

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

Anh Duc Do^{a,e*}, Minh Tam Pham^b, Thi Hang Dinh^c, The Chi Ngo^d, Quoc Dat Luu^e, Ngoc Thach Pham^f, Dieu Linh Ha^g and Hong Nhat Vuong^h

^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

Article history:

Received November 28, 2019

Received in revised format:

December 28, 2019

Accepted January 24, 2020

Available online

February 1, 2020

Keywords:

AHP

Lecturer evaluation

Interval complex neutrosophic

TOPSIS

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.

© 2020 by the authors; licensee Growing Science, Canada.

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

* Corresponding author. Tel: +84.813686666

E-mail address: anhduc1510@gmail.com (A. D. Do)

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 absences 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 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 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.

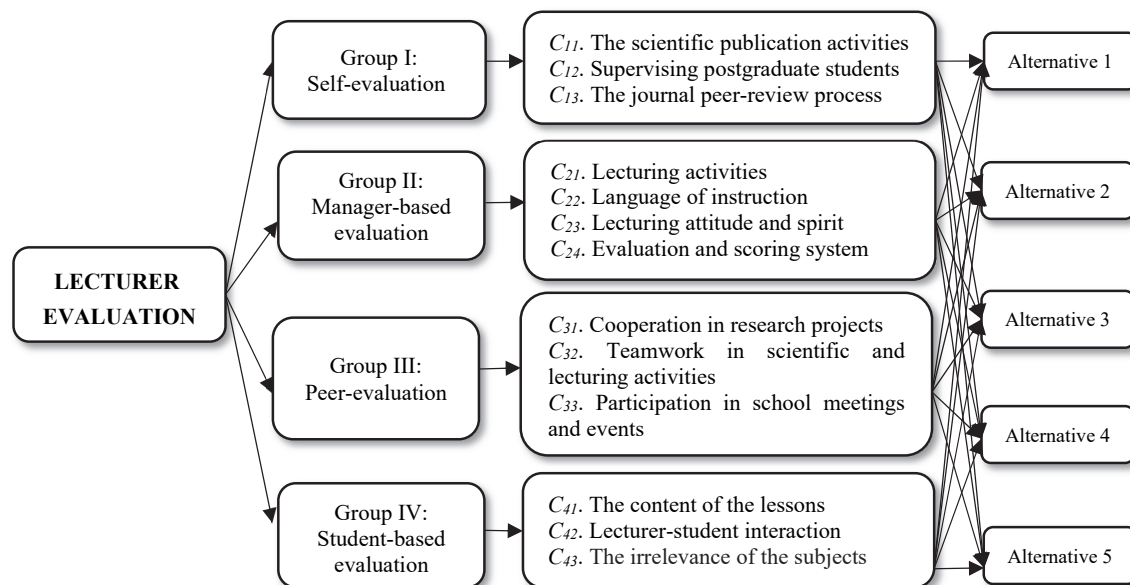


Fig. 1. The criteria used in lecturer evaluation

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

	Description/ Unit	TRUTH (T)	INDETERMINAC Y (I)	FALSITY (F)
1. Self-evaluation (C₁)				
1.1. The scientific publication activities (C₁₁) (+) (Wu et al. 2012; Zare et al. 2016; Jaramillo et al. 2017; Zerem 2017)				
<i>Real part:</i> The number of articles at international standard level (h ₁₁)/ Total number of articles expected published per year (t ₁₁)	The ratio is between the number of articles over 02 articles per year and the total in a year.	h ₁₁ which completed/ t ₁₁	h ₁₁ which are submitting or processing/ t ₁₁	h ₁₁ which did not complete or rejected/ t ₁₁
<i>Imaginary part:</i> The duration to carry out researches (months) (h' ₁₁)/12 months	The duration to carry out researches (months)/12 months.	h' ₁₁ are under 4 months/12 months	h' ₁₁ are from 04-06 months/12 months	h' ₁₁ are over 12 months/12 months
1.2. Supervising postgraduate students (C₁₂) (+) (Wu et al. 2012; Wisker 2012; Sharp et al. 2017)				
<i>Real part:</i> The number trainees who were guided (h ₁₂)/ Total trainees per year (t ₁₂)	The ratio is between the number of trainees over 05 trainees per year and the total in a year.	h ₁₂ completed/ t ₁₂	h ₁₂ uncompleted or processing/ t ₁₂	h ₁₂ could not complete or rejected/ t ₁₂
<i>Imaginary part:</i> The number of standard graduation reports (h' ₁₂)/ total graduation reports (t' ₁₂)	The ratio is between the number of standard graduation reports over 05 reports per year and the total in a year.	h' ₁₂ published to articles /t' ₁₂	h' ₁₂ can publish lately to articles /t' ₁₂	h' ₁₂ did not publish to articles/t' ₁₂
1.3. The journal peer-review process (C₁₃) (+) (Wu et al. 2012; Zare et al. 2016)				
<i>Real part:</i> The number of journal publications reviewed (h ₁₃)/ Total number of journal publications were suggested per year (t ₁₃)	The ratio is between the number of journal publications over 02 publications per year and the total in a year.	h ₁₃ which completed/ t ₁₃	h ₁₃ which are submitting or processing/ t ₁₃	h ₁₃ which did not complete or rejected/ t ₁₃
<i>Imaginary part:</i> The duration to become the journal reviewer (months) (h' ₁₃)/Total duration joined in the scientific publication process (t' ₁₃)	The duration to become the journal reviewer (months)/ Total duration joined in the scientific publication process.	h' ₁₃ are under 03 years/ t' ₁₃	h' ₁₃ are from 03-05 years/ t' ₁₃	h' ₁₃ are over 05 years/ t' ₁₃
2. Evaluation of Management (C₂)				
2.1. Lecturing activities (C₂₁) (+) (Wisker 2012; Jaramillo et al. 2017)				
<i>Real part:</i> The number of lessons per year (h ₂₁)/Total standard lessons (t ₂₁)	The ratio is between the number of lessons per year and the total in a year.	h ₂₁ had over 70 lessons per year/ t ₂₁	h ₂₁ from 50-70 lessons per year / t ₂₁	h ₂₁ had under 50 lessons per year / t ₂₁
<i>Imaginary part:</i> The number of subjects which lecturers were assigned (h' ₂₁)/ Total subjects which student registered (t' ₂₁)	The ratio is between the number of subjects which lecturers were assigned and total subjects which student registered in a year.	h' ₂₁ with only 1 class of registration /t' ₂₁	h' ₂₁ with 03-05 classes of registration /t' ₂₁	h' ₂₁ with over 05 classes of registration /t' ₂₁
2.2. Lecturing styles (C₂₂) (+) (Wu et al. 2012; Bradford 2015)				
<i>Real part:</i> The number of courses taught in English (h ₂₂)/ Total courses (t ₂₂)	The ratio is between the number of number of courses taught in English and total courses in a year.	h ₂₂ had over 50 peoples/t ₂₂	h ₂₂ had from 30-50 peoples/t ₂₂	h ₂₂ had under 30 peoples/t ₂₂
<i>Imaginary part:</i> The number of students who did not understand lessons (h' ₂₂)/ Total students (t' ₂₂)	The ratio between the number of who did not understand lessons and total students in a year.	h' ₂₂ failed had score at good and excellent level in the English program /t' ₂₂	h' ₂₂ had score at medium level in the English program /t' ₂₂	h' ₂₂ failed in the English program /t' ₂₂
2.3. Lecturing attitude and spirit (C₂₃) (-) (Frunzã 2014; Ziad 2016; Wang and Hsieh 2017)				
<i>Real part:</i> The number of lessons which lecturers came to class lately (h ₂₃)/Total lessons (t ₂₃)	The ratio is between the number of lessons which lecturers came to class lately and total in a year.	h ₂₃ accounts for under 30% of t ₂₃	h ₂₃ accounts for 30-50% of t ₂₃	h ₂₃ accounts for over 50% of t ₂₃
<i>Imaginary part:</i> The duration which lecturers came to class lately (h' ₂₃)/Total duration of lessons (t' ₂₃)	The ratio is between the duration which lecturers came to class lately and total duration in a year.	h' ₂₃ accounts for under 20% of t' ₂₃	h' ₂₃ accounts for 20-50% of t' ₂₃	h' ₂₃ accounts for over 50% of t' ₂₃
2.4. Score evaluation process for students (C₂₄) (+) (Bradford 2015; Tondeur et al. 2017)				
<i>Real part:</i> The number of exams which lectures organized (h ₂₄)/Total standard exams (t ₂₄)	The ratio is between the number of exams which lecturers organized and total in a year.	h ₂₄ accounts for over 80% t ₂₄	h ₂₄ accounts for 60-80% t ₂₄	h ₂₄ accounts for under 60% t ₂₄
<i>Imaginary part:</i> The average duration which lecturers paid each exam (h' ₂₄)	The difference between the average duration which lecturers paid each exam and the exam time before.	h' ₂₄ are under 01 month	h' ₂₄ are from 01-02 months	h' ₂₄ are over 02 months
3. Peer-evaluation (C₃)				
3.1. Cooperation in research projects (C₃₁) (+) (Wager and Kleinert 2012; Wu et al. 2012; Hein et al. 2015)				
<i>Real part:</i> The number of co-worker cooperation projects (h ₃₁)/ Total projects (t ₃₁)	The ratio is between the number of co-worker cooperation projects over 02 projects and the total in a year.	h ₃₁ at ministerial level/ t ₃₁	h ₃₁ at school level/t ₃₁	h ₃₁ did not belong to the above two categories / t ₃₁
<i>Imaginary part:</i> The average duration to carry out each project (months) (h' ₃₁)/12 months	The ratio between the average duration to carry out project and the total in a year	h' ₃₁ at over 12 months/12 months	h' ₃₁ at over 05-07 months/12 months	h' ₃₁ at under 05 months/12 months
3.2. Teamwork in scientific and lecturing activities (C₃₂) (+) (Johnson et al. 2012; Wardil and Hauert 2015)				
<i>Real part:</i> The number of initiatives to improve lecturing effectiveness (h ₃₂)/ Total initiatives (t ₃₂)	The ratio is between the number of initiatives to improve lecturing effectiveness and total initiatives in a year.	h ₃₂ at ministerial level/t ₃₂	h ₃₂ at school level /t ₃₂	h ₃₂ did not belong to the above two categories /t ₃₂
<i>Imaginary part:</i> The duration to complete initiatives (months) (h' ₃₂)/12 months	The ratio between the number of initiatives to duration and total in a year.	h' ₃₂ which are under 05 months/12 months	h' ₃₂ which are from 05-07 months/12 months	h' ₃₂ which are over 12 months/12 months
3.3. Participation school meetings and events (C₃₃) (+) (Frunzã 2014; Ziad 2016)				
<i>Real part:</i> The number of participation school meetings and events (h ₃₃)/ Total compulsory school meetings and events (t ₃₃)	The ratio is between the number of participation school meetings and events and total compulsory school meetings and events in a year.	h ₃₃ accounts for over 80% t ₃₃	h ₃₃ accounts for 60-80% t ₃₃	h ₃₃ accounts for under 60% t ₃₃
<i>Imaginary part:</i> The number of school meetings and events were on time (h' ₃₃)/ Total compulsory school meetings and events (t' ₃₃)	The ratio is between the number of school meetings and events be on time and total compulsory school meetings and events in a year.	h' ₃₃ which accounts for over 80% t' ₃₃	h' ₃₃ which accounts for 60-80% t' ₃₃	h' ₃₃ which accounts for under 60% t' ₃₃

