



# Article AHP, Fuzzy SAW, and Fuzzy WPM for the Evaluation of Cultural Websites

# Katerina Kabassi \*<sup>D</sup>, Christos Karydis and Athanasios Botonis

Department of Environment, Ionian University, Zakynthos, 29100, Greece; c.karydis@ionio.gr (C.K.); nassbotonis@gmail.com (A.B.)

\* Correspondence: kkabassi@ionio.gr

Received: 13 December 2019; Accepted: 18 February 2020; Published: 23 February 2020



**Abstract:** The evaluation of cultural websites is a complicated procedure that depends on several criteria. The main contribution of this paper is on showing how this process may be implemented. For this reason, different multi-criteria decision-making methods are combined and compared for processing the results of the evaluation. More specifically, the paper provides the criteria for evaluating a website of cultural heritage and presents the application of Analytical Hierarchy Process for estimating the weights of the criteria. Then, it compares two different multi-criteria decision-making models, fuzzy simple additive weighting and fuzzy weighted product model, for being combined with Analytical Hierarchy Process and processing the results of the evaluation. The evaluation involves 29 websites of the conservation labs of the international museums and useful conclusions are extracted about the application of all methods and the combination of Analytical Hierarchy Process with fuzzy simple additive weighting and fuzzy weighted product model.

Keywords: virtual reality; museum tours; evaluation; multi-criteria decision making; fuzzy logic

# 1. Introduction

Websites play an important role in promoting culture to everyone. For this purpose, cultural websites are targeted to users of different background knowledge and characteristics. The great variety of the characteristics of the users interacting with the websites of cultural content, results in not being able to address the needs of all users and as a result most users encounter problems during their interaction with such websites. Therefore, the need for evaluating such websites is great. Indeed, Jesús et al. [1], after the evaluation of two museum websites, pointed out that the usability, the content, and other characteristics of the website positively and significantly influence users' intentions of revisiting it and visiting the museum physically. The problem is even greater when one refers to more specialized parts of the museum. Such a case is the electronic presence of a conservation lab.

A conservation lab's main functions are to receive art, provide a workspace for the conservation of art, maintain safety standards, and provide for the storage of conservation records. Different specializations will have different requirements for particular equipment in their labs including architecture, book and paper, electronic media, objects, paintings, photographic materials, scientific research, textiles, and wooden artifacts. However, even if the conservation labs in museums play an important role in the preservation of the collections and their artifacts, mainly this role is sometimes overlooked. The museums' conservation labs and the treatments which are carried out on the artifacts, many times are not obvious for the public. Nevertheless, their content may be of interest to students, researchers, archaeologists, tourists, and artists for further education and preservation guidelines purposes.

Many researchers have highlighted the need for evaluating e-museum websites in order to confirm that websites are actually accessible and usable [2,3]. For this purpose, different models for evaluation

of cultural websites have been proposed. For example, Di Blas et al. [4] have developed MilanoLugano Evaluation Method (MiLE), a model for weighting attributes in a usability evaluation. However, they did not use a specific decision-making model but only weighting between pairs of attributes. Similarly, models such as MuseumQual [5] and Museum's Sites Evaluation Framework (MUSEF) [6] use multiple criteria but not a specific theory for their model. Indeed, the majority of the models for website evaluation do not use a theory for combining the evidence collected during the experiment. Therefore, the main scope of this paper is to present a model for the evaluation of the websites of cultural content. As a testbed we use the 29 websites of the museums' conservation labs that had electronic presence on the internet. The specific websites have been selected because they contain both specialized and non-specialized content and, therefore, their content aims at targeting professionals, tourists, and all different kinds of users.

Taking into account that the evaluation of a website is a multi-dimensional problem that is based on several criteria and the fact that there is a lack of evaluation models for cultural websites that use theories, we consider multi-criteria decision making (MCDM) methods in a model for evaluating websites of cultural content. The Analytic Hierarchy Process (AHP) [7] is really strong in making decisions by making pairwise comparisons of uncertain, qualitative, and quantitative factors. The method provides a well-designed procedure for calculating the weights of the criteria used in the reasoning process of experts [8]. However, the implementation of AHP is getting complicated as the number of alternatives increases. Therefore, in many cases [9,10], AHP is combined with other methods.

