international standard level (h11)/Total number of articles published expected per year (t11)</td>
<td>h11 which completed/ t11</td>
<td>h11 which did not complete or rejected/ t11</td>
<td>h11</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The duration to carry out researches (months) (h12)/Total standard level (h 11)/ Total number of articles at international standard level (h13)</td>
<td>h12 are under 05 months/12 months</td>
<td>h12 are 04-06 months/12 months</td>
<td>h12 are over 12 months/12 months</td>
</tr>
<tr>
<td>1.2. Supervising postgraduate students (C12) (+) (Wu et al. 2012; Wisker 2012; Sharp et al. 2017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of trainees who were guided (h21)/Total trainees per year (t21)</td>
<td>h21 which completed/ t21</td>
<td>h21 which did not complete or rejected/ t21</td>
<td>h21</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The number of standard graduation reports (h22)/Total graduation reports (t22)</td>
<td>h22 published to articles (h23)</td>
<td>h22 can publish lately to articles (h24)</td>
<td>h22 did not publish to articles (t22)</td>
</tr>
<tr>
<td>1.3. The journal peer-review process (C13) (+) (Wu et al. 2012; Zare et al. 2016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of journal publications reviewed (h31)/Total number of journal publications were suggested per year (t31)</td>
<td>h31 which completed/ t31</td>
<td>h31 which did not complete or rejected/ t31</td>
<td>h31</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The duration to become the journal reviewer (months) (h32)/Total duration joined in the scientific publication process (t33)</td>
<td>h32 are under 03 years/ t32</td>
<td>h32 are from 03-05 years/ t32</td>
<td>h32 are over 05 years/ t32</td>
</tr>
<tr>
<td>2. Lecturing activities (C2) (+) (Wu et al. 2012; Jaramillo et al. 2017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of lessons per year (h41)/Total standard lessons (t41)</td>
<td>h41 had over 70 lessons per year/ t41</td>
<td>h41 from 50-70 lessons per year/ t42</td>
<td>h41 which are submitting or processing/ t43</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The number of subjects which lecturers were assigned (h42)/Total subjects which student registered (t43)</td>
<td>h42 accounts for under 05% of t42</td>
<td>h42 accounts for 01-02% of t42</td>
<td>h42 accounts for over 50% of t42</td>
</tr>
<tr>
<td>2.2. Lecturing styles (C22) (+) (Wu et al. 2012; Bradford 2015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of courses taught in English (h51)/Total courses (t51)</td>
<td>h51 had over 50 peoples/ t51</td>
<td>h51 had from 30-50 peoples/ t52</td>
<td>h51 had under 30 peoples/t52</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The number of students who did not understand lessons (h52)/Total students (t53)</td>
<td>h52 failed had score at medium level in the English program (h53)</td>
<td>h52 failed in the English program (t54)</td>
<td>h52</td>
</tr>
<tr>
<td>2.3. Lecturing attitude and spirit (C23) (+) (Frunza 2014; Zeyd 2016; Wang and Hsieh 2017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of lessons which lecturers came to class lately (h61)/Total lessons (t61)</td>
<td>h61 accounts for over 50% of t61</td>
<td>h61 accounts for 30-50% of t62</td>
<td>h61 accounts for under 30% of t63</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The duration which lecturers came to class lately and total duration in a year (t64)</td>
<td>h64 accounts for over 20% of t64</td>
<td>h64 accounts for 01-02 months</td>
<td>h64 accounts for over 05 months/12 months</td>
</tr>
<tr>
<td>2.4. Score evaluation process for students (C24) (+) (Bradford 2015; Tondeur et al. 2017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of exams which lecturers organized (h71)/Total standard exams (t71)</td>
<td>h71 accounts for over 80% (h72)</td>
<td>h71 accounts for 60-80% (h73)</td>
<td>h71 accounts for under 60% (h74)</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The average duration which lecturers paid each exam (h74)</td>
<td>h74 are under 01 month</td>
<td>h74 are from 01-02 months</td>
<td>h74 are over 02 months</td>
</tr>
<tr>
<td>3. Peer-evaluation (C3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1. Cooperation in research projects (C31) (+) (Wager and Kleinert 2012; Wu et al. 2012; Hein et al. 2015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of co-worker cooperation projects (h81)/Total projects (t81)</td>
<td>h81 at ministerial level/ t81</td>
<td>h81 at school level/ t82</td>
<td>h81 did not belong to the above two categories/ t83</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The average duration to carry out each project (months) (h84)/12 months</td>
<td>h84 at over 12 months/12 months</td>
<td>h84 at over 05-07 months/12 months</td>
<td>h84 which are over 12 months/12 months</td>
</tr>
<tr>
<td>3.2. Teamwork in scientific and lecturing activities (C32) (+) (Johnson et al. 2012; Wardil and Hauert 2015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of initiatives to improve lecturing effectiveness (h91)/Total initiatives (t91)</td>
<td>h91 at ministerial level/ t91</td>
<td>h91 at school level/ t92</td>
<td>h91 did not belong to the above two categories/ t93</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The duration to complete initiatives (months) (h94)/12 months</td>
<td>h94 which are under 05 months/12 months</td>
<td>h94 which are from 05-07 months/12 months</td>
<td>h94 which are over 12 months/12 months</td>
</tr>
<tr>
<td>3.3. Participation school meetings and events (C33) (+) (Frunza 2014; Zeyd 2016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real part:</strong> The number of participation school meetings and events (h101)/Total compulsory school meetings and events (t101)</td>
<td>h101 accounts for over 80% (h102)</td>
<td>h101 accounts for 60-80% (h103)</td>
<td>h101 accounts for under 60% (h104)</td>
</tr>
<tr>
<td><strong>Imaginary part:</strong> The number of school meetings and events were on time (h104)/Total compulsory school meetings and events (t104)</td>
<td>h104 which accounts for over 80% (h105)</td>
<td>h104 which accounts for 60-80% (h106)</td>
<td>h104 which accounts for under 60% (h107)</td>
</tr>
</tbody>
</table>
3. Methodology

3.1. Interval Complex Neutrosophic Set

The neutrosophic set is a generalization of the classic set, fuzzy set (Zadeh, 1965), interval-valued fuzzy set (Turksen 1986) and intuitionistic fuzzy set (Atanassov, 1986) that was proposed by Smarandache (1998). Many real-life problems have not only truth and falsehood, but indeterminacy between several suitable opinions (Ali et al., 2018). This method represents an extension of the standard interval [0, 1] used for interval fuzzy sets. To deal with this situation, the concept of interval neutrosophic sets (INSs) can be used to make these values intervals rather than real numbers. Furthermore, the Hamming and Euclidean distances between the INS and the similarity measures are based on distances (Ye, 2014). Moreover, based on complex numbers, Ali and Smarandache introduced the complex neutrosophic set to handle the amplitude and phase terms of the set’s members (Ali & Smarandache, 2017). In real problems, it is difficult to find a crisp neutrosophic membership degree with unclear information, so Ali proposed the interval complex neutrosophic set (ICNS) (Ali et al., 2018). In this set, the terms of the ICNS can handle unsure values in the membership. This section provides some basic definitions of the neutrosophic set proposed by Smarandache (1998).

3.2. The definition, operation rules and distance of ICNS

The interval neutrosophic linguistic set, developed based on the theory of the INS, allows solving complex problems in quantitative assessments, as shown in the following (Ye, 2014).