Table 1**The hierarchical model for lecturer evaluation and calculation methods**

	Description/ Unit	TRUTH (T)	INDETERMINAC Y (I)	FALSITY (F)
4. Evaluation of Student (C₄)				
4.1. The content of the lessons (C₄₁) (+) (Nilson 2016; Brookfield 2017; Shingphachanh 2018)				
<i>Real part:</i> Making content comprehensible to students	The ratio is between the number of students understood lessons (h ₄₁) and total students (t ₄₁) in a year.	h ₄₁ which accounts for over 80% t ₄₁	h ₄₁ which accounts for 50- 80% t ₄₁	h ₄₁ which accounts for under 50% t ₄₁
<i>Imaginary part:</i> The duration to finish subjects (h' ₄₁)/The maximum duration to finish subjects (t' ₄₁)	The ratio between the theoretical learning duration and the maximum duration to finish subjects in a year	h' ₄₁ which account for under 40% t' ₄₁	h' ₄₁ which account for 40-70% t' ₄₁	h' ₄₁ hich account for over 70% t' ₄₁
4.2. Lecturer-student interaction (C₄₂) (+) (Kupers et al. 2015; Aldrup et al. 2018)				
<i>Real part:</i> Active learning encouraged	The ratio is between the number of discussing lessons which student have the interaction (h ₄₂) and total lessons (t ₄₂) in a year.	h ₄₂ accounts for over 60% t ₄₂	h ₄₂ accounts for 30- 60% t ₄₂	h ₄₂ accounts for under 30% t ₄₂
<i>Imaginary part:</i> The number of lessons which students asked the questions (h' ₄₂)/Total lessons (t' ₄₂)	The ratio is between the number of lessons which students asked the questions and total lessons in a year.	h' ₄₂ have under 05 questions from students/t' ₄₂	h' ₄₂ have from 05- 10 questions from students/t' ₄₂	h' ₄₂ have over 10 questions from students/t' ₄₂
4.3. The irrelevance of subjects (C₄₃) (-) (Nilson 2016; Brookfield 2017; Shingphachanh 2018)				
<i>Real part:</i> The number of lessons had irrelevance of subject to practice (h ₄₃)/Total lessons (t ₄₃)	The ratio is between the number of lessons had irrelevance of subject to practice and the total in a year.	h ₄₃ accounts for under 30% t ₄₃	h ₄₃ accounts for 30- 60% t ₄₃	h ₄₃ accounts for over 60% t ₄₃
<i>Imaginary part:</i> The number of students who complained the irrelevance of lesson into the reality (h' ₄₃)/ Total students (t' ₄₃)	The ratio is between the number of students had irrelevance of lessons to practice and the total in a year.	h' ₄₃ accounts for under 30% t' ₄₃	h' ₄₃ accounts for 30- 60% t' ₄₃	h' ₄₃ accounts for over 60% t' ₄₃

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) \bar{A} in X is:

$$\bar{A} = \left\{ (x, T_{\bar{A}}(x), I_{\bar{A}}(x), F_{\bar{A}}(x)) \mid x \in X \right\},$$

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

$$0^- \leq T_{\bar{A}}(x) + I_{\bar{A}}(x) + F_{\bar{A}}(x) \leq 3^+.$$

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 \bar{A} in X is:

$$\bar{A} = \{(x, T_{\bar{A}}(x), I_{\bar{A}}(x), F_{\bar{A}}(x)) \mid x \in X\},$$

where, the functions $T_{\bar{A}}(x), I_{\bar{A}}(x), F_{\bar{A}}(x) \subseteq [0, 1]$ define the degree of truth-membership function, indeterminacy-membership function, and the falsity-membership function respectively, so:

$$0 \leq \sup(T_{\bar{A}}(x)) + \sup(I_{\bar{A}}(x)) + \sup(F_{\bar{A}}(x)) \leq 3.$$

Definition 3. *Complex fuzzy set*

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

$$\bar{S} = \{(x, \eta_{\bar{S}}(x)) \mid x \in U\}.$$

Definition 4. *Interval-valued complex neutrosophic set* (Ali et al. 2018)

An interval-valued complex fuzzy set \bar{A} is defined over a universe of discourse X by a membership function

$$\mu_{\bar{A}} : X \rightarrow \Gamma^{[0,1]} \times R,$$

$$\mu_{\bar{A}} = r_{\bar{A}}(x) \cdot e^{j \cdot \omega_{\bar{A}}(x)}$$

In the above equation, $\Gamma^{[0,1]}$ is the collection of interval fuzzy sets and R is the set of real numbers. $r_{\bar{A}}(x)$ is the interval-valued membership function while $e^{j \cdot \omega_{\bar{A}}(x)}$ is the phase term, with $j = \sqrt{-1}$.

Definition 5. *Union of interval complex neutrosophic sets* (ICNSs)

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

Let $\mu_{\bar{A}} = r_{\bar{A}}(x) \cdot e^{j \cdot \omega_{\bar{A}}(x)}$ and $\mu_{\bar{B}} = r_{\bar{B}}(x) \cdot e^{j \cdot \omega_{\bar{B}}(x)}$ be the complex-valued membership functions of \bar{A} and \bar{B} , respectively. The membership union of $\bar{A} \cup \bar{B}$ is given by $\mu_{\bar{A} \cup \bar{B}} = [r_{\bar{A}}(x) + r_{\bar{B}}(x)] \cdot e^{j \cdot \omega_{\bar{A} \cup \bar{B}}(x)}$. Since, $r_{\bar{A}}(x)$ and $r_{\bar{B}}(x)$ are real-valued and belong to $[0, 1]$, the operators max and min can be applied to them. For calculating phase term $\omega_{\bar{A} \cup \bar{B}}(x)$, they give several methods. Let \bar{A} and \bar{B} be two IVCNSs in X , where

$$\bar{A} = \{(x, T_{\bar{A}}(x), I_{\bar{A}}(x), F_{\bar{A}}(x)) \mid x \in X\} \text{ and } \bar{B} = \{(x, T_{\bar{B}}(x), I_{\bar{B}}(x), F_{\bar{B}}(x)) \mid x \in X\}.$$

Then, the union of two interval neutrosophic sets was defined as follows:

$$\bar{A} \cup \bar{B} = \{(x, T_{\bar{A} \cup \bar{B}}(x), I_{\bar{A} \cup \bar{B}}(x), F_{\bar{A} \cup \bar{B}}(x)) \mid x \in X\}. \text{ Then}$$

$$T_{\bar{A} \cup \bar{B}}(x) = \left[\inf p_{\bar{A} \cup \bar{B}}(x), \sup p_{\bar{A} \cup \bar{B}}(x) \right] \cdot e^{j \cdot \mu_{\bar{A} \cup \bar{B}}(x)}$$

$$I_{\bar{A} \cup \bar{B}}(x) = \left[\inf q_{\bar{A} \cup \bar{B}}(x), \sup q_{\bar{A} \cup \bar{B}}(x) \right] \cdot e^{j \cdot \nu_{\bar{A} \cup \bar{B}}(x)}$$

$$F_{\bar{A} \cup \bar{B}}(x) = \left[\inf r_{\bar{A} \cup \bar{B}}(x), \sup r_{\bar{A} \cup \bar{B}}(x) \right] \cdot e^{j \cdot \omega_{\bar{A} \cup \bar{B}}(x)}$$

where

$$\inf p_{\bar{A} \cup \bar{B}}(x) = \vee \left(\inf p_{\bar{A}}(x), \inf p_{\bar{B}}(x), \sup p_{\bar{A} \cup \bar{B}}(x) \right) = \vee \left(\sup p_{\bar{A}}(x), \sup p_{\bar{B}}(x) \right),$$

$$\inf q_{\bar{A} \cup \bar{B}}(x) = \wedge \left(\inf q_{\bar{A}}(x), \inf q_{\bar{B}}(x), \sup q_{\bar{A} \cup \bar{B}}(x) \right) = \wedge \left(\sup q_{\bar{A}}(x), \sup q_{\bar{B}}(x) \right),$$

$$\inf r_{\bar{A} \cup \bar{B}}(x) = \wedge \left(\inf r_{\bar{A}}(x), \inf r_{\bar{B}}(x), \sup r_{\bar{A} \cup \bar{B}}(x) \right) = \wedge \left(\sup r_{\bar{A}}(x), \sup r_{\bar{B}}(x) \right),$$