Taking into account the above, we have made two different combinations of methods. The first combination involved AHP and fuzzy simple additive weighting (SAW) [11]. The second combination involved AHP and fuzzy weighted product model (WPM) [12,13]. Both these combinations have been used in the implementation of the evaluation results and the processing of its results. Fuzzy SAW (FSAW) and fuzzy WPM (FWPM) are simple multi-criteria decision-making methods, meaning that their implementation does not get more complicated with the increase of the alternatives. As a result, AHP has been used for the estimation of the weights of the criteria. Then the main evaluation experiment took place with the participation of 81 users who visited and evaluated the 29 websites using linguistic terms. The results collected were in turn processed with fuzzy SAW and fuzzy WPM in order to compare them and extract useful conclusions about their application and their combination with AHP.

## 2. Research Aim

Kabassi [14] presented a review of the evaluation experiments of museums' websites, gathered the criteria used in different expert-based or empirical evaluation experiments and created the final set of criteria used for the evaluation of websites. In this review, the fuzzy theory is only used in the Fuzzy Quality Tree for Web Inspection (FQT4Web) methodology [15] in combination with a hierarchical tree. We use fuzzy theory in combination with two multi-criteria decision-making theories in a model for cultural websites' evaluation.

Our research described in the present paper is based on previous research on the evaluation of websites of cultural content presented in [9,14,16]. Kabassi et al. [17] present the evaluation of museums' conservation labs using eight criteria and the WPM model. However, the set of criteria was considered limited and a better set may have given better results. Furthermore, in those evaluation experiments, the fuzzy theory was not employed and giving crisp values to the criteria was considered hard by the evaluators. In Kabassi [9], on the other hand, a framework has been proposed for the evaluation of museums' websites using AHP and a technique for the order of preference by similarity to an ideal solution [18] using the set of criteria proposed in [14].

The application of multi-criteria decision-making theories for the evaluation of museums' websites revealed a greater role that decision making can play in evaluation experiments. At the same time, important questions were raised: (1) multi–criteria decision making can be used in evaluation experiments of websites with cultural content in general? (2) can the set of criteria be improved?

(3) can fuzzy multi-criteria decision-making methods be successfully used? (4) which method gives better results?

In an effort to address the above questions, in this paper, we present the evaluation experiment of twenty-nine websites of conservation labs and try to explore the suitability of the set of criteria proposed in [14] for the evaluation of different websites of cultural content. Furthermore, we use fuzzy multi-criteria decision making in the evaluation process and check the effectiveness of such methods by combining them with AHP. Finally, we compare two different fuzzy multi-criteria decision-making methods by applying them in a similar evaluation experiment and compare their application and results. These combinations of methods have never been compared before and have not been used for evaluating websites of cultural content.

The paper is organized as follows: Section 3 presents the material and methods used in this paper. More specifically, it presents the criteria used for the evaluation of the websites, as well as the multi-criteria decision-making methods that are used to combine these criteria and evaluate the websites. In Section 3, we present the experiment design, which includes first the estimation of the weights using AHP and then both fuzzy SAW and fuzzy WPM for the processing of the data collected during the evaluation. The extracted results are presented in Section 5. In Section 6, we present a comparison of the two methods. The results of the evaluation and the conclusions drawn by this work are discussed and analyzed in Section 7 and we give the conclusions drawn in Section 8.

#### 3. Materials and Methods

In this section, we present the materials and methods used in the implementation of the proposed model for the evaluation of the websites of conservation labs in museums. For the evaluation of websites of museums' conservation labs we have used the set of criteria proposed in [14] for the expert-based evaluation experiments.