**Definition 1. Neutrosophic set** (Smarandache 1998):

Let $X$ be a universe of discourse, with a generic element in $X$ denoted by $x$. A neutrosophic set (NS) $\vec{A}$ in $X$ is:

$$\vec{A} = \{(x, T_\vec{A}(x), I_\vec{A}(x), F_\vec{A}(x)) | x \in X\},$$

where, the functions $T_\vec{A}(x), I_\vec{A}(x), F_\vec{A}(x)$ of $[0^-, 1^+]$ define the degree of truth-membership function, indeterminacy-membership function, and the falsity-membership function respectively. There is no restriction $T_\vec{A}(x), I_\vec{A}(x), F_\vec{A}(x)$ so (Wang et al., 2010):

$$0^- \leq T_\vec{A}(x) + I_\vec{A}(x) + F_\vec{A}(x) \leq 3^+.$$
Definition 2. Interval neutrosophic set:
Let \( X \) be a universe of discourse, with a generic element in \( X \) denoted by \( x \). An interval neutrosophic set \( \overline{A} \) in \( X \) is:
\[
\overline{A} = \left\{(x, T_{\overline{A}}(x), I_{\overline{A}}(x), F_{\overline{A}}(x)) \mid x \in X \right\},
\]
where, the functions \( T_{\overline{A}}(x), I_{\overline{A}}(x), F_{\overline{A}}(x) \subseteq [0,1] \) define the degree of truth-membership function, indeterminacy-membership function, and the falsity-membership function respectively, so:
\[
0 \leq \sup \left( T_{\overline{A}}(x) \right) + \sup \left( I_{\overline{A}}(x) \right) + \sup \left( F_{\overline{A}}(x) \right) \leq 3.
\]

Definition 3. Complex fuzzy set
A complex fuzzy set \( \overline{S} \), defined on a universe of discourse \( U \), is characterized by a membership function \( \eta_{\overline{S}}(x) \) that assigns to any element \( x \in U \) a complex-valued grade of membership in \( \overline{S} \). The values lie within the unit circle in the complex plane, and thus, all forms \( p_{\overline{S}}(x) \cdot e^{j\mu_{\overline{S}}(x)} \) where \( p_{\overline{S}}(x) \) and \( \mu_{\overline{S}}(x) \) are both real-valued and \( p_{\overline{S}}(x) \in [0,1] \) with \( j = \sqrt{-1} \). The term \( p_{\overline{S}}(x) \) is termed as amplitude term, and \( e^{j\mu_{\overline{S}}(x)} \) is termed as phase term. The complex fuzzy set can be represented as:
\[
\overline{S} = \left\{(x, \eta_{\overline{S}}(x)) \mid x \in U \right\}.
\]

An interval-valued complex fuzzy set \( \overline{A} \) is defined over a universe of discourse \( X \) by a membership function
\[
\mu_{\overline{A}} : X \to \Gamma^{[0,1]} \times R,
\]
\[
\mu_{\overline{A}} = r_{\overline{A}}(x) \cdot e^{j\omega_{\overline{A}}(x)}
\]
In the above equation, \( \Gamma^{[0,1]} \) is the collection of interval fuzzy sets and \( R \) is the set of real numbers. \( r_{\overline{A}}(x) \) is the interval-valued membership function while \( e^{j\omega_{\overline{A}}(x)} \) is the phase term, with \( j = \sqrt{-1} \).

Definition 5. Union of interval complex neutrosophic sets (ICNSs)
Two complex fuzzy sets \( \overline{A} \) and \( \overline{B} \) were defined by Ramot et al. as follows:
Let \( \mu_{\overline{A}} = r_{\overline{A}}(x) \cdot e^{j\omega_{\overline{A}}(x)} \) and \( \mu_{\overline{B}} = r_{\overline{B}}(x) \cdot e^{j\omega_{\overline{B}}(x)} \) be the complex-valued membership functions of \( \overline{A} \) and \( \overline{B} \), respectively. The, the membership union of \( \overline{A} \cup \overline{B} \) is given by \( \mu_{\overline{A} \cup \overline{B}} = \left[ r_{\overline{A}}(x) + r_{\overline{B}}(x) \right] \cdot e^{j\omega_{\overline{A} \cup \overline{B}}(x)} \). Since, \( r_{\overline{A}}(x) \) and \( r_{\overline{B}}(x) \) are real-valued and belong to \([0,1] \), the operators max and min can be applied to them. For calculating phase term \( \omega_{\overline{A} \cup \overline{B}}(x) \), they give several methods. Let \( \overline{A} \) and \( \overline{B} \) be two IVCNSs in \( X \), where
\[
\overline{A} = \left\{(x, T_{\overline{A}}(x), I_{\overline{A}}(x), F_{\overline{A}}(x)) \mid x \in X \right\} \quad \text{and} \quad \overline{B} = \left\{(x, T_{\overline{B}}(x), I_{\overline{B}}(x), F_{\overline{B}}(x)) \mid x \in X \right\}.
\]
Then, the union of two interval neutrosophic sets was defined as follows:
\[
\overline{A} \cup \overline{B} = \left\{(x, T_{\overline{A} \cup \overline{B}}(x), I_{\overline{A} \cup \overline{B}}(x), F_{\overline{A} \cup \overline{B}}(x)) \mid x \in X \right\}.
\]
\[ T_{\odot, \odot}(x) = \left[ \inf p_{\odot, \odot}(x), \sup p_{\odot, \odot}(x) \right] \cdot e^{j\pi_\odot(x)} \]
\[ I_{\odot, \odot}(x) = \left[ \inf q_{\odot, \odot}(x), \sup q_{\odot, \odot}(x) \right] \cdot e^{j\rho_\odot(x)} \]
\[ F_{\odot, \odot}(x) = \left[ \inf r_{\odot, \odot}(x), \sup r_{\odot, \odot}(x) \right] \cdot e^{j\alpha_\odot(x)} \]

where

\[ \inf p_{\odot, \odot}(x) = \sqrt{\left( \inf p_{\odot}(x), \inf p_{\odot}(x), \sup p_{\odot, \odot}(x) \right)} = \sqrt{\left( \sup p_{\odot}(x), \sup p_{\odot}(x) \right)} , \]
\[ \inf q_{\odot, \odot}(x) = \wedge \left( \inf q_{\odot}(x), \inf q_{\odot}(x), \sup q_{\odot, \odot}(x) \right) = \wedge \left( \sup q_{\odot}(x), \sup q_{\odot}(x) \right) , \]
\[ \inf r_{\odot, \odot}(x) = \wedge \left( \inf r_{\odot}(x), \inf r_{\odot}(x), \sup r_{\odot, \odot}(x) \right) = \wedge \left( \sup r_{\odot}(x), \sup r_{\odot}(x) \right) , \]

where \( \wedge \) and \( \vee \) denote the max and min operator respectively. To calculate the phase term

\[ e^{j\pi_\odot(x)} , e^{j\rho_\odot(x)} , e^{j\alpha_\odot(x)} . \]

**Definition 6. Intersection of interval-valued complex neutrosophic sets (ICNSs)**

Two complex fuzzy sets \( \widetilde{A} \) and \( \widetilde{B} \) were defined by Ramot et al. as follows:

Let \( \mu_\widetilde{A} = r_\widetilde{A}(x) \cdot e^{j\pi_\widetilde{A}(x)} \) and \( \mu_\widetilde{B} = r_\widetilde{B}(x) \cdot e^{j\rho_\widetilde{B}(x)} \) be the complex-valued membership functions of \( \widetilde{A} \) and \( \widetilde{B} \), respectively. Let \( \widetilde{A} \) and \( \widetilde{B} \) be two IVCNSs in \( X \), where

\[ \widetilde{A} = \{(x, T_\widetilde{A}(x), I_\widetilde{A}(x), F_\widetilde{A}(x)) \mid x \in X \} \quad \text{and} \quad \widetilde{B} = \{(x, T_\widetilde{B}(x), I_\widetilde{B}(x), F_\widetilde{B}(x)) \mid x \in X \} \]

Then, the intersection of two interval neutrosophic sets was defined as follows: Let \( \widetilde{A} \) and \( \widetilde{B} \) be two IVCNSs in \( X \). Then:

\[ \widetilde{A} \odot \widetilde{B} = \{(T_{\odot, \odot}(x), I_{\odot, \odot}(x), F_{\odot, \odot}(x)) \mid x \in X \}, \]

where

\[ T_{\odot, \odot}(x) = \left[ \inf p_{\odot, \odot}(x), \sup p_{\odot, \odot}(x) \right] \cdot e^{j\pi_{\odot, \odot}(x)} , \]
\[ I_{\odot, \odot}(x) = \left[ \inf q_{\odot, \odot}(x), \sup q_{\odot, \odot}(x) \right] \cdot e^{j\rho_{\odot, \odot}(x)} , \]
\[ F_{\odot, \odot}(x) = \left[ \inf r_{\odot, \odot}(x), \sup r_{\odot, \odot}(x) \right] \cdot e^{j\alpha_{\odot, \odot}(x)} , \]

\[ \inf p_{\odot, \odot}(x) = \bigvee \left( \inf p_{\odot}(x), \inf p_{\odot}(x), \sup p_{\odot, \odot}(x) \right) = \bigvee \left( \sup p_{\odot}(x), \sup p_{\odot}(x) \right) , \]
\[ \inf q_{\odot, \odot}(x) = \bigwedge \left( \inf q_{\odot}(x), \inf q_{\odot}(x), \sup q_{\odot, \odot}(x) \right) = \bigwedge \left( \sup q_{\odot}(x), \sup q_{\odot}(x) \right) , \]
\[ \inf r_{\odot, \odot}(x) = \bigwedge \left( \inf r_{\odot}(x), \inf r_{\odot}(x), \sup r_{\odot, \odot}(x) \right) = \bigwedge \left( \sup r_{\odot}(x), \sup r_{\odot}(x) \right) , \]

\( \wedge \) and \( \vee \) denote the max and min operator respectively. To calculate the phase term

\[ e^{j\pi_{\odot, \odot}(x)} , e^{j\rho_{\odot, \odot}(x)} , e^{j\alpha_{\odot, \odot}(x)} . \]

**3.3. The operation rules of interval-valued complex neutrosophic set**

**Definition 7. The operational rules of the interval complex neutrosophic sets**

Let \( \widetilde{A} = \left[ [ T^L_{\widetilde{A}}, T^U_{\widetilde{A}} ], [ I^L_{\widetilde{A}}, I^U_{\widetilde{A}} ], [ F^L_{\widetilde{A}}, F^U_{\widetilde{A}} ] \right] \) and \( \widetilde{B} = \left[ [ T^L_{\widetilde{B}}, T^U_{\widetilde{B}} ], [ I^L_{\widetilde{B}}, I^U_{\widetilde{B}} ], [ F^L_{\widetilde{B}}, F^U_{\widetilde{B}} ] \right] \) be two interval neutrosophic sets over \( X \) which are defined by \( [ T^L_{\widetilde{A}}, T^U_{\widetilde{A}} ] = [ p^L_{\widetilde{A}}, p^U_{\widetilde{A}} ] \cdot e^{j[\pi^L_{\widetilde{A}}, \pi^U_{\widetilde{A}}]} \), \( [ I^L_{\widetilde{A}}, I^U_{\widetilde{A}} ] = [ q^L_{\widetilde{A}}, q^U_{\widetilde{A}} ] \cdot e^{j[\rho^L_{\widetilde{A}}, \rho^U_{\widetilde{A}}]} \), \( [ F^L_{\widetilde{A}}, F^U_{\widetilde{A}} ] = [ r^L_{\widetilde{A}}, r^U_{\widetilde{A}} ] \cdot e^{j[\alpha^L_{\widetilde{A}}, \alpha^U_{\widetilde{A}}]} \) respectively. Then, the operational rules of ICNS are defined as follows. The complement of \( x \) is:
\[ \bar{A} = \left( [p_A^L, p_A^U], [q_A^L, q_A^U], [r_A^L, r_A^U] \right) \quad \text{and} \quad \bar{B} = \left( [p_B^L, p_B^U], [q_B^L, q_B^U], [r_B^L, r_B^U] \right) \]

(i) The complement of \( \bar{A} \) is defined as:
\[
\bar{A} = \left( [r_A^L, r_A^U], e^{j[2-\mu_A^L(x),2-\mu_A^U(x)]}, [1-q_A^L, 1-q_A^U] \cdot e^{j[2+\nu_A^L(x),2+\nu_A^U(x)]}, [p_A^L, p_A^U], e^{j[2-\omega_A^L(x),2-\omega_A^U(x)]} \right)
\]

(ii) The addition of \( \bar{A} \) and \( \bar{B} \) is defined as:
\[
\bar{A} \oplus \bar{B} = \left( [p_A^L + p_B^L, p_A^U + p_B^U], [q_A^L + q_B^L, q_A^U + q_B^U], [r_A^L + r_B^L, r_A^U + r_B^U], e^{j[\omega_A^L(x)+\omega_B^L(x),\omega_A^U(x)+\omega_B^U(x)]} \right)
\]

(iii) The bounded between \( \bar{A} \) and \( \bar{B} \) is defined as:
\[
\bar{A} - \bar{B} = \max \left( 0, [p_A^L - p_B^L, p_A^U - p_B^U], [q_A^L - q_B^L, q_A^U - q_B^U], [r_A^L - r_B^L, r_A^U - r_B^U], e^{j[\omega_A^L(x)-\omega_B^L(x),\omega_A^U(x)-\omega_B^U(x)]} \right)
\]

(iv) The product of \( \bar{A} \) and \( \bar{B} \) is defined as:
\[
\bar{A} \otimes \bar{B} = \left( [p_A^L p_B^L, p_A^U p_B^U], e^{j[\omega_A^L(x)\omega_B^L(x),\omega_A^U(x)\omega_B^U(x)]}, [q_A^L q_B^L, q_A^U q_B^U], e^{j[\omega_A^L(x)\omega_B^U(x)+\omega_A^U(x)\omega_B^L(x)]}, [r_A^L r_B^L, r_A^U r_B^U], e^{j[\omega_A^L(x)\omega_B^U(x)\omega_A^U(x)\omega_B^L(x)]} \right)
\]

(v) The scalar multiplication of \( \bar{A} \) is defined as:
\[
k \otimes \bar{A} = \left( [1-(1-p_A^L)^k,1-(1-p_A^U)^k], e^{j[\omega_A^L(x)k,\omega_A^U(x)k]}, [q_A^L]^k, e^{j[k+\nu_A^L(x)k, k+\nu_A^U(x)k]}, [r_A^L]^k, e^{j[k+\omega_A^L(x)k, k+\omega_A^U(x)k]} \right)
\]

3.4. The difference between two interval-valued complex neutrosophic sets

**Definition 8.** The difference between two interval-valued complex neutrosophic sets

Let \( \bar{A} = \left( [T_A^L, T_A^U], [I_A^L, I_A^U], [F_A^L, F_A^U] \right) \) and \( \bar{B} = \left( [T_B^L, T_B^U], [I_B^L, I_B^U], [F_B^L, F_B^U] \right) \) be two interval neutrosophic sets over \( X \) which are defined by \( [T_A^L, T_A^U], [I_A^L, I_A^U], [F_A^L, F_A^U] \) and \( [T_B^L, T_B^U], [I_B^L, I_B^U], [F_B^L, F_B^U] \), respectively.