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

$$e^{j \cdot \mu_{\bar{A} \cup \bar{B}}(x)}, e^{j \cdot \nu_{\bar{A} \cup \bar{B}}(x)}, e^{j \cdot \omega_{\bar{A} \cup \bar{B}}(x)}.$$

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

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

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

$$\bar{A} = \left\{ (x, T_{\bar{A}}(x), I_{\bar{A}}(x), F_{\bar{A}}(x)) \mid x \in X \right\} \text{ and } \bar{B} = \left\{ (x, T_{\bar{B}}(x), I_{\bar{B}}(x), F_{\bar{B}}(x)) \mid x \in X \right\}$$

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

$$\bar{A} \cap \bar{B} = \left\{ (T_{\bar{A} \cap \bar{B}}(x), I_{\bar{A} \cap \bar{B}}(x), F_{\bar{A} \cap \bar{B}}(x)) \mid x \in X \right\},$$

where

$$T_{\bar{A} \cap \bar{B}}(x) = \left[\inf p_{\bar{A} \cap \bar{B}}(x), \sup p_{\bar{A} \cap \bar{B}}(x) \right] \cdot e^{j \cdot \mu_{\bar{A} \cap \bar{B}}(x)},$$

$$I_{\bar{A} \cap \bar{B}}(x) = \left[\inf q_{\bar{A} \cap \bar{B}}(x), \sup q_{\bar{A} \cap \bar{B}}(x) \right] \cdot e^{j \cdot \nu_{\bar{A} \cap \bar{B}}(x)},$$

$$F_{\bar{A} \cap \bar{B}}(x) = \left[\inf r_{\bar{A} \cap \bar{B}}(x), \sup r_{\bar{A} \cap \bar{B}}(x) \right] \cdot e^{j \cdot \omega_{\bar{A} \cap \bar{B}}(x)},$$

$$\inf p_{\bar{A} \cap \bar{B}}(x) = \wedge \left(\inf p_{\bar{A}}(x), \inf p_{\bar{B}}(x), \sup p_{\bar{A} \cap \bar{B}}(x) \right) = \vee \left(\sup p_{\bar{A}}(x), \sup p_{\bar{B}}(x) \right),$$

$$\inf q_{\bar{A} \cap \bar{B}}(x) = \vee \left(\inf q_{\bar{A}}(x), \inf q_{\bar{B}}(x), \sup q_{\bar{A} \cap \bar{B}}(x) \right) = \wedge \left(\sup q_{\bar{A}}(x), \sup q_{\bar{B}}(x) \right),$$

$$\inf r_{\bar{A} \cap \bar{B}}(x) = \vee \left(\inf r_{\bar{A}}(x), \inf r_{\bar{B}}(x), \sup r_{\bar{A} \cap \bar{B}}(x) \right) = \wedge \left(\sup r_{\bar{A}}(x), \sup r_{\bar{B}}(x) \right),$$

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

$$e^{j \cdot \mu_{\bar{A} \cap \bar{B}}(x)}, e^{j \cdot \nu_{\bar{A} \cap \bar{B}}(x)}, e^{j \cdot \omega_{\bar{A} \cap \bar{B}}(x)}.$$

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

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

Let $\bar{A} = \left([T_{\bar{A}}^L, T_{\bar{A}}^U], [I_{\bar{A}}^L, I_{\bar{A}}^U], [F_{\bar{A}}^L, F_{\bar{A}}^U] \right)$ and $\bar{B} = \left([T_{\bar{B}}^L, T_{\bar{B}}^U], [I_{\bar{B}}^L, I_{\bar{B}}^U], [F_{\bar{B}}^L, F_{\bar{B}}^U] \right)$ be two interval neutrosophic sets over X which are defined by $[T_{\bar{A}}^L, T_{\bar{A}}^U] = [p_{\bar{A}}^L, p_{\bar{A}}^U] \cdot e^{j\pi[\mu_{\bar{A}}^L(x), \mu_{\bar{A}}^U(x)]}$,

$[I_{\bar{A}}^L, I_{\bar{A}}^U] = [q_{\bar{A}}^L, q_{\bar{A}}^U] \cdot e^{j\pi[\nu_{\bar{A}}^L(x), \nu_{\bar{A}}^U(x)]}$, $[F_{\bar{A}}^L, F_{\bar{A}}^U] = [r_{\bar{A}}^L, r_{\bar{A}}^U] \cdot e^{j\pi[\omega_{\bar{A}}^L(x), \omega_{\bar{A}}^U(x)]}$ respectively. Then, the operational rules of ICNS are defined as follows. The complement of x is:

$$\bar{A} = \left(\left[p_A^L, p_A^U \right], \left[q_A^L, q_A^U \right], \left[r_A^L, r_A^U \right] \right) \text{ and } \bar{B} = \left(\left[p_B^L, p_B^U \right], \left[q_B^L, q_B^U \right], \left[r_B^L, r_B^U \right] \right)$$

(i) The complement of \bar{A} is defined as:

$$\bar{A} = \left(\left[r_A^L, r_A^U \right] \cdot e^{j\pi[2-\mu_A^L(x), 2-\mu_A^U(x)]}, \left[1-q_A^L, 1-q_A^U \right] \cdot e^{j\pi[2-\nu_A^L(x), 2-\nu_A^U(x)]}, \left[p_A^L, p_A^U \right] \cdot e^{j\pi[2-\omega_A^L(x), 2-\omega_A^U(x)]} \right) \tag{1}$$

(ii) The addition of \bar{A} and \bar{B} is defined as:

$$\bar{A} \oplus \bar{B} = \left(\left[p_A^L + p_B^L - p_A^L p_B^L, p_B^U + p_A^U - p_A^U p_B^U \right] \cdot e^{j\pi[\mu_A^L(x)+\mu_B^L(x), \mu_A^U(x)+\mu_B^U(x)]}, \right. \\ \left. \left[q_A^L q_B^L, q_A^U q_B^U \right] \cdot e^{j\pi[\nu_A^L(x)+\nu_B^L(x), \nu_A^U(x)+\nu_B^U(x)]}, \left[r_A^L r_B^L, r_A^U r_B^U \right] \cdot e^{j\pi[\omega_A^L(x)+\omega_B^L(x), \omega_A^U(x)+\omega_B^U(x)]} \right) \tag{2}$$

(iii) The bounded between \bar{A} and \bar{B} is defined as:

$$\bar{A} - \bar{B} = \max \left(\begin{array}{l} 0, \left[p_A^L - p_B^L, p_B^U - p_A^U \right] \cdot e^{j\pi \cdot \max(0, [\mu_A^L(x) - \mu_B^L(x), \mu_A^U(x) - \mu_B^U(x)])} \\ \left[q_A^L - q_B^L, q_B^U - q_A^U \right] \cdot e^{j\pi \cdot \max(0, [\nu_A^L(x) - \nu_B^L(x), \nu_A^U(x) - \nu_B^U(x)])} \\ \left[r_A^L - r_B^L, r_A^U - r_B^U \right] \cdot e^{j\pi \cdot \max(0, [\omega_A^L(x) - \omega_B^L(x), \omega_A^U(x) - \omega_B^U(x)])} \end{array} \right) \tag{3}$$

(iv) The product of \bar{A} and \bar{B} is defined as:

$$\bar{A} \otimes \bar{B} = \left(\begin{array}{l} \left[p_A^L p_B^L, p_A^U p_B^U \right] \cdot e^{j\pi[\mu_A^L(x), \mu_B^L(x), \mu_A^U(x), \mu_B^U(x)]}, \\ \left[q_A^L + q_B^L - q_A^L q_B^L, q_A^U + q_B^U - q_A^U q_B^U \right] \cdot e^{j\pi[\nu_A^L(x)+\nu_B^L(x), \nu_A^U(x)+\nu_B^U(x)]}, \\ \left[r_A^L + r_B^L - r_A^L r_B^L, r_A^U + r_B^U - r_A^U r_B^U \right] \cdot e^{j\pi[\omega_A^L(x)+\omega_B^L(x), \omega_A^U(x)+\omega_B^U(x)]} \end{array} \right) \tag{4}$$

(v) The scalar multiplication of \bar{A} is defined as:

$$k \otimes \bar{A} = \left(\begin{array}{l} \left[1 - (1 - p_A^L)^k, 1 - (1 - p_A^U)^k \right] \cdot e^{j\pi[k \cdot \mu_A^L(x), k \cdot \mu_A^U(x)]} \\ \left[(q_A^L)^k, (q_A^U)^k \right] \cdot e^{j\pi[k \cdot \nu_A^L(x), k \cdot \nu_A^U(x)]}, \\ \left[(r_A^L)^k, (r_A^U)^k \right] \cdot e^{j\pi[k \cdot \omega_A^L(x), k \cdot \omega_A^U(x)]} \end{array} \right) \quad (k > 0) \tag{5}$$

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

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

Let $\bar{A} = \left(\left[T_A^L, T_A^U \right], \left[I_A^L, I_A^U \right], \left[F_A^L, F_A^U \right] \right)$ and $\bar{B} = \left(\left[T_B^L, T_B^U \right], \left[I_B^L, I_B^U \right], \left[F_B^L, F_B^U \right] \right)$ be two interval neutrosophic sets over X which are defined by $\left[T_A^L, T_A^U \right] = \left[p_A^L, p_A^U \right] \cdot e^{j\pi[\mu_A^L(x), \mu_A^U(x)]}$, $\left[I_A^L, I_A^U \right] = \left[q_A^L, q_A^U \right] \cdot e^{j\pi[\nu_A^L(x), \nu_A^U(x)]}$, $\left[F_A^L, F_A^U \right] = \left[r_A^L, r_A^U \right] \cdot e^{j\pi[\omega_A^L(x), \omega_A^U(x)]}$, respectively.