1. Category 1: Content. In this category, all criteria are related to the content of a website.

- i. **c11: Currency/clarity/text comprehension.** This criterion checks the currency and the clarity of the text. Currency refers to how successful is the system in providing up-to-date information, and how successfully it can reflect the current state of the world that it represents. Clarity refers to how comprehensible the texts provided to the users are. For this purpose, the quality and the style are checked as well as the way the content is organized and designed in order to make the website credible and trustworthy.
- ii. **c12: Completeness/richness.** This criterion checks whether a website has adequate information on the subject.
- iii. c13: Quality content. This criterion involves the accuracy and understandability of content.
- iv. **c14: Support of research.** Checks whether the website provides information for the support of research.
- 2. Category 2: Usability. All the criteria that are related to Usability.
- i. **c21: Consistency.** Consistency means that similar pieces of information are dealt with in similar fashions [4].
- ii. **c22:** Accessibility. Accessibility measures how easily and intuitively accessible is the website's information for any user.
- iii. **c23: Structure/navigation.** The structure of the information provided plays an important role in the success of a website. Therefore, the organization of the content pieces should be in such a way that the navigation of the user to the content of the website is easy.
- iv. c24: Easy to use/simplicity. The user interface should be simple and easy to use.
- v. **c25: User interface/overall presentation/design.** This criterion checks whether the overall presentation is attractive and engaging.

- vi. **c26: Efficiency.** This criterion shows whether actions within the website can be performed successfully and quickly [4].
- 3. Category 3: Functionality. Criteria that are related to the functionality of the website.
- i. **c31:** Multilingualism. The information should be given in more than one language [4].
- ii. c32: Multimedia. Different media should be used to convey the information [4].
- iii. **c33: Interactivity.** This criterion checks whether the content of the website is comprehensive and useful, nicely presented, and easy to explore and use.
- iv. **c34:** Adaptivity. Adaptivity is the ability of the system to adapt to users' characteristics such as needs and interests while adaptability refers to the ability of users to adapt the user interface to their own preferences.

# 3.1. AHP

The steps of the implementation of AHP that were used for the estimation of the weights of the criteria were the following:

- 1. Form the set of evaluators. For the estimation of the weights of the criteria, it is important to have the view of experts on the field. Therefore, evaluators are only human experts. The selection of expert-based evaluations has many advantages [19] and the correct choice of the expert would give reliable and valid results.
- 2. Setting up a pairwise comparison matrix of criteria. In this step, a comparison matrix is formed so that the heuristics are pairwise compared. More specifically, a V from the scale that is presented in Table 1 is assigned to the comparison result of two elements P and Q at first, then the value of comparison of Q and P is a reciprocal value of V, i.e., 1/V. The value of the comparison between P and P is one (see Table 1).
- 3. **Calculating weights of criteria:** After making pairwise comparisons, estimations are made that result in the final set of weights of the criteria.

	Р	Q
Р	1	V
Q	1/V	1

**Table 1.** Matrix for the pairwise comparison of criteria.

As soon as, the weights have been calculated, the websites of conservation labs in museums are evaluated and the analysis of the results of the evaluation is made based on fuzzy SAW or fuzzy WPM.

# 3.2. Fuzzy SAW and Fuzzy WPM

Fuzzy SAW and fuzzy WPM have many steps in common and mainly differ in the pre-last step of their implementation. For this purpose, we give the steps that are common and those that differentiate in the implementation of these two methods. Both methods presuppose that the weights of the criteria have been estimated and use them to evaluate the alternative website of conservation labs in museums. The steps of fuzzy SAW and fuzzy WPM are the following:

- 1. **Forming a new set of evaluators.** In this phase of the evaluation experiment, the set of evaluators were formed, following the taxonomy of types of users of cultural websites proposed by Sweetnam et al. [20]. This group may be the same as the one formed in step 1 of the application of AHP or may be different.
- 2. Assigning values to the criteria. In order to make this process easier for the users, especially for those that do not have experience in multi-criteria analysis, the users could use linguistic terms for characterizing the fourteen criteria presented at the beginning of this section. Therefore,

evaluators use linguistic terms to give values to the criteria. The linguistic terms used in this method have been proposed by [21]. All criteria are assigned fuzzy values. The linguistic terms and the corresponding fuzzy values, which are used for rating all criteria, are presented in Table 2.