The difference between \( \bar{A} \) and \( \bar{B} \) is defined as follows:
\[
d(\bar{A}, \bar{B}) = \frac{1}{3n} \left( \max \left| T_A^L - T_B^L \right|, \max \left| T_A^U - T_B^U \right| + \max \left| I_A^L - I_B^L \right|, \max \left| I_A^U - I_B^U \right| + \max \left| F_A^L - F_B^L \right|, \max \left| F_A^U - F_B^U \right| \right)
\]
130

\[
\begin{align*}
&= \max \left( \max \left( p_T^U - p_T^L \cdot e^{i \pi \max \left( \frac{1}{2} \mu_T^U - \frac{1}{2} \mu_T^L \right)} \right), \right. \\
&\left. \quad \max \left( q_T^U - q_T^L \cdot e^{i \pi \max \left( \frac{1}{2} \nu_T^U - \frac{1}{2} \nu_T^L \right)} \right) \right) + \\
&\quad \max \left( r_T^U - r_T^L \cdot e^{i \pi \max \left( \frac{1}{2} \omega_T^U - \frac{1}{2} \omega_T^L \right)} \right) \right)
\end{align*}
\]

3.5. Linguistic Variables

A linguistic variable is a variable whose values are words or sentences in a natural or artificial language. For instance, some matters are characterized by linguistic terms, such as unsatisfied, fair and satisfied. Each linguistic variable can be assigned one or more linguistic values, which are in turn connected to a numeric value through the mechanism of membership functions. The linguistic terms shown in the following were used to quantify each attribute to normalize the decision-making matrix. Then, the linguistic variables of each criterion were weighted for importance. Accordingly, this study classified the data resources into five scales to express un-quantified matters, as shown in Table 2.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Linguistic variables for weighting importance based on interval complex neutrosophic values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linguistic variables</td>
</tr>
<tr>
<td>Far Below Standards (FBS)</td>
<td>(0.1, 0.2), [0.1, 0.2], [0.7, 0.8]</td>
</tr>
<tr>
<td>Below Standards (BS)</td>
<td>(0.2, 0.3), [0.2, 0.3], [0.5, 0.6]</td>
</tr>
<tr>
<td>Meets Standards (MS)</td>
<td>(0.3, 0.4), [0.3, 0.4], [0.3, 0.4]</td>
</tr>
<tr>
<td>Above Standards (AS)</td>
<td>(0.5, 0.6), [0.2, 0.3], [0.2, 0.3]</td>
</tr>
<tr>
<td>Far Above Standards (FAS)</td>
<td>(0.7, 0.8), [0.1, 0.2], [0.1, 0.2]</td>
</tr>
</tbody>
</table>

4. A novel hierarchical MCDM-model-based ICNS

4.1. Brief description of the hierarchical MCDM model

The TOPSIS method does not consider a hierarchical structure for the main criteria and sub-criteria. Thus, the relationship between overall weight and global weight lacks information about (i) the comparative analysis of different criteria and (ii) the approach separates qualitative and quantitative variables (Wang & Chan, 2013). These reasons make the TOPSIS technique more compatible with decision-making problems than others (Bottani & Rizzi, 2006). Especially, the accuracy of this framework depends on three issues: organizing the model’s structure; choosing and selecting the criteria set; and weighting each criterion for calculation in the TOPSIS technique. This paper proposes a hierarchical MCDM approach to present the framework for evaluating lecturer capacity. First, the AHP is used to determine the weights of criteria through pair-wise comparisons. Then, the TOPSIS method is used to acquire the comparative ratings of alternatives for lecturer performance. This evaluation of lecturer performance will lead to better results for two reasons: (i) the AHP technique can describe the correlations between criteria in the model; (ii) the aim of the TOPSIS technique is to convert multiple choices to a single choice.

Many researchers have highlighted the drawback of TOPSIS regarding the weight allocated to each response (Bottani & Rizzi, 2006; Wang & Chan, 2013). These weights all must sum up to one and may vary from person to person. So, these weights were determined using the AHP method for each response in order to consider which information is superior for solving complex decision problems (Saaty 1980). Any complex problem can be decomposed into several sub-problems with hierarchical levels using the AHP technique, where each level represents a set of criteria or attributes relative to other problems. This hierarchy of criteria and sub-criteria could be either quantitative or qualitative in nature. This can be done by introducing pair-wise comparisons between criteria, as assessed by professionals or experts in the corresponding areas. Implementing this lecturer evaluation could help determine the benefits and costs of educational activities. This study proposes a hierarchical TOPSIS
approach in the interval-valued complex neutrosophic environment. This approach was extended by applying interval linguistic variables and complex numbers.

4.2. Analysis of criteria weights with the AHP technique

In this methodology, the problem is structured in a hierarchy of different levels consisting of the main goal, criteria, sub-criteria, and alternatives. When assessing the improvement areas for implementing lecturer evaluation, it is essential to know the relative importance of the criteria and sub-criteria. In other words, assessors must determine the weights of the main criteria. The AHP is a capable technique for comparing the short- and long-term impacts of the gauge year. Additionally, the subjective assessments were converted to numerical values and processed to rank each alternative on the scale. The AHP method solved the problem with four necessary steps (Saaty, 1980):

**Step 1**: Define the problem, determine the goal of analysis and build the hierarchical structure model to evaluate the quality of the lecturer. First, the patterns needed to be defined. The criteria, sub-criteria, and alternatives were also determined. Second, all information was put in the hierarchical structures of the AHP technique. These illustrated the range of the problem from general to more detailed (Fabjanowicz et al., 2018). Note that the quality of performance affects the correctness of results, especially the consistency between the pair-wise comparisons of elements.

**Step 2**: The construct pair-wise comparison. This study involved collecting data from decision-makers and consultations with experts to compare alternatives. The relative importance of different elements was determined using the standard scoring values given in Table 3. The pair-wise comparison matrices for all factors were composed using expert opinions.