The difference between \bar{A} and \bar{B} is defined as follows:

$$d(\bar{A}, \bar{B}) = \frac{1}{3n} \left(\max(|T_A^L - T_B^L|, |T_A^U - T_B^U|) + \max(|I_A^L - I_B^L|, |I_A^U - I_B^U|) + \max(|F_A^L - F_B^L|, |F_A^U - F_B^U|) \right) \tag{6}$$

$$= \max \left(\begin{array}{l} \max \left(\left| p_A^L - p_B^L \right| \cdot e^{j\pi \cdot \max(0, \mu_A^L - \mu_B^L)}, \left| p_A^U - p_B^U \right| \cdot e^{j\pi \cdot \max(0, \mu_A^U - \mu_B^U)} \right) + \\ \max \left(\left| q_A^L - q_B^L \right| \cdot e^{j\pi \cdot \max(0, \nu_A^L - \nu_B^L)}, \left| q_A^U - q_B^U \right| \cdot e^{j\pi \cdot \max(0, \nu_A^U - \nu_B^U)} \right) + \\ \max \left(\left| r_A^L - r_B^L \right| \cdot e^{j\pi \cdot \max(0, \omega_A^L - \omega_B^L)}, \left| r_A^U - r_B^U \right| \cdot e^{j\pi \cdot \max(0, \omega_A^U - \omega_B^U)} \right) \end{array} \right).$$

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 2

Linguistic variables for weighting importance based on interval complex neutrosophic values

Linguistic variables	Amplitude term	Phase term
Far Below Standards (FBS)	([0.1, 0.2], [0.1, 0.2], [0.7, 0.8])	([0.1, 0.2], [0.2, 0.3], [0.7, 0.8])
Below Standards (BS)	([0.2, 0.3], [0.2, 0.3], [0.5, 0.6])	([0.2, 0.3], [0.3, 0.4], [0.5, 0.6])
Meets Standards (MS)	([0.3, 0.4], [0.3, 0.4], [0.3, 0.4])	([0.3, 0.4], [0.4, 0.5], [0.3, 0.4])
Above Standards (AS)	([0.5, 0.6], [0.2, 0.3], [0.2, 0.3])	([0.5, 0.6], [0.3, 0.4], [0.2, 0.3])
Far Above Standards (FAS)	([0.7, 0.8], [0.1, 0.2], [0.1, 0.2])	([0.7, 0.8], [0.2, 0.3], [0.1, 0.2])

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 3

Standard values of the relative importance of factors (Saaty 1980)

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

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).

$$w_{ii} = \frac{a_{ii}}{\sum_{i=1}^n a_{ni}} \quad (7)$$

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

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

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

Consistency Ratio (<i>CR</i>)	λ_{\max}	Consistency Index
$CR = \frac{CI}{RI}$ (8)	$\lambda_{\max} = \frac{\sum_1^n y}{n}$ (9)	$CI = \frac{\lambda_{\max} - n}{\sum_{i=1}^n n - 1}$ (10)

In which: CI: The matrix's consistency index
RI: The average consistency of random matrices
 λ_{\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 ≤ 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)

n	1	2	3	4	5	6	7	8	9	10
<i>RI</i>	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

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 logicity, 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, \bar{D} , based on a given set of criteria and sub-criteria. The data of decision matrix \bar{D} come from different sources. $\bar{D} = [d_{ij}]_{m \times n}$ is the decision matrix, where $d = ([T_{ij}^L, T_{ij}^U], [I_{ij}^L, I_{ij}^U], [F_{ij}^L, F_{ij}^U])$ and takes the form of the INVs for alternative A_i with respect to attribute C_j in which $[T_{ij}^L, T_{ij}^U] = [p_{ij}^L, p_{ij}^U], e^{j\pi[\mu_{ij}^L(x), \mu_{ij}^U(x)]}$, $[I_{ij}^L, I_{ij}^U] = [q_{ij}^L, q_{ij}^U], e^{j\pi[v_{ij}^L(x), v_{ij}^U(x)]}$, $[F_{ij}^L, F_{ij}^U] = [r_{ij}^L, r_{ij}^U]$ and $e^{j\pi[\omega_{ij}^L(x), \omega_{ij}^U(x)]}$, respectively.

For a multiple attribute decision problem, let $A = (A_1, A_2, \dots, A_n)$ be a discrete set of alternatives, $C = (C_1, C_2, \dots, C_n)$ be the set of attributes, $W = (w_1, w_2, \dots, w_n)^T$ 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 \bar{R} , as follows:

$$\bar{R} = [r_{ij}]_{m \times n} \text{ is the decision matrix, where } r_{ij} = \left([T_{ij}^{L*}, T_{ij}^{U*}], [I_{ij}^{L*}, I_{ij}^{U*}], [F_{ij}^{L*}, F_{ij}^{U*}] \right).$$

$$\bar{R} = \begin{cases} d_{ij} & \text{if criterion } C_{ij} \text{ is benefit type} \\ \bar{d}_{ij} & \text{if criterion } C_{ij} \text{ is cost type} \end{cases} \quad (11)$$

where x_{\max} and x_{\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.

$$\begin{cases} S_j^+ = \left([\max T_{ij}^{L**}, \max T_{ij}^{U**}], [\min I_{ij}^{L**}, \min I_{ij}^{U**}], [\min F_{ij}^{L**}, \min F_{ij}^{U**}] \right) \\ S_j^- = \left([\min T_{ij}^{L**}, \min T_{ij}^{U**}], [\max I_{ij}^{L**}, \max I_{ij}^{U**}], [\max F_{ij}^{L**}, \max F_{ij}^{U**}] \right) \end{cases} \quad j = 1, 2, \dots, n \quad (12)$$

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:

$$\begin{cases} D_i^+ = \sum_{j=1}^n d(S_j, S_j^+) \\ D_i^- = \sum_{j=1}^n d(S_j, S_j^-) \end{cases} \quad i = 1, 2, \dots, m \quad (13)$$

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 $\bar{A} = \left([T_{\bar{A}}^L, T_{\bar{A}}^U], [I_{\bar{A}}^L, I_{\bar{A}}^U], [F_{\bar{A}}^L, F_{\bar{A}}^U] \right)$ and $\bar{B} = \left([T_{\bar{B}}^L, T_{\bar{B}}^U], [I_{\bar{B}}^L, I_{\bar{B}}^U], [F_{\bar{B}}^L, F_{\bar{B}}^U] \right)$ be two interval neutrosophic sets over X which are defined by $[T_{\bar{A}}^L, T_{\bar{A}}^U] = [p_{\bar{A}}^L, p_{\bar{A}}^U] \cdot e^{j\pi[\mu_{\bar{A}}^L(x), \mu_{\bar{A}}^U(x)]}$, $[I_{\bar{A}}^L, I_{\bar{A}}^U] = [q_{\bar{A}}^L, q_{\bar{A}}^U] \cdot e^{j\pi[v_{\bar{A}}^L(x), v_{\bar{A}}^U(x)]}$, $[F_{\bar{A}}^L, F_{\bar{A}}^U] = [r_{\bar{A}}^L, r_{\bar{A}}^U] \cdot e^{j\pi[\omega_{\bar{A}}^L(x), \omega_{\bar{A}}^U(x)]}$, respectively. Let $x_{\bar{X}}^a$ and $x_{\bar{X}}^p$ be the amplitude and phase terms of interval complex number \bar{X} . The comparison between \bar{A} and \bar{B} is defined as follows:

$$\bar{A} > \bar{B} \Leftrightarrow \begin{cases} x_{\bar{A}}^a > x_{\bar{B}}^a \\ x_{\bar{A}}^a = x_{\bar{B}}^a, x_{\bar{A}}^p > x_{\bar{B}}^p \end{cases} \quad (14)$$

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 $D^{(k)} = (d_{ij}^{(k)})_{m \times n}$ ($k = 1, 2, \dots, l$) 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.