- 3. **Linguistic terms are transformed into fuzzy numbers.** Each linguistic term is assigned to a fuzzy number, which is a vector  $\tilde{a} = (a_1, a_2, a_3)$ . The matches are presented in Table 2 [21].
- 4. **Construction of the MCDM matrix.** A fuzzy multi-criteria decision-making problem can be expressed in a matrix format. Each element of the matrix is a fuzzy number. However, in order to aggregate all the values of the decision-makers in one single value, the geometric mean is used. The geometric mean of two fuzzy numbers  $\tilde{a} = (a_1, a_2, a_3)$  and  $\tilde{b} = (b_1, b_2, b_3)$  is calculated as follows:

$$\tilde{D} = \begin{array}{cccc}
 & C_1 & C_2 & C_n \\
\tilde{A}_1 & \begin{pmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \tilde{x}_{1n} \\
\tilde{x}_{21} & \tilde{x}_{22} & \tilde{x}_{2n} \\
 & & \tilde{x}_{ij} \\
\tilde{x}_{m1} & \tilde{x}_{m2} & \tilde{x}_{mn} \\
\end{array}\right), i = 1, 2, \dots, m; j = 1, 2, \dots, n, \ \tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij}) \quad (1)$$

where *i* shows the alternative and *j* shows the criterion. Each  $\tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$  is a triangular fuzzy number.

# 5. Implement Fuzzy SAW or Fuzzy WPM

- 5.1 SAW.
  - 5.1.i The normalization of fuzzy numbers. To avoid the complicated normalization formula normally used, Chen [21] proposes a linear scale transformation in order to transform the various criteria scales into a comparable scale. The particular normalization method aims at preserving the property that the ranges of normalized triangular fuzzy numbers belong to [0,1]. The normalization of a fuzzy number  $\hat{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$  is given by the formula:

$$\tilde{r}_{ij} = (\frac{a_{ij}}{c_i^*}, \frac{b_{ij}}{c_i^*}, \frac{c_{ij}}{c_i^*}),$$
(2)

where  $c_{i}^{*}$  is the maximum value of all  $c_{ij}$  (3).

5.1.ii Calculating the weighted normalized fuzzy numbers of the MCDM matrix. Derive total fuzzy scores for individual alternatives by multiplying the fuzzy rating matrix by their respective weight vectors. Obtained total fuzzy score vector by multiplying the fuzzy rating matrix  $\tilde{D}$  by the corresponding weight vector W, i.e.,

$$\bar{F}_{SAW} = \bar{D} \otimes W^{T} = \begin{bmatrix} \bar{r}_{11} & \bar{r}_{12} & \dots & \bar{r}_{1n} \\ \bar{r}_{21} & \bar{r}_{22} & \dots & \bar{r}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots \\ \bar{r}_{m1} & \bar{r}_{m2} & \dots & \bar{r}_{mn} \end{bmatrix} \otimes \begin{bmatrix} W_{1} \\ W_{2} \\ \vdots \\ \vdots \\ W_{n} \end{bmatrix} = \begin{bmatrix} \bar{r}_{11} \otimes W_{1} \oplus \bar{r}_{12} \otimes W_{2} \oplus \dots \bar{r}_{1n} \otimes W_{n} \\ \bar{r}_{21} \otimes W_{1} \oplus \bar{r}_{22} \otimes W_{2} \oplus \dots \bar{r}_{2n} \otimes W_{n} \\ \vdots \\ \vdots \\ \bar{r}_{m1} \otimes W_{1} \oplus \bar{r}_{m2} \otimes W_{2} \oplus \dots \bar{r}_{mn} \otimes W_{n} \end{bmatrix} = \begin{bmatrix} \bar{f}_{1} \\ \bar{f}_{2} \\ \vdots \\ \vdots \\ \bar{r}_{m1} \otimes W_{1} \oplus \bar{r}_{m2} \otimes W_{2} \oplus \dots \bar{r}_{mn} \otimes W_{n} \end{bmatrix} = \begin{bmatrix} \bar{f}_{1} \\ \bar{f}_{2} \\ \vdots \\ \vdots \\ \bar{f}_{n} \end{bmatrix}$$
(3)

where  $\tilde{f}_i = (a_i, b_i, c_i), i = 1, 2, ..., m$ .