<table>
<thead>
<tr>
<th>Intensity of importance</th>
<th>Definition</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Equal importance</td>
<td>Two activities contribute equally to the objective.</td>
</tr>
<tr>
<td>5</td>
<td>Essential or strong importance</td>
<td>Experience and judgment strongly favor one activity over another.</td>
</tr>
<tr>
<td>7</td>
<td>Very strong importance</td>
<td>An activity is favored very strongly over another, its dominance.</td>
</tr>
<tr>
<td>9</td>
<td>Demonstrated importance</td>
<td>The evidence favoring one activity over another is the highest possible order of affirmation.</td>
</tr>
<tr>
<td>2, 4, 6, 8</td>
<td>Intermediate values between the two adjacent judgment</td>
<td>When compromise is needed</td>
</tr>
</tbody>
</table>

This scale aims to determine how many times a more important or dominant element is prioritized over another element concerning the criterion or property with which that element is associated in non-trivial comparisons, according to the formula presented below (Saaty, 2008).

\[ w_{ij} = \frac{a_{ij}}{\sum_{j=1}^{n} a_{ni}} \]  \hspace{1cm} (7)

Then the crisp pair-wise comparisons were constructed such that:

\[ A = (\tilde{a}_{ij})_{n \times n} = \begin{bmatrix} 1 & a_{12} & \cdots & n \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix} \]

**Step 3**: Check the Consistency Ratio (CR). If matrix A is a consistent matrix, then the maximum eigenvalue of A should equal its number of orders. However, in practice, the pair-wise comparison matrix cannot achieve complete consistency. This shows the consistency of elements when the
decision-makers make pair-wise comparisons. The execution of this step in the algorithm confirmed that each matrix was within a permissive CR. Note that the result of multiplying the values in diagonal elements is 1 as shown in Table 4.

### Table 4
The formula of the Consistency Ratio (CR) in the AHP method

<table>
<thead>
<tr>
<th>Consistency Ratio (CR)</th>
<th>( \lambda_{\text{max}} )</th>
<th>Consistency Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>( CR = \frac{CI}{RI} )</td>
<td>( \lambda_{\text{max}} = \frac{\sum y}{n} )</td>
<td>( CI = \frac{\lambda_{\text{max}} - n}{n-1} )</td>
</tr>
</tbody>
</table>

In which:  
CI: The matrix’s consistency index  
RI: The average consistency of random matrices  
\( \lambda_{\text{max}} \): The average of the entries in consistency vector \( y \) and given by the formula.  
n: The number of criteria or systems

If the consistency ratio \( \leq 0.1 \), the evaluation within the matrix is acceptable. In contrast, if CR is more than 0.1, the judgments are untrustworthy, because they are too close to randomness and the assessment is valueless or must be repeated. The average consistency of random matrices (RI) as shown in Table 5.

### Table 5
The average consistency of Random Matrices (RI) (Saaty 1980)

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>RI</td>
<td>0</td>
<td>0</td>
<td>0.58</td>
<td>0.9</td>
<td>1.12</td>
<td>1.24</td>
<td>1.32</td>
<td>1.41</td>
<td>1.45</td>
<td>1.49</td>
</tr>
</tbody>
</table>

4.3. Hierarchical TOPSIS model based on ICNS

The classical TOPSIS approach was identified as the best alternative based on the estimation of Euclidean distance. This solution describes the relationship between the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution. The ideal classical TOPSIS method can be presented using the following steps:

The advantages of TOPSIS are its logicality, rationality and computational simplicity. The classical TOPSIS method and its extensions have demonstrated their capabilities and potential in dealing with MCDM problems in various fields. These extensions involve integrating the interval complex neutrosophic set and TOPSIS.

**Step 1: Construct the decision-making matrix**

The process starts with the construction of a decision matrix, \( D \), based on a given set of criteria and sub-criteria. The data of decision matrix \( D \) come from different sources. \( D = [d_{ij}]_{m \times n} \) is the decision matrix, where \( d = ([T^L_{ij}, T^U_{ij}], [I^L_{ij}, I^U_{ij}], [F^L_{ij}, F^U_{ij}]) \) and takes the form of the INVs for alternative \( A_i \) with respect to attribute \( C_j \) in which \( [T^L_{\pi}, T^U_{\pi}] = [P^L_{\pi}, P^U_{\pi}], e^{j\pi[\alpha_{\pi}(x), \beta_{\pi}(x)]}, [I^L_{\pi}, I^U_{\pi}] = [q^L_{\pi}, q^U_{\pi}], e^{j\pi[\nu_{\pi}(x), \xi_{\pi}(x)]}, [F^L_{\pi}, F^U_{\pi}] = [r^L_{\pi}, r^U_{\pi}] \) and \( e^{j\pi[\omega_{\pi}(x), \lambda_{\pi}(x)]} \), respectively.

For a multiple attribute decision problem, let \( A = (A_1, A_2, ..., A_n) \) be a discrete set of alternatives, \( C = (C_1, C_2, ..., C_n) \) be the set of attributes, \( W = (w_1, w_2, ..., w_n)^T \) be the weighting vector of the attributes and \( \sum_{j=1}^{n} w_j = 1, w_j \geq 0 \), \( w_j \) be unknown. In general, the criteria can be classified into two categories: benefit and cost.
Step 2: Normalize the decision matrix (Chi & Liu 2013)

In general, the criteria can be classified into two categories: benefit and cost. The benefit criterion means that a higher value is better, while for the cost criterion the opposite holds true. Therefore, it is necessary to normalize them in order to transform them into a dimensionless matrix that allows comparing the various criteria. In this research, the normalized decision matrix is denoted by $\mathbf{R}$, as follows:

$$\mathbf{R} = \left[ r_{ij} \right]_{m \times n}$$

is the decision matrix, where $r_{ij} = \left( \left[ T_{ij}^{L}, T_{ij}^{U} \right], \left[ L_{ij}^{L}, L_{ij}^{U} \right], \left[ F_{ij}^{L}, F_{ij}^{U} \right] \right)$.

$$\mathbf{R} = \begin{cases} \frac{d_{ij}}{} & \text{if criterion C}_j \text{ is benefit type} \\ \frac{1}{d_{ij}} & \text{if criterion C}_j \text{ is cost type} \end{cases}$$

where $x_{\text{max}}$ and $x_{\text{min}}$ respectively present the largest and lowest values of each sub-criterion.

Step 3: Determine the positive ideal solution (PIS) and negative ideal solution (NIS)

The study involved selecting the virtual positive ideal solution and negative ideal solution by selecting the best values for each attribute from all alternatives.

$$\begin{align*}
S_j^* &= \left[ \left[ \text{max } T_{ij}^{L*}, \text{max } T_{ij}^{U*} \right], \left[ \text{min } I_{ij}^{L*}, \text{min } I_{ij}^{U*} \right], \left[ \text{min } F_{ij}^{L*}, \text{min } F_{ij}^{U*} \right] \right] \\
S_j^- &= \left[ \left[ \text{min } T_{ij}^{L*}, \text{min } T_{ij}^{U*} \right], \left[ \text{max } I_{ij}^{L*}, \text{max } I_{ij}^{U*} \right], \left[ \text{max } F_{ij}^{L*}, \text{max } F_{ij}^{U*} \right] \right]
\end{align*}$$

Step 4: Calculate the separation measures from the positive ideal solution (PIS) and negative ideal solution (NIS)

The distance between the alternative $A_i$ and the positive ideal solution/negative ideal solution is:

$$\begin{align*}
D_i^+ &= \sum_{j=1}^{n} \frac{d_{ij}}{d_{ij}^+} \\
D_i^- &= \sum_{j=1}^{n} \frac{d_{ij}}{d_{ij}^-}
\end{align*}$$

Step 5: Select the best alternative

Rule 1: When $d_i^+$ is smaller, $A_i$ is nearer to the best $A^*$, i.e., $A_i$ is the best solution. Rule 2: When $d_i^-$ is bigger, $A_i$ is farther from the worst $A^*$, i.e., $A_i$ is the best solution.