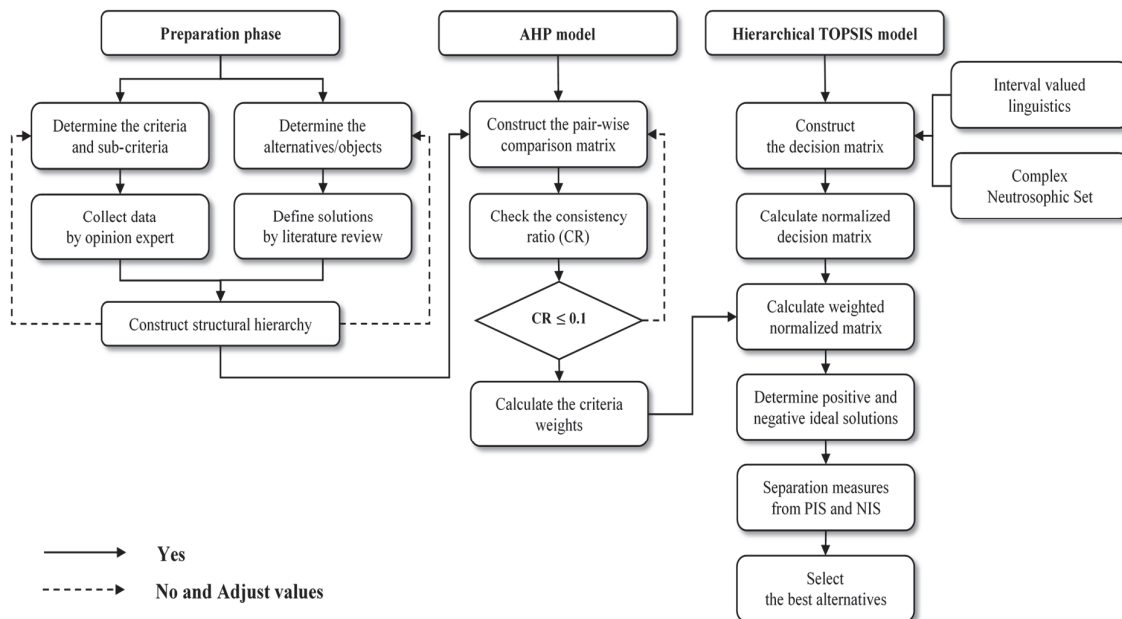


Fig 2. The methodology framework

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

Criteria	C_1	C_2	C_3	C_4	Weight
C_1	1	1/4	1/2	1/5	0.084
C_2	4	1	3	2	0.454
C_3	2	1/3	1	1/2	0.154
C_4	5	1/2	2	1	0.309

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

Table 7

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

Sub-criteria	Pair-wise comparison matrix			Local Weight	Overall Weight	Sub-criteria	Pair-wise comparison matrix				Local Weight	Overall Weight
	C_{11}	C_{12}	C_{13}				C_{21}	C_{22}	C_{23}	C_{24}		
C_{11}	1	1/3	2	0.230	0.018	C_{21}	1	2	3	5	0.471	0.214
C_{12}	3	1	5	0.640	0.053	C_{22}	1/2	1	2	4	0.284	0.129
C_{13}	1/2	1/5	1	0.130	0.011	C_{23}	1/3	1/2	1	3	0.171	0.078
W = 0.083; $\lambda_{max} = 3.005, CI = 0.003, RI = 0.58, CR = 0.005 < 0.10$						C_{24}	1/5	1/4	1/3	1	0.074	0.034
						W = 0.454; $\lambda_{max} = 4.066, CI = 0.0219, RI = 0.900, CR = 0.0240 < 0.10$						
Sub-criteria	Pair-wise comparison matrix			Local Weight	Overall Weight	Sub-criteria	Pair-wise comparison matrix			Local Weight	Overall Weight	
	C_{31}	C_{32}	C_{33}				C_{41}	C_{42}	C_{43}			
C_{31}	1	1/2	4	0.324	0.050	C_{41}	1	3	4	0.623	0.193	
C_{32}	2	1	6	0.587	0.090	C_{42}	1/3	1	2	0.239	0.074	
C_{33}	1/4	1/6	1	0.089	0.014	C_{43}	1/4	1/2	1	0.137	0.042	
W = 0.154; $\lambda_{max} = 3.012, CI = 0.006, RI = 0.580, CR = 0.011 < 0.10$						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

Criteria	A_1	A_2	A_3	A_4	A_5	Criteria	A_1	A_2	A_3	A_4	A_5
C_{11}	MS	AS	BS	AS	MS	C_{31}	MS	AS	AS	AS	MS
C_{12}	MS	BS	MS	AS	AS	C_{32}	MS	FAS	MS	MS	FAS
C_{13}	MS	AS	BS	AS	BS	C_{33}	FBS	MS	MS	MS	MS
C_{21}	FAS	AS	FAS	MS	AS	C_{41}	BS	FAS	FAS	MS	MS
C_{22}	BS	MS	AS	MS	AS	C_{42}	AS	FAS	FAS	FAS	FAS
C_{23}	BS	MS	MS	FAS	FAS	C_{43}	AS	AS	FAS	FAS	FAS
C_{24}	BS	AS	AS	MS	MS						

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 11
The weight matrix of the attributes in the decision matrix