5.2 WPM. Calculating the weighted normalized fuzzy numbers of the alternatives. The classic WPM compares alternatives in pairs by calculating a ratio  $P(A_K/A_L)$ . If this ratio is greater than or equal to the value one, then it indicates that alternative  $A_K$  is preferred than  $A_L$ . However, when the alternatives are too many these pair-wise comparisons are complicated

and time-consuming. Therefore, we use an alternative application of WPM, which is proposed by Triantafyllou [12]. In this alternative approach of WPM, the decision-maker use only products without the previous ratios. Therefore, for each alternative the following value is calculated:

$$\tilde{f}_{i} = \prod_{j=1}^{n} (\tilde{x}_{ij})^{w_{j}}, i = 1, 2, \dots, m; j = 1, 2, \dots, n$$
(4)

The term  $\tilde{f}_i$  denotes the total performance value of the alternative  $A_i$ .  $\tilde{x}_{ij}$  is a triangular number  $\tilde{f}_i = \prod_{j=1}^n ((a_i, b_i, c_i)^{w_j})$ . These values are used for the formation of the final fuzzy rating matrix:

$$\tilde{F}_{WPM} = \begin{bmatrix} \tilde{f_1} \\ \tilde{f_2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \tilde{f_n} \end{bmatrix}$$

6. Compute a crisp value. A value is calculated for alternative website using a defuzzification method and select the alternative(s) with the maximum total score. Rank total fuzzy scores *f*<sub>1</sub>, *f*<sub>2</sub>, ..., *f*<sub>m</sub> by the signed distance to determine the best location. Four defuzzification methods are most commonly used: The centroid method, mean of maximal (MOM), a-cut method, and signed distance method [11,22–26]. All these methods share advantages and disadvantages [27], but Yao and Chiang [28] propose the signed distance method, which is also used by Chou et al. [11] in fuzzy SAW. The crisp total scores of individual locations are calculated by the following defuzzification equation:

$$d(\tilde{f}_i) = \frac{1}{3}(a_i + b_i + c_i), \ i = 1, 2, \dots, m$$
(5)

where  $d(\tilde{f}_i)$  gives the defuzzified value (crisp value) of the total fuzzy score of the alternative  $A_i$ . The ranking of the websites can then be based on the above crisp value of the total scores for individual alternative websites.

Linguistic Term	Fuzzy Number
Very Poor	(0,0,1)
Poor	(0,1,3)
Fair	(3,5,7)
Good	(7,9,10)
Very Good	(9,10,10)

Table 2. Linguistic variables and fuzzy numbers.

#### 4. Experiment Design

For the implementation of the experiment, we conducted a search on the websites of conservation labs in museums all over the world. This search revealed that only a few museums provide adequate information about their conservation labs and have separate sections on their websites or a completely new website. This process revealed twenty-nine museum conservation labs websites that belong to the following international and national museums: (1) Archaeological Museum of Thessaloniki, (2) Australian Museum, (3) Barberini – Corsini Gallery – Roma, (4) Benaki Museum, (5) Boston Museum of Fine Arts, (6) British Museum, (7) Brooklyn museum, (8) Byzantine & Christian Museum

in Athens, (9) De Young Museum of Fine Arts, (10) Galleria Nazionale d'Arte Moderna, (11) Getty Institution, (12) Guggenheim Museum, (13) Hermitage Museum, (14) Metropolitan Museum, (15) MoMa, (16) Museo Del Prado, (17) Museum of Byzantine Culture in Thessaloniki, (18) Museum of Islamic Art – Doha, (19) National Gallery of Greece, (20) National Museum New Delhi, (21) NTNU University museum, (22) Oriental Institute Museum, (23) Rijksmuseum, (24) Smithsonian museum, (25) Tate Modern, (26) Tokyo National Museum, (27) University of Michigan Museum of Art, (28) Vatican Museum, and (29) Victoria & Albert Museum.