Let $\mathbf{A} = \left( \left[ T_{A}^{L}, T_{A}^{U} \right], \left[ I_{A}^{L}, I_{A}^{U} \right], \left[ F_{A}^{L}, F_{A}^{U} \right] \right)$ and $\mathbf{B} = \left( \left[ T_{B}^{L}, T_{B}^{U} \right], \left[ I_{B}^{L}, I_{B}^{U} \right], \left[ F_{B}^{L}, F_{B}^{U} \right] \right)$ be two interval neutrosophic sets over $X$ which are defined by $\left[ T_{A}^{L}, T_{A}^{U} \right] = \left[ P_{A}^{L}, P_{A}^{U} \right] e^{i \left[ \mu_{A}(x) \phi_{A}(x) \right]}$, $\left[ I_{A}^{L}, I_{A}^{U} \right] = \left[ q_{A}^{L}, q_{A}^{U} \right] e^{i \left[ \nu_{A}(x) \psi_{A}(x) \right]}$, $\left[ F_{A}^{L}, F_{A}^{U} \right] = \left[ r_{A}^{L}, r_{A}^{U} \right] e^{i \left[ \omega_{A}(x) \theta_{A}(x) \right]}$, respectively. Let $x_{\mathbf{A}}^L$ and $x_{\mathbf{A}}^U$ be the amplitude and phase terms of interval complex number $\mathbf{A}$. The comparison between $\mathbf{A}$ and $\mathbf{B}$ is defined as follows:

$$\mathbf{A} > \mathbf{B} \iff \begin{cases} x_{\mathbf{A}}^L > x_{\mathbf{B}}^L \\ x_{\mathbf{A}}^U = x_{\mathbf{B}}^U \\ x_{\mathbf{A}}^L > x_{\mathbf{B}}^U \end{cases}$$

4.4. Steps in the novel hierarchical TOPSIS model

From the proposed algorithm of the AHP technique and the integration of the interval complex neutrosophic set and TOPSIS, it was possible to compute the basic steps of the novel hierarchical TOPSIS models as follows:
**Step 1**: Determined the criteria, sub-criteria and alternatives/objects of the decision-making problems and establish a hierarchical structure.

**Step 2**: Obtained the degree of importance and the performance of the alternatives/objects for all criteria using expert opinions.

**Step 3**: After collecting information from experts, the pair-wise comparison matrix was constructed based on the standard values of the relative importance of factors.

**Step 4**: Used the Lambda-max method to calculate the weight of each criterion given by the experts. The consistency index of a comparison matrix must be below 0.1. If it exceeded this value, the decision maker checked the pair-wise comparison matrix in Step 3.

**Step 5**: Applied the AHP method to integrate all expert opinions to obtain a weight for each aggregative criterion.

**Step 6**: Constructed the decision matrix from the input data that were collected in Step 1 based on the interval-valued linguistics and complex neutrosophic set. The matrix was built by putting \( D^{(k)} = (d^{(k)}_{ij})_{m \times n} \) \((k = 1,2,\ldots,l)\) into a collective interval complex neutrosophic decision matrix \( D = (d_{ij})_{m \times n} \).

**Step 7**: Established a normalized performance matrix where two types of criteria (benefit and cost criteria) were converted.

**Step 8**: Calculated the weighted normalized matrix.

**Step 9**: Determined the maximum and minimum of the differences and calculated the positive and negative ideal solutions (PIS and NIS).

**Step 10**: Used the separation measures of each point to PIS and NIS and compared them to select the best alternatives.

Fig. 2 shows the methodology framework for lecturer evaluation using a hierarchical TOPSIS model and Interval-valued Complex Neutrosophic Set.
5. Application of the proposed hierarchical multi-criteria model for lecturer evaluation

This section presents a numerical example to demonstrate the proposed method. A university needs to choose the best lecturer from five possible options, denoted as $A_1$, $A_2$, $A_3$, $A_4$, and $A_5$. There are 13 sub-criteria in four groups in the database; these represent the average values of each surveyed group. The four criteria are: $C_1$, Self-evaluation; $C_2$, Manager-based evaluation; $C_3$, Peer-evaluation; $C_4$, Student-based evaluation. Sub-criteria $C_{23}$ and $C_{43}$ are cost criteria and the others are benefit criteria. In order to select the best lecturer, the hierarchical TOPSIS in the interval-valued complex neutrosophic set is applied as follows.

**Steps 1-5:** Calculate the weight of criteria and sub-criteria

Tables 6 and 7 show the weight of the sub-criteria and criteria based on the AHP methodology using formulas (8-10).

**Table 6**
The values of pair-wise comparison matrices and weights of criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$C_4$</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>1</td>
<td>1/4</td>
<td>1/2</td>
<td>1/5</td>
<td>0.084</td>
</tr>
<tr>
<td>$C_2$</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0.454</td>
</tr>
<tr>
<td>$C_3$</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
<td>1/2</td>
<td>0.154</td>
</tr>
<tr>
<td>$C_4$</td>
<td>5</td>
<td>1/2</td>
<td>2</td>
<td>1</td>
<td>0.309</td>
</tr>
</tbody>
</table>

$\lambda_{max} = 4.090, CI = 0.030, RI = 0.900, CR = 0.003 < 0.10$ (Consistency OK)

**Table 7**
The values of pair-wise comparison matrices and weights of sub-criteria

<table>
<thead>
<tr>
<th>Sub-criteria</th>
<th>Pair-wise comparison matrix</th>
<th>Local Weight</th>
<th>Overall Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{11}$</td>
<td>1 1/3 2</td>
<td>0.230</td>
<td>0.018</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>3 1 5</td>
<td>0.640</td>
<td>0.053</td>
</tr>
<tr>
<td>$C_{13}$</td>
<td>1/2 1/5 1</td>
<td>0.130</td>
<td>0.011</td>
</tr>
</tbody>
</table>

$W = 0.083; \lambda_{max} = 3.005, CI = 0.003, RI = 0.900, CR = 0.005 < 0.10$

<table>
<thead>
<tr>
<th>Sub-criteria</th>
<th>Pair-wise comparison matrix</th>
<th>Local Weight</th>
<th>Overall Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{21}$</td>
<td>1 1/3 2</td>
<td>0.471</td>
<td>0.214</td>
</tr>
<tr>
<td>$C_{22}$</td>
<td>1/2 1/2 1/2</td>
<td>0.284</td>
<td>0.129</td>
</tr>
<tr>
<td>$C_{23}$</td>
<td>1/3 1/2 1</td>
<td>0.171</td>
<td>0.078</td>
</tr>
</tbody>
</table>

$W = 0.471; \lambda_{max} = 4.066, CI = 0.0219, RI = 0.900, CR = 0.0240 < 0.10$

<table>
<thead>
<tr>
<th>Sub-criteria</th>
<th>Pair-wise comparison matrix</th>
<th>Local Weight</th>
<th>Overall Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{31}$</td>
<td>1 1/2 4</td>
<td>0.324</td>
<td>0.050</td>
</tr>
<tr>
<td>$C_{32}$</td>
<td>1 2 6</td>
<td>0.587</td>
<td>0.090</td>
</tr>
<tr>
<td>$C_{33}$</td>
<td>1/4 1/6 1</td>
<td>0.089</td>
<td>0.014</td>
</tr>
</tbody>
</table>