	A_1		A_2		A_3		A_4		A_5	
	Amplitude term	Phase term	Amplitude term	Phase term	Amplitude term	Phase term	Amplitude term	Phase term	Amplitude term	Phase term
C_{11}	{[0.006, 0.009], [0.979, 0.984], [0.979, 0.984]}	{[0.005, 0.007], [0.007, 0.009], [0.005, 0.007]}	{[0.012, 0.016], [0.971, 0.979], [0.971, 0.979]}	{[0.009, 0.011], [0.005, 0.007], [0.004, 0.005]}	{[0.004, 0.006], [0.971, 0.979], [0.988, 0.991]}	{[0.004, 0.005], [0.005, 0.007], [0.009, 0.011]}	{[0.012, 0.016], [0.971, 0.979], [0.971, 0.979]}	{[0.009, 0.011], [0.005, 0.007], [0.004, 0.005]}	{[0.006, 0.009], [0.979, 0.984], [0.979, 0.984]}	{[0.005, 0.007], [0.007, 0.009], [0.005, 0.007]}
C_{12}	{[0.019, 0.027], [0.938, 0.953], [0.938, 0.953]}	{[0.016, 0.021], [0.021, 0.027], [0.016, 0.021]}	{[0.012, 0.019], [0.918, 0.938], [0.964, 0.973]}	{[0.011, 0.016], [0.016, 0.021], [0.027, 0.032]}	{[0.019, 0.027], [0.938, 0.953], [0.938, 0.953]}	{[0.016, 0.021], [0.021, 0.027], [0.016, 0.021]}	{[0.036, 0.047], [0.918, 0.938], [0.918, 0.938]}	{[0.027, 0.032], [0.016, 0.021], [0.011, 0.016]}	{[0.036, 0.047], [0.918, 0.938], [0.918, 0.938]}	{[0.027, 0.032], [0.016, 0.021], [0.011, 0.016]}
C_{13}	{[0.004, 0.006], [0.987, 0.990], [0.987, 0.990]}	{[0.003, 0.004], [0.004, 0.006], [0.003, 0.004]}	{[0.008, 0.010], [0.982, 0.987], [0.982, 0.987]}	{[0.006, 0.007], [0.003, 0.004], [0.002, 0.003]}	{[0.002, 0.004], [0.982, 0.987], [0.992, 0.994]}	{[0.002, 0.003], [0.003, 0.004], [0.006, 0.007]}	{[0.008, 0.010], [0.982, 0.987], [0.982, 0.987]}	{[0.006, 0.007], [0.003, 0.004], [0.002, 0.003]}	{[0.002, 0.004], [0.982, 0.987], [0.992, 0.994]}	{[0.002, 0.003], [0.003, 0.004], [0.006, 0.007]}
C_{21}	{[0.227, 0.291], [0.611, 0.709], [0.611, 0.709]}	{[0.150, 0.171], [0.043, 0.064], [0.021, 0.043]}	{[0.138, 0.178], [0.709, 0.773], [0.709, 0.773]}	{[0.107, 0.128], [0.064, 0.086], [0.043, 0.064]}	{[0.227, 0.291], [0.611, 0.709], [0.611, 0.709]}	{[0.125, 0.171], [0.043, 0.064], [0.021, 0.043]}	{[0.073, 0.104], [0.773, 0.822], [0.773, 0.822]}	{[0.064, 0.086], [0.086, 0.107], [0.064, 0.086]}	{[0.138, 0.178], [0.709, 0.773], [0.709, 0.773]}	{[0.107, 0.128], [0.064, 0.086], [0.043, 0.064]}
C_{22}	{[0.028, 0.045], [0.813, 0.856], [0.914, 0.936]}	{[0.026, 0.039], [0.039, 0.052], [0.065, 0.077]}	{[0.045, 0.064], [0.856, 0.889], [0.856, 0.889]}	{[0.039, 0.052], [0.052, 0.065], [0.039, 0.052]}	{[0.086, 0.111], [0.813, 0.856], [0.813, 0.856]}	{[0.065, 0.077], [0.039, 0.052], [0.026, 0.039]}	{[0.045, 0.064], [0.856, 0.889], [0.856, 0.889]}	{[0.039, 0.052], [0.052, 0.065], [0.039, 0.052]}	{[0.086, 0.111], [0.813, 0.856], [0.813, 0.856]}	{[0.065, 0.077], [0.039, 0.052], [0.026, 0.039]}
C_{23}	{[0.053, 0.069], [0.973, 0.983], [0.947, 0.961]}	{[0.133, 0.140], [0.125, 0.133], [0.109, 0.117]}	{[0.027, 0.039], [0.961, 0.973], [0.910, 0.931]}	{[0.125, 0.133], [0.117, 0.125], [0.125, 0.133]}	{[0.027, 0.039], [0.961, 0.973], [0.910, 0.931]}	{[0.125, 0.133], [0.117, 0.125], [0.125, 0.133]}	{[0.008, 0.017], [0.983, 0.992], [0.973, 0.983]}	{[0.094, 0.101], [0.133, 0.140], [0.140, 0.148]}	{[0.008, 0.017], [0.983, 0.992], [0.973, 0.983]}	{[0.094, 0.101], [0.133, 0.140], [0.140, 0.148]}
C_{24}	{[0.008, 0.012], [0.947, 0.960], [0.977, 0.983]}	{[0.007, 0.010], [0.010, 0.014], [0.017, 0.020]}	{[0.023, 0.031], [0.947, 0.960], [0.947, 0.960]}	{[0.017, 0.020], [0.010, 0.014], [0.007, 0.010]}	{[0.023, 0.031], [0.947, 0.960], [0.947, 0.960]}	{[0.017, 0.020], [0.010, 0.014], [0.007, 0.010]}	{[0.012, 0.017], [0.960, 0.969], [0.960, 0.969]}	{[0.010, 0.014], [0.014, 0.017], [0.010, 0.014]}	{[0.012, 0.017], [0.960, 0.969], [0.960, 0.969]}	{[0.010, 0.014], [0.014, 0.017], [0.010, 0.014]}
C_{31}	{[0.018, 0.025], [0.942, 0.955], [0.942, 0.955]}	{[0.015, 0.020], [0.020, 0.025], [0.015, 0.020]}	{[0.034, 0.045], [0.923, 0.942], [0.923, 0.942]}	{[0.025, 0.030], [0.015, 0.020], [0.010, 0.015]}	{[0.034, 0.045], [0.923, 0.942], [0.923, 0.942]}	{[0.025, 0.030], [0.015, 0.020], [0.010, 0.015]}	{[0.034, 0.045], [0.923, 0.942], [0.923, 0.942]}	{[0.025, 0.030], [0.015, 0.020], [0.010, 0.015]}	{[0.018, 0.025], [0.942, 0.955], [0.942, 0.955]}	{[0.015, 0.020], [0.020, 0.025], [0.015, 0.020]}
C_{32}	{[0.032, 0.045], [0.897, 0.921], [0.897, 0.921]}	{[0.027, 0.036], [0.036, 0.045], [0.027, 0.036]}	{[0.103, 0.135], [0.813, 0.865], [0.813, 0.865]}	{[0.063, 0.072], [0.018, 0.027], [0.009, 0.018]}	{[0.032, 0.045], [0.897, 0.921], [0.897, 0.921]}	{[0.027, 0.036], [0.036, 0.045], [0.027, 0.036]}	{[0.032, 0.045], [0.897, 0.921], [0.897, 0.921]}	{[0.027, 0.036], [0.036, 0.045], [0.027, 0.036]}	{[0.103, 0.135], [0.813, 0.865], [0.813, 0.865]}	{[0.063, 0.072], [0.018, 0.027], [0.009, 0.018]}
C_{33}	{[0.001, 0.003], [0.968, 0.978], [0.995, 0.997]}	{[0.001, 0.003], [0.003, 0.004], [0.010, 0.011]}	{[0.005, 0.007], [0.983, 0.987], [0.983, 0.987]}	{[0.004, 0.006], [0.006, 0.007], [0.004, 0.006]}	{[0.005, 0.007], [0.983, 0.987], [0.983, 0.987]}	{[0.004, 0.006], [0.006, 0.007], [0.004, 0.006]}	{[0.005, 0.007], [0.983, 0.987], [0.983, 0.987]}	{[0.004, 0.006], [0.006, 0.007], [0.004, 0.006]}	{[0.005, 0.007], [0.983, 0.987], [0.983, 0.987]}	{[0.004, 0.006], [0.006, 0.007], [0.004, 0.006]}
C_{41}	{[0.042, 0.067], [0.733, 0.793], [0.875, 0.906]}	{[0.039, 0.058], [0.058, 0.077], [0.097, 0.116]}	{[0.207, 0.267], [0.641, 0.733], [0.641, 0.733]}	{[0.135, 0.154], [0.039, 0.058], [0.019, 0.039]}	{[0.207, 0.267], [0.641, 0.733], [0.641, 0.733]}	{[0.135, 0.154], [0.039, 0.058], [0.019, 0.039]}	{[0.067, 0.094], [0.793, 0.838], [0.793, 0.838]}	{[0.058, 0.077], [0.077, 0.097], [0.058, 0.077]}	{[0.067, 0.094], [0.793, 0.838], [0.793, 0.838]}	{[0.058, 0.077], [0.077, 0.097], [0.058, 0.077]}
C_{42}	{[0.050, 0.066], [0.888, 0.915], [0.888, 0.915]}	{[0.037, 0.044], [0.022, 0.030], [0.015, 0.022]}	{[0.085, 0.112], [0.843, 0.888], [0.843, 0.888]}	{[0.052, 0.059], [0.015, 0.022], [0.007, 0.015]}	{[0.085, 0.112], [0.843, 0.888], [0.843, 0.888]}	{[0.052, 0.059], [0.015, 0.022], [0.007, 0.015]}	{[0.085, 0.112], [0.843, 0.888], [0.843, 0.888]}	{[0.052, 0.059], [0.015, 0.022], [0.007, 0.015]}	{[0.085, 0.112], [0.843, 0.888], [0.843, 0.888]}	{[0.052, 0.059], [0.015, 0.022], [0.007, 0.015]}
C_{43}	{[0.009, 0.015], [0.985, 0.991], [0.971, 0.979]}	{[0.059, 0.063], [0.067, 0.071], [0.071, 0.076]}	{[0.009, 0.015], [0.985, 0.991], [0.971, 0.979]}	{[0.059, 0.063], [0.067, 0.071], [0.071, 0.076]}	{[0.004, 0.009], [0.991, 0.996], [0.985, 0.991]}	{[0.050, 0.055], [0.071, 0.076], [0.076, 0.080]}	{[0.004, 0.009], [0.991, 0.996], [0.991, 0.996]}	{[0.050, 0.055], [0.071, 0.076], [0.071, 0.076]}	{[0.004, 0.009], [0.991, 0.996], [0.991, 0.996]}	{[0.050, 0.055], [0.071, 0.076], [0.071, 0.076]}

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 12
The positive and negative ideal solutions (PIS and NIS)

	S^+		S^-	
	Amplitude term	Phase term	Amplitude term	Phase term
C_{11}	{[0.227, 0.291], [0.987, 0.991], [0.995, 0.997]}	{[0.150, 0.171], [0.125, 0.133], [0.109, 0.117]}	{[0.001, 0.003], [0.611, 0.709], [0.611, 0.709]}	{[0.001, 0.003], [0.003, 0.004], [0.003, 0.004]}
C_{12}	{[0.207, 0.267], [0.985, 0.991], [0.983, 0.987]}	{[0.135, 0.154], [0.117, 0.125], [0.125, 0.133]}	{[0.005, 0.007], [0.641, 0.733], [0.641, 0.733]}	{[0.004, 0.006], [0.003, 0.004], [0.002, 0.003]}
C_{13}	{[0.227, 0.291], [0.991, 0.996], [0.992, 0.994]}	{[0.150, 0.007], [0.004, 0.006], [0.006, 0.007]}	{[0.002, 0.004], [0.611, 0.709], [0.611, 0.709]}	{[0.002, 0.003], [0.003, 0.004], [0.004, 0.006]}
C_{21}	{[0.085, 0.112], [0.991, 0.996], [0.985, 0.991]}	{[0.094, 0.101], [0.133, 0.140], [0.140, 0.148]}	{[0.004, 0.007], [0.773, 0.822], [0.773, 0.822]}	{[0.004, 0.006], [0.003, 0.004], [0.002, 0.003]}
C_{22}	{[0.107, 0.128], [0.133, 0.140], [0.140, 0.148]}	{[0.107, 0.128], [0.133, 0.140], [0.140, 0.148]}	{[0.002, 0.004], [0.709, 0.773], [0.709, 0.773]}	{[0.002, 0.003], [0.003, 0.004], [0.004, 0.006]}

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.

Table 13

The separation measures of each point to PIS and NIS

	d^+					d^-				
	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5
Amplitude term	5.288	5.075	5.597	4.056	21.358	8.180	7.022	7.815	4.001	5.522
Phase term	3.79	3.782	2.490	3.411	3.729	1.353	1.329	1.328	1.329	1.328

Step 11: Rank the alternatives: $A_1 > A_3 > A_2 > A_5 > A_4$.