## 4.1. AHP

As soon as the set of the websites of conservation labs of museums is formed, AHP is applied to calculate the weights of the criteria used in the reasoning process of the decision-makers. For the application of AHP, the following steps are implemented:

- 1. The set of evaluators was formed. The group of evaluators participated four professional conservators and four software engineers, three of which had experience in a University Department of Conservation of Antiquities & Works of Art. 87.5% of the experts were Greek and the rest of them, English. 75% of the evaluators were 35–45 years old and the others belonged to the age group 45–55. Finally, 62.5% of the experts were male.
- 2. A pairwise comparison matrix of criteria was set up. Each one of the eight experts were asked to fill the pairwise comparison matrix so that the criteria would be pairwise compared. As a result, eight matrices were completed. In order to find the final pairwise comparison matrix of criteria, we calculated the geometric mean of the values of all corresponding cells belonging to the eight matrixes completed by the experts.
- 3. The weights of the criteria were calculated. In this step, the principal eigenvalue and the corresponding normalized right eigenvector of the comparison matrix gave the relative importance of the various criteria being compared. The elements of the normalized eigenvector were the weights of criteria. In terms of simplicity, we had used the Priority Estimation Tool (PriEst) (Sirah et al. 2015), an open-source decision-making software that implements the AHP method, for making the calculations of AHP. The weights of the criteria were estimated as follows:  $w_{c1} = 0.292$ ,  $w_{c2} = 0.534$ ,  $w_{c3} = 0.174$ ,  $w_{c11} = 0.34$ ,  $w_{c12} = 0.186$ ,  $w_{c13} = 0.325$ ,  $w_{c14} = 0.149$ ,  $w_{c21} = 0.172$ ,  $w_{c22} = 0.15$ ,  $w_{c23} = 0.214$ ,  $w_{c24} = 0.213$ ,  $w_{c25} = 0.164$ ,  $w_{c26} = 0.088$ ,  $w_{c31} = 0.242$ ,  $w_{c32} = 0.315$ ,  $w_{c33} = 0.196$ ,  $w_{c3} = 0.247$ .

The next steps of AHP would require the completion of eight 29 by 29 matrixes. This considered very complicated and hard to implement even for the experts on the field. Another difficulty that cultural experts often encounter when evaluating alternative websites, is quantifying criteria. For this purpose, we have used alternatively, fuzzy SAW and fuzzy WPM to evaluate the websites of museums' conservation labs. Then, we compared these two fuzzy MCDM models to see which one seemed more appropriate for completing AHP.

## 4.2. Fuzzy SAW vs. Fuzzy WPM

In this part of the evaluation, experts and non-experts were selected to participate in the experiment. More specifically, the new group of the evaluator consisted of 81 users of different knowledge and levels of expertise in conservation and software engineering. They were asked to visit the 29 websites of conservation labs of museums and answer a questionnaire with multiple-choice questions. Each question corresponded to the evaluation of each criterion. This process resulted in one decision matrix with the linguistic terms for the 81 users. The linguistic terms in the 81 matrices created by the evaluators were then replaced with fuzzy numbers. The fuzzy numbers that correspond to the linguistic terms are presented in Table 2. In order to aggregate all the values of the 81 matrices created by the decision-makers in one single matrix, the geometric mean was used. This process resulted in one single decision matrix:

$$\tilde{D} = \begin{array}{ccc} & C_{1} & C_{2} & C_{14} \\ \tilde{D} = \begin{array}{ccc} A_{1} & \left( \begin{array}{ccc} \tilde{x}_{11} & \tilde{x}_{12} & \tilde{x}_{114} \\ \tilde{x}_{21} & \tilde{x}_{22} & \tilde{x}_{214} \\ & & \tilde{x}_{ij} \\ & & \tilde{x}_{291} & \tilde{x}_{292} & \tilde{x}_{2914} \end{array} \right), \ i = 1.2, \dots, 29; \ j = 1, 2, \dots, 14, \ \tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij}) \quad (6)$$

where *i* shows the alternative museum and *j* shows one of the 14 criteria. Each element of the matrix is a triangular fuzzy number. In our evaluation  $\tilde{x}_{21} = (4.92, 6.97, 8.64)$ .

As far as the application of SAW is concerned, the values of the criteria were transformed into a comparable scale using the formula:  $\tilde{r}_{ij} = (\frac{a_{ij}}{c_j^*}, \frac{b_{ij}}{c_j^*}, \frac{c_{ij}}{c_j^*})$ , where  $c_j^*$  is the maximum value of all  $c_{ij}$ . As a result, all triangular fuzzy numbers that existed in the decision matrix belonged to [0,1]. In our evaluation, for example,  $\tilde{r}_{21} = (0.55, 0.78, 0.96)$ . These normalized values were further used for the application of fuzzy SAW.