$W = 0.154; \lambda_{max} = 3.012, CI = 0.006, RI = 0.580, CR = 0.011 < 0.10$

<table>
<thead>
<tr>
<th>Sub-criteria</th>
<th>Pair-wise comparison matrix</th>
<th>Local Weight</th>
<th>Overall Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{41}$</td>
<td>1 1/3 2</td>
<td>0.623</td>
<td>0.193</td>
</tr>
<tr>
<td>$C_{42}$</td>
<td>1/3 2 1</td>
<td>0.239</td>
<td>0.074</td>
</tr>
<tr>
<td>$C_{43}$</td>
<td>1/4 1/2 1</td>
<td>0.137</td>
<td>0.042</td>
</tr>
</tbody>
</table>

$W = 0.309; \lambda_{max} = 3.025, CI = 0.013, RI = 0.580, CR = 0.022 < 0.10$

**Step 6:** Construct the decision matrix from the collected input data, based on the interval-valued linguistics and complex neutrosophic set as seen in Table 2. The hierarchical linguistic variables for the importance weight of each criterion was presented and the ratings of alternatives under all criteria by decision-makers in Table 8 and Table 9.

**Table 8**
Hierarchical linguistic variables for the importance weights of criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{11}$</td>
<td>MS</td>
<td>AS</td>
<td>BS</td>
<td>AS</td>
<td>MS</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>MS</td>
<td>BS</td>
<td>MS</td>
<td>AS</td>
<td>AS</td>
</tr>
<tr>
<td>$C_{13}$</td>
<td>MS</td>
<td>AS</td>
<td>BS</td>
<td>AS</td>
<td>BS</td>
</tr>
<tr>
<td>$C_{14}$</td>
<td>FAS</td>
<td>AS</td>
<td>FAS</td>
<td>MS</td>
<td>AS</td>
</tr>
<tr>
<td>$C_{15}$</td>
<td>BS</td>
<td>MS</td>
<td>AS</td>
<td>MS</td>
<td>MS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{21}$</td>
<td>MS</td>
<td>AS</td>
<td>AS</td>
<td>AS</td>
<td>MS</td>
</tr>
<tr>
<td>$C_{22}$</td>
<td>MS</td>
<td>FAS</td>
<td>MS</td>
<td>MS</td>
<td>FAS</td>
</tr>
<tr>
<td>$C_{23}$</td>
<td>FBS</td>
<td>MS</td>
<td>MS</td>
<td>MS</td>
<td>MS</td>
</tr>
<tr>
<td>$C_{24}$</td>
<td>BS</td>
<td>FAS</td>
<td>FAS</td>
<td>FAS</td>
<td>MS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{31}$</td>
<td>AS</td>
<td>AS</td>
<td>FAS</td>
<td>FAS</td>
<td>FAS</td>
</tr>
<tr>
<td>$C_{32}$</td>
<td>AS</td>
<td>FAS</td>
<td>FAS</td>
<td>FAS</td>
<td>FAS</td>
</tr>
<tr>
<td>$C_{33}$</td>
<td>AS</td>
<td>AS</td>
<td>FAS</td>
<td>FAS</td>
<td>FAS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{41}$</td>
<td>AS</td>
<td>AS</td>
<td>FAS</td>
<td>FAS</td>
<td>FAS</td>
</tr>
</tbody>
</table>
**Table 9**
The ratings of alternatives under all criteria by decision-makers

<table>
<thead>
<tr>
<th>$A_i$</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
<th>$A_6$</th>
<th>$A_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{11}$</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.50,0.60</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
<tr>
<td>$C_{13}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
<tr>
<td>$C_{14}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
<tr>
<td>$C_{15}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
</tbody>
</table>

**Step 7:** Establish a normalized performance matrix in which two or three cost criteria are converted using formula (11), as shown in Table 10.

**Table 10**
The normalized matrix after converting the benefit and cost criteria

<table>
<thead>
<tr>
<th>$A_i$</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
<th>$A_6$</th>
<th>$A_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{11}$</td>
<td>0.50,0.60</td>
<td>0.60,0.70</td>
<td>0.70,0.80</td>
<td>0.80,0.90</td>
<td>0.50,0.60</td>
<td>0.60,0.70</td>
<td>0.70,0.80</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
<tr>
<td>$C_{13}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
<tr>
<td>$C_{14}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
<tr>
<td>$C_{15}$</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
<td>0.50,0.60</td>
<td>0.20,0.30</td>
<td>0.30,0.40</td>
<td>0.40,0.50</td>
</tr>
</tbody>
</table>

Note: The ratings in Tables 9 and 10 are hypothetical examples for demonstration purposes.
Step 8: Calculate the weighted normalized matrix by multiplying the weight of each criterion and the corresponding values from the decision-making in step 7 (shown in Table 11).

<table>
<thead>
<tr>
<th>$A_i$</th>
<th>$A_j$</th>
<th>$A_k$</th>
<th>$A_l$</th>
<th>$A_m$</th>
<th>$A_n$</th>
<th>$A_o$</th>
<th>$A_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase term</td>
<td>Amplitude term</td>
<td>Phase term</td>
<td>Amplitude term</td>
<td>Phase term</td>
<td>Amplitude term</td>
<td>Phase term</td>
<td>Amplitude term</td>
</tr>
<tr>
<td>$c_1$</td>
<td>(0.06, 0.09), (0.05, 0.07), (0.05, 0.06), (0.10, 0.12)</td>
<td>(0.04, 0.05), (0.04, 0.05), (0.04, 0.05), (0.04, 0.05)</td>
<td>(0.10, 0.12), (0.10, 0.12), (0.10, 0.12), (0.10, 0.12)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.06, 0.09), (0.06, 0.09), (0.06, 0.09), (0.06, 0.09)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.98, 0.99), (0.98, 0.99), (0.98, 0.99), (0.98, 0.99)</td>
</tr>
<tr>
<td>$c_2$</td>
<td>(0.04, 0.05), (0.04, 0.05), (0.04, 0.05), (0.04, 0.05)</td>
<td>(0.10, 0.12), (0.10, 0.12), (0.10, 0.12), (0.10, 0.12)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.06, 0.09), (0.06, 0.09), (0.06, 0.09), (0.06, 0.09)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.99, 0.99), (0.99, 0.99), (0.99, 0.99), (0.99, 0.99)</td>
</tr>
<tr>
<td>$c_3$</td>
<td>(0.10, 0.12), (0.10, 0.12), (0.10, 0.12), (0.10, 0.12)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.06, 0.09), (0.06, 0.09), (0.06, 0.09), (0.06, 0.09)</td>
<td>(0.06, 0.09), (0.06, 0.09), (0.06, 0.09), (0.06, 0.09)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.05, 0.07), (0.05, 0.07), (0.05, 0.07), (0.05, 0.07)</td>
<td>(0.99, 0.99), (0.99, 0.99), (0.99, 0.99), (0.99, 0.99)</td>
</tr>
</tbody>
</table>

Step 9: Determine the maximum and minimum of the differences and calculate the positive and negative ideal solutions (PIS and NIS) using formula (12) as shown in Table 12.