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

Comparison of interval complex neutrosophic TOPSIS with other methods of lecturer evaluation

Method	Object	Aggregative value					Ranking
		A_1	A_2	A_3	A_4	A_5	
Proposed method	D^+	$5.288.e^{3.790}$	$5.075.e^{3.782}$	$5.597.e^{2.490}$	$4.056.e^{3.411}$	$21.358.e^{3.729}$	$A_1 > A_3 > A_2 >$
	D^-	$8.180.e^{1.353}$	$7.022.e^{1.329}$	$7.815.e^{1.328}$	$4.001.e^{1.329}$	$5.522.e^{1.328}$	$A_5 > A_4$
Interval neutrosophic TOPSIS (Chi and Liu, 2013)	D^+	0.384	0.140	0.221	0.395	0.339	$A_2 > A_3 > A_5 >$
	D^-	0.237	0.481	0.400	0.225	0.281	$A_1 > A_4$
AHP (Saaty, 1980)		5.164	7.060	7.460	6.040	6.906	$A_3 > A_2 > A_5 > A_4 > A_1$
TOPSIS (Hwang and Yoon, 1981)	D^+	0.098	0.043	0.033	0.079	0.059	$A_3 > A_2 > A_5 >$
	D^-	0.051	0.090	0.104	0.051	0.070	$A_4 > A_1$

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

- Akram, M., Shahzadi, S., & Smarandache, F. (2018). Multi-Attribute Decision-Making Method Based on Neutrosophic Soft Rough Information. *Axioms*, 7(1), 19. doi:10.3390/axioms7010019
- Aldrup, K., Klusmann, U., Lüdtke, O., Göllner, R., & Trautwein, U. (2018). Student misbehavior and teacher well-being: Testing the mediating role of the teacher-student relationship. *Learning and Instruction*, 58, 126-136.
- Ali, M., Dat, L. Q., Son, L. H., & Smarandache, F. (2018). Interval Complex Neutrosophic Set: Formulation and Applications in Decision-Making. *International Journal of Fuzzy Systems*, 20(3), 986-999.
- Ali, M., & Smarandache, F. (2017). Complex neutrosophic set. *Neural Computing and Applications*, 28(7), 1817-1834.
- Alias, M., Masek, A., & Salleh, H. H. M. (2015). Self, Peer and Teacher Assessments in Problem Based Learning: Are They in Agreements? *Procedia - Social and Behavioral Sciences*, 204, 309-317.
- Almeida, J. de C. (2017). Teacher Performance Evaluation: The Importance of Performance Standards. *International Journal for Cross-Disciplinary Subjects in Education*, 8(1), 2973-2981.
- Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 20(1), 87-96.
- Bambaeroo, F., & Shokrpour, N. (2017). The impact of the teachers' non-verbal communication on success in teaching. *Journal of Advances in Medical Education and Professionalism*, 5(2), 51-59.
- Baykasoğlu, A., Kaplanoğlu, V., Durmuşoğlu, Z. D. U., & Şahin, C. (2013). Integrating fuzzy DEMATEL and fuzzy hierarchical TOPSIS methods for truck selection. *Expert Systems with Applications*, 40(3), 899-907.
- Biggs, J. B., & Collis, K. F. (2014). *Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome)*. Academic Press, 245 pages.
- Bohlmann, N. L., & Weinstein, R. S. (2013). Classroom context, teacher expectations, and cognitive level: Predicting children's math ability judgments. *Journal of Applied Developmental Psychology*, 34(6), 288-298.
- Bottani, E., & Rizzi, A. (2006). A fuzzy TOPSIS methodology to support outsourcing of logistics services. *Supply Chain Management: An International Journal*, 11(4), 294-308. doi: 10.1108/13598540610671743
- Bradford, A. (2015). Adopting English-Taught Degree Programs. *International Higher Education*, 69, 8-10.
- Brookfield, S. D. (2017). *Becoming a Critically Reflective Teacher*. John Wiley and Sons.
- Buttram, J. L., & Wilson, B. L. (1987). Promising trends in teacher evaluation. *Educational Leadership*, 44(7), 4-6.
- Cegarra-Navarro, J.-G., Soto-Acosta, P., and Martinez-Caro, E. (2016). Linking counter-knowledge to goal orientation through an unlearning context - A study from a Spanish University. *Learning and Individual Differences*, 45, 260-267.
- Cegarra-Sánchez, J., & Cegarra-Navarro, J.-G. (2017). Making meaning out of noise: A knowledge management core competence for higher education students. *VINE Journal of Information and Knowledge Management Systems*, 47(4), 506-521.
- Chappuis, S., Commodore, C., and Stiggins, R. (2016). *Balanced Assessment Systems: Leadership, Quality, and the Role of Classroom Assessment*. Corwin Press.
- Cheng, M. M. H., Chan, K.-W., Tang, S. Y. F., and Cheng, A. Y. N. (2009). Pre-service teacher education students' epistemological beliefs and their conceptions of teaching. *Teaching and Teacher Education*, 25(2), 319-327.
- Chi, P., & Liu, P. (2013). An extended TOPSIS method for the multiple attribute decision making problems based on interval neutrosophic set. *Neutrosophic Sets and Systems*, 1, 1-8.
- Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002). Teacher evaluation: A review of the literature. *Annual Meeting of the American Educational Research Association*, New Orleans, LA, 1-18.

- Cuevas, R., Ntoumanis, N., Fernandez-Bustos, J. G., and Bartholomew, K. (2018). Does teacher evaluation based on student performance predict motivation, well-being, and ill-being? *Journal of School Psychology, 68*, 154-162.
- Danielson, C. (2000). *Teacher evaluation to enhance professional practice*. Retrieved from <http://ebookcentral.proquest.com/lib/ucm/detail.action?docID=280406>
- Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education, 40*(3), 291-309.
- Davey, B. (1991). Evaluating teacher competence through the use of performance assessment tasks: An overview. *Journal of Personnel Evaluation in Education, 5*(2), 121-132.
- Derrington, M. L., and Campbell, J. W. (2015). Implementing new teacher evaluation systems: Principals' concerns and supervisor support. *Journal of Educational Change, 16*(3), 305-326.
- Fabjanowicz, M., Bystrzanowska, M., Namieśnik, J., Tobiszewski, M., and Płotka-Wasyłka, J. (2018). An analytical hierarchy process for selection of the optimal procedure for resveratrol determination in wine samples. *Microchemical Journal, 142*, 126-134.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., and Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1-9.
- Fischer, C., Fishman, B., Dede, C., Eisenkraft, A., Frumin, K., Foster, B., ... McCoy, A. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform. *Teaching and Teacher Education, 72*, 107-121.
- Frunzã, V. (2014). Implications of Teaching Styles on Learning Efficiency. *Procedia - Social and Behavioral Sciences, 127*, 342-346.
- Gormally, C., Evans, M., and Brickman, P. (2014). Feedback about teaching in higher Ed: Neglected opportunities to promote change. *CBE Life Sciences Education, 13*(2), 187-199.
- Hein, N., Kroenke, A., & Júnior, M. M. R. (2015a). Professor assessment using multi-criteria decision analysis. *Procedia Computer Science, 55*, 539-548.
- Hwang, C.-L., & Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey*. Berlin Heidelberg: Springer-Verlag.
- Iantorno, S. E., Andras, L. M., and Skaggs, D. L. (2016). Variability of reviewers' comments in the peer review process for orthopaedic research. *Spine Deformity, 4*(4), 268-271.
- Ishizaka, A., & Labib, A. (2009). Analytic hierarchy process and expert choice: Benefits and limitations. *OR Insight, 22*(4), 201-220.
- Jaramillo, I. F., Pico, R. B., & Marquez, C. V. (2017). A model for faculty evaluation in higher education ecuadorian through multi-criteria decision Analysis. *Indian Journal of Science and Technology, 10*(18), 1-8.
- Jiayi, W., & Ling, C. (2012). Reviewing teacher evaluation of rewards and punishments: The overview of Chinese teacher evaluation research. *Education Research International, 1*-16. doi: 10.1155/2012/184640
- Johnson, E. D., Al-Mahmood, R., & Maierb, A. G. (2012). Student and staff perceptions of teamwork in groupwriting for science honours. *International Journal of Innovation in Science and Mathematics Education, 20*(4), 25-41.
- Karthikeyan, R., Venkatesan, K. G. S., & Chandrasekar, A. (2016). A comparison of strengths and weaknesses for analytical hierarchy process. *Journal of Chemical and Pharmaceutical Sciences 9*(3), S-12-S-15.
- Kilic, A. (2010). Learner-centered micro teaching in teacher education. *International Journal of Instruction, 3*(1) 77-100.
- King, F. (2014). Evaluating the impact of teacher professional development: An evidence-based framework. *Professional Development in Education, 40*(1), 89-111.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*(3), 231-251.

- Kupers, E., van Dijk, M., & van Geert, P. (2015). Within-teacher differences in one-to-one teacher-student interactions in instrumental music lessons. *Learning and Individual Differences, 37*, 283-289.
- Kurtz, S., Draper, J., Silverman, J., Draper, J., & Silverman, J. (2017). *Teaching and Learning Communication Skills in Medicine*. CRC Press, 2nd edition, 388 Pages.
- Lans, R. M. van der, Grift, W. J. C. M. van de, & Veen, K. van. (2018). Developing an instrument for teacher feedback: Using the rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education, 86*(2), 247-264.
- Lazarides, R., Viljaranta, J., Aunola, K., & Nurmi, J.-E. (2018). Teacher ability evaluation and changes in elementary student profiles of motivation and performance in mathematics. *Learning and Individual Differences, 67*, 245-258.
- Li, G., Gang KOU, G., and Peng, Y. (2015). Dynamic fuzzy multiple criteria decision making for performance evaluation. *Technological and Economic Development of Economy, 21*(5), 705-719.
- Liu, S., & Teddlie, C. (2007). A follow-up study on teacher evaluation in China: Historical analysis and latest trends. *Journal of Personnel Evaluation in Education, 18*(4), 253-272.
- Liu, S., & Zhao, D. (2013). Teacher evaluation in China: Latest trends and future directions. *Educational Assessment, Evaluation and Accountability, 25*(3), 231-250.
- Malakolunthu, S., & Vasudevan, V. (2012). Teacher evaluation practices in Malaysian primary schools: Issues and challenges. *Asia Pacific Education Review, 13*(3), 449-456.
- Malen, B., Rice, J. K., Matlach, L. K. B., Bowsher, A., Hoyer, K. M., and Hyde, L. H. (2015). Developing organizational capacity for implementing complex education reform initiatives: Insights from a multiyear study of a teacher incentive fund program. *Educational Administration Quarterly, 51*(1), 133-176.
- Malik, M. M., Abdallah, S., and Hussain, M. (2016). Assessing supplier environmental performance: Applying Analytical Hierarchical Process in the United Arab Emirates healthcare chain. *Renewable and Sustainable Energy Reviews, 55*, 1313-1321.
- Maltarich, M. A., Nyberg, A. J., Reilly, G., Abdulsalam, D. "Dee," & Martin, M. (2017). Pay-for-performance, sometimes: An Interdisciplinary Approach to Integrating Economic Rationality with Psychological Emotion to Predict Individual Performance. *Academy of Management Journal, 60*(6), 2155-2174.
- Marzano, R. J., & Toth, M. D. (2013). *Teacher Evaluation that Makes a Difference: A New Model for Teacher Growth and Student Achievement*. Association for Supervision and Curriculum Development, 192 pages.
- Muijs, D., & Reynolds, D. (2017). *Effective Teaching: Evidence and Practice*. SAGE.
- Nahid, B. S., Nasr isfahani, A., Rouhollahi, A., & Khalili, R. (2016). Effective teaching methods in higher education: Requirements and barriers. *Journal of Advances in Medical Education and Professionalism, 4*(4), 170-178.
- Nilson, L. B. (2016). *Teaching at Its Best: A Research-Based Resource for College Instructors*. John Wiley and Sons, 400 pages.
- Odden, A. (2014). Lessons Learned About Standards-Based Teacher Evaluation Systems. *Peabody Journal of Education, 79*(4), 126-137.
- OECD. (2009). *Teacher Evaluation: Current Practices in OECD Countries and a Literature Review*. OECD Education Working papers, 49 pages.
- OECD (Ed.). (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. OECD reviews of evaluation and assessment in education, 670 pages.
- Ovando, M. N. (2001). Teachers' perceptions of a learner-centered teacher evaluation system. *Journal of Personnel Evaluation in Education, 15*, 213-231.
- Parrish, D. R. (2016). Principles and a model for advancing future-oriented and student-focused teaching and learning. *Procedia - Social and Behavioral Sciences, 228*, 311-315.
- Ramot, D., Milo, R., Friedman, M., and Kandel, A. (2002). Complex fuzzy sets. *IEEE Transactions on Fuzzy Systems, 10*(2), 171-186.

- Reddy, L. A., Dudek, C. M., Peters, S., Alperin, A., Kettler, R. J., and Kurz, A. (2018). Teachers' and school administrators' attitudes and beliefs of teacher evaluation: A preliminary investigation of high poverty school districts. *Educational Assessment, Evaluation and Accountability*, 30(1), 47-70.
- Saaty, T. L. (1980). The analytic hierarchy process. *Newyork, NY: McGraw-Hill Inc*, 17-34.
- Saaty, Thomas L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), 83-98.
- Said, B., Ye, J., & Smarandache, F. (2015). An extended TOPSIS method for multiple attribute decision making based on interval neutrosophic uncertain linguistic variables. *Neutrosophic Sets and Systems*, 8, 22-31.
- Schön, D. A. (2017). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, 384 pages.
- Sharp, J. A., Peters, J., Howard, K., Peters, J., and Howard, K. (2017). *The Management of a Student Research Project*. Taylor and Fancis, 278 pages.
- Shingphachanh, S. (2018). Teachers' understanding and concerns about the practices of lesson study in suburb schools in Laos. *International Journal for Lesson and Learning Studies*, 7(2), 150-162.
- Singh, I., & Jha, A. (2014). Difference in Self-reported and Students-rated Teacher Effectiveness among Medical and Engineering Faculty Members: Need for Direct Informal Feedback. *American Journal of Educational Research*, 2(5), 272-277.
- Skedsmo, G., & Huber, S. G. (2018). Teacher evaluation: The need for valid measures and increased teacher involvement. *Educational Assessment, Evaluation and Accountability*, 30(1), 1-5.
- Smarandache, F. (1998). A unifying field in logics. Neutrosophy: Neutrosophic probability, set and logic. *American Research Press, Rehoboth*, 105 pages.
- Sonnert, G., Hazari, Z., & Sadler, P. M. (2018). Evaluating the quality of middle school mathematics teachers, using videos rated by college students. *Studies in Educational Evaluation*, 58, 60-69.
- Steinberg, M. P., and Garrett, R. (2016). Classroom Composition and Measured Teacher Performance: What Do Teacher Observation Scores Really Measure? *Educational Evaluation and Policy Analysis*, 38(2), 293-317.
- Taut, S., Santelices, M. V., Araya, C., & Manzi, J. (2011). Perceived effects and uses of the national teacher evaluation system in Chilean elementary schools. *Studies in Educational Evaluation*, 37(4), 218-229.
- Thomas, S., Chie, Q. T., Abraham, M., Jalarajan Raj, S., & Beh, L.-S. (2014). A Qualitative Review of Literature on Peer Review of Teaching in Higher Education: An Application of the SWOT Framework. *Review of Educational Research*, 84(1), 112-159.
- Tondeur, J., van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Educational Technology Research and Development*, 65(3), 555-575.
- Torkabadi, A. M., Pourjavad, E., & Mayorga, R. V. (2018). An integrated fuzzy MCDM approach to improve sustainable consumption and production trends in supply chain. *Sustainable Production and Consumption*, 16, 99-109.
- Turksen, I. B. (1986). Interval valued fuzzy sets based on normal forms. *Fuzzy Sets and Systems*, 20(2), 191-210.
- Tuytens, M., & Devos, G. (2017). The role of feedback from the school leader during teacher evaluation for teacher and school improvement. *Teachers and Teaching*, 23(1), 6-24.
- Wager, E., & Kleinert, S. (2012). Cooperation between research institutions and journals on research integrity cases: Guidance from the committee on publication ethics. *Saudi Journal of Anaesthesia*, 6(2), 155-160.
- Wang, H., Smarandache, F., Zhang, Y., & Sunderraman, R. (2010). Single valued Neutrosophic sets. *Technical Sciences and Applied Mathematics*, 10-14.
- Wang, J.-W., Cheng, C.-H., & Huang, K.-C. (2009). Fuzzy hierarchical TOPSIS for supplier selection. *Applied Soft Computing*, 9(1), 377-386.

- Wang, T.-Y., & Hsieh, F.-J. (2017). Taiwanese high school students' perspectives on effective mathematics teaching behaviors. *Studies in Educational Evaluation*, 55, 35-45.
- Wang, X., & Chan, H. K. (2013). A hierarchical fuzzy TOPSIS approach to assess improvement areas when implementing green supply chain initiatives. *International Journal of Production Research*, 51(10), 3117-3130.
- Wardil, L., & Hauert, C. (2015). Cooperation and coauthorship in scientific publishing. *Physical Review E*, 91(1), 1-6.
- William, D., Thompson, M., & Thompson, M. (2017). Integrating assessment with learning: what will it take to make it work? In: Dwyer, C A, (ed.) *The Future of Assessment: Shaping Teaching and Learning*. (pp. 53-82). Lawrence Erlbaum Associates: Mahwah, New Jersey.
- Wisker, G. (2012). *The Good Supervisor: Supervising Postgraduate and Undergraduate Research for Doctoral Theses and Dissertations*. Palgrave Macmillan, 400 pages.
- Wolf, L. A. (2016). The peer review process. *Journal of Emergency Nursing*, 42(5), 454-456.
- Wu, H.-Y., Chen, J.-K., Chen, I.-S., and Zhuo, H.-H. (2012). Ranking universities based on performance evaluation by a hybrid MCDM model. *Measurement*, 45(5), 856-880.
- Wu, M.-J., Huang, C.-Y., Kao, Y.-S., Lue, Y.-F., Chen, L.-C. (2018). Developing a professional performance evaluation system for pre-Service automobile repair vocational high school teachers in Taiwan. *Sustainability*, 10(10), 3537.
- Yang, W., & Pang, Y. (2018). New multiple attribute decision making method based on DEMATEL and TOPSIS for multi-valued interval Neutrosophic sets. *Symmetry*, 10(4), 115.
- Ye, J. (2014). Some aggregation operators of interval neutrosophic linguistic numbers for multiple attribute decision making. *Journal of Intelligent and Fuzzy Systems*, 27(5), 2231-2241.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353.
- Zare, M., Pahl, C., Rahnama, H., Nilashi, M., Mardani, A., Ibrahim, O., & Ahmadi, H. (2016). Multi-criteria decision-making approach in E-learning: A systematic review and classification. *Applied Soft Computing*, 45, 108-128.
- Zerem, E. (2017). The ranking of scientists based on scientific publications assessment. *Journal of Biomedical Informatics*, 75, 107-109.
- Zhang, H., Wang, J., & Chen, X. (2014). Interval Neutrosophic sets and their application in multicriteria decision making problems. *The Scientific World Journal*, 2014, 1-15.
- Zhou, H., Liu, Q., Tian, J., & Li, Q. (2018). Rights and interests guarantee of private school teachers. In H. Zhou, Q. Liu, J. Tian, & Q. Li (Eds.), *Private Education in China: Achievement and Challenge*, 169-211.
- Zyad, H. (2016). Integrating Computers in the Classroom: Barriers and Teachers' Attitudes. *International Journal of Instruction*, 9(1), 65-78.