<table>
<thead>
<tr>
<th>$S^+$</th>
<th>$S^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase term</td>
<td>Amplitude term</td>
</tr>
<tr>
<td>$c_1$</td>
<td>(0.02, 0.04), (0.02, 0.03), (0.02, 0.03), (0.02, 0.03)</td>
</tr>
<tr>
<td>$c_2$</td>
<td>(0.00, 0.00), (0.00, 0.00), (0.00, 0.00), (0.00, 0.00)</td>
</tr>
</tbody>
</table>

Step 10: Use the separation measures of each point to PIS and NIS and compare them to select the best alternatives using formula (13) that was shown in Table 13.
Step 11: Rank the alternatives: $A_1 > A_2 > A_3 > A_4$.

A total of 13 evaluation dimensions exist for the four options in this study. Tables 6 and 7 illustrate that a consistency ratio of 0.10 or less is acceptable. If the consistency ratio is greater than 0.10, the weight assignments must be re-evaluated within the pair-wise comparisons matrix. It is clear that the highest number of results is for the manager-based evaluation ($C_2$) (0.454), while self-evaluation has the lowest ($C_1$) (0.084). Therefore, the manager aspects (lecturer activities and lecturer style) and the content of the lessons are the major factors for the results of evaluation, with 0.214, 0.129 and 0.193, respectively. Managers and students are the direct and objective assessors of the quality of teaching and the capacity of the lecturer. Hence, these are the main subjects who are affected by the results of this assessment.

After identifying the relative weights of the evaluation criteria, five lecturers of the University of Economics and Business, Vietnam were ranked. We first summarize all the input data and then standardized the different units of each evaluation criterion based on a linguistic variable. Next, we utilized the hierarchical TOPSIS method in the interval-valued complex neutrosophic environment by using the relative weights acquired from the AHP in the previous Section to calculate the weighted values of each lecturer. The results derived for this approach show that $A_1$’s teaching activities have the highest level and is recommended to enhance capacity. This is due to the fact that there is a focus on investment for teaching and scientific activities (at FAS and MS, respectively). Additionally, this lecturer has good lecturer-student interactions and improves the quality of the subjects. The medium grades belong to $A_2$, $A_3$, and $A_5$, who have the low values of input data for most criteria. However, from the results indicated in Table 10, we discovered that all five criteria in $A_4$ achieve the meets standards (MS) grade, which indicates that this alternative does not focus on cooperation with co-workers and behavior. Thus, based on the results of the evaluation, this subject was rated with the worst grade. Therefore, the proposed approach is useful for lecturer evaluation and can improve effectiveness and sustained competitive advantage. The implementation of this assessment will not only improve the lecturers’ performance but also enhance the University’s brand image.

6. Discussion and Conclusions

6.1 Comparing with other models

In this Section, the proposed method is compared with other methods. Table 14 lists the results of the comparison. The proposed method and the classic TOPSIS method can solve problems in uncertain environments. However, the TOPSIS and AHP techniques have some disadvantages in terms of calculation methods and results. Besides, the extent of the interval neutrosophic TOPSIS does not consider the capacity of each lecturer in the specific time period. The comparisons between the previous methods and the proposed hierarchical TOPSIS method are summarized in Table 14.

Table 14

<table>
<thead>
<tr>
<th>Method</th>
<th>Object</th>
<th>Aggregative value</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td></td>
<td>$A_1$ 5.288</td>
<td>$A_2$ 5.075</td>
</tr>
<tr>
<td>Interval neutrosophic TOPSIS</td>
<td></td>
<td>$D^+$ 5.075</td>
<td>$D^-$ 5.288</td>
</tr>
<tr>
<td>AHP (Saaty, 1980)</td>
<td></td>
<td>$A_1$ 5.075</td>
<td>$A_2$ 5.075</td>
</tr>
<tr>
<td>TOPSIS (Hwang and Yoon, 1981)</td>
<td></td>
<td>$D^+$ 0.098</td>
<td>$D^-$ 0.051</td>
</tr>
</tbody>
</table>

Note: For comparison, in the TOPSIS technique, the criteria weights are those of AHP.
Table 14 shows the comparison between the proposed method and three different methods for the aggregative value and the ranking of total alternatives. The table shows that the TOPSIS and AHP methods have the same results in classifying the abilities of the lecturers. Accordingly, $A_3$ is the best lecturer ($n=7.460$ for the AHP results), while the worst lecturer is $A_1$ ($n=5.164$ for the AHP results). Additionally, in terms of the results for the interval neutrosophic TOPSIS, there is a difference with that of the below techniques. $A_2$ is evaluated with high values as the best alternative, and $A_3$ and $A_5$ follow. Moreover, the lowest level for lecturer evaluation is for $A_4$, instead of $A_1$ in the results of the TOPSIS and AHP methods. In contrast, $A_1$ has the highest rating in the integrated hierarchical TOPSIS and the interval-valued complex neutrosophic set. Furthermore, similar to the interval neutrosophic TOPSIS, $A_4$ has a medium rating (at the final class). When the study used different methods, the ranking of the alternatives was disturbed. This demonstrates that the number of criteria and the sample size play important roles in decision-making problems. Thus, depending on the complexity of the issues, an assessor may require customizable selections and a suitable method.

6.2 Limitations of this model

The current study has several limitations. First, lecturer evaluation is a difficult issue in employee recruitment in the educational system (Malen et al. 2015). Any assessment must be effective because it includes the possibility of job loss (Kunter and Baumert 2006; Odden 2014). For some lecturers, a low rating may motivate them to enhance their professional level. Others may try to build nice relationships with management, fellow workers and students. Hence, determining the weaknesses of lecturers and universities is key to supporting improvement, but it also brings negative impacts such as competition between colleagues (Fischer et al. 2018). Additionally, this approach should be tested for a long time before being applied (Derrington and Campbell 2015). Secondly, although the TOPSIS method in interval-valued complex neutrosophic is effective, this method requires complex calculations to ensure exact results (Chi and Liu 2013). Thus, the method depends on the professional knowledge of experts to identify the importance of each criterion (Malik et al., 2016). Third, limitations in the number of samples may undermine the robustness of the findings, such as low reliability of the cognitive measure ($n = 05$). It would be beneficial for further research to use larger samples from different educational levels. Moreover, the study findings rely on four objective reports to create a set of criteria. Further research can evaluate many methods for selecting the data, such as interviews and survey reports. The survey method could be used to obtain the information and data required by an investigator (Wu et al. 2012). Finally, this study is limited by the fact that it only reflects some subjects’ responses, which may or may not share the perspectives of the evaluated lecturers. Still, school leaders are objective observers of lecturer-level effects, perhaps more so than the lecturers themselves (Taut et al. 2011). For this reason, we recommend further work on the coverage and depth of the problem.

6.3 Conclusions

The implementation of lecturer evaluation can generate competition between lecturers in a university. It also requires a strategic approach to fairness and transparency. Because of its complexity, assessment has become a hot topic in education and management systems. This study has presented a comprehensive assessment based on a hierarchical structure for the criteria set. The hierarchical TOPSIS approach was developed with the interval-valued complex neutrosophic set to create an assessment for determining lecturer capability. The proposed hierarchical approach was further compared with the related methods to demonstrate the advantages and applicability. The results show that the proposed approach is efficient and can be used to solve other decision-making problems. However, there remains certain limitation of the study and future work are proposed to make lecture evaluation more accurate, which will support the dynamic decision-making procedure in education contexts in the real world.

Acknowledgements

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 503.01-2018.03.
References