Taking into account that criteria have different weights in the reasoning process of the decision-makers, we calculated the values of each criterion with the corresponding weight of the criterion. Then these values were combined together and as a result, a fuzzy number was estimated for each alternative website. Finally, the weighted normalized fuzzy MCDM matrix  $\tilde{F}_{SAW} = [\tilde{f_i}], i = 1, 2, ..., 29$  was constructed.

As far as the application of WPM is concerned, numbers were not normalized. Each criterion value was power raised to the corresponding weight. Then these values were multiplied according to the model of WPM using Equation (5). This resulted in the construction of the fuzzy MCDM matrix  $\tilde{F}_{WPM} = [\tilde{f}_i], i = 1, 2, ..., 29$ .

Tables  $\tilde{F}_{SAW}$  and  $\tilde{F}_{WPM}$  contain the results of the application of fuzzy SAW and fuzzy WPM. In these tables, the element  $\tilde{f}_i$  was estimated either with SAW in the first matrix or with WPM in the second matrix.

	1.39	2.01	2.66	]	2.02		55.02	166.98	427.69	]	216.56	]
	1.51	2.20	2.79		2.17		74.48	239.50	507.86		273.95	
	1.60	2.26	2986		2.24		91.53	264.30	543.15		299.66	
	1.66	2.33	2.83		2.27		97.43	283.38	524.36		301.72	
	1.03	1.60	2.29		1.64		20.78	77.50	265.79		121.36	
	1.34	1.94	2.63		1.97		51.50	165.62	421.45		212.86	
	1.63	2.26	2.80		2.23		86.00	255.48	487.97		266.48	
	1.27	1.90	2.52		1.90		40.12	133.76	358.27		177.39	
	1.12	1.67	2.38		1.72		27.36	85.55	296.78		136.56	
	1.08	1.70	2.41		1.73		23.65	100.35	314.50		146.17	
	1.36	1.98	2.61		1.98		48.98	151.31	394.09		198.13	
	1.07	1.66	2.37		1.70		22.17	86.62	293.51		134.10	
	1.12	1.72	2.42		1.75		27.24	98.72	316.61		147.52	
	1.32	1.98	2.60		1.97		43.15	151.33	390.64		195.04	
$\tilde{F}_{SAW} =$	1.15	1.73	2.44	$F_{SAW} =$	1.78	$\tilde{F}_{WPM} =$	30.51	101.57	324.79	$F_{WPM} =$	152.29	
	1.31	1.96	2.62		1.96		44.67	149.74	402.75		199.05	
	1.32	2.01	2.64		1.99		50.83	184.74	430.18		221.92	
	1.24	1.84	2.53		1.87		33.95	117.40	354.50		168.62	
	1.24	1.89	2.49		1.87		37.08	130.74	346.01		171.28	
	0.79	1.25	2.00		1.35		8.92	34.04	171.00		71.32	
	1.39	2.06	2.69		2.05		53.16	191.00	445.66		229.94	
	1.56	2.23	2.80		2.20		81.17	245.99	511.74		279.63	
	0.95	1.40	2.16		1.50		15.97	51.01	219.20		95.39	
	1.19	1.81	2.47		1.82		32.90	117.67	336.42		162.33	
	1.38	2.06	2.70		2.05		54.81	192.43	456.51		234.58	
	0.64	1.15	1.89		1.23		5.30	28.85	151.14		61.76	
	0.79	1.32	2.07		1.39		9.63	45.67	197.00		84.10	
	1.07	1.78	2.47		1.77		25.46	127.15	349.93		164.52	
	1.25	1.95	2.58		1.93		41.94	167.47	399.30		202.90	

# 5. Results

Tables  $\tilde{F}_{SAW} = [\tilde{f}_i], i = 1, 2, ..., 29$  and  $\tilde{F}_{WPM} = [\tilde{f}_i], i = 1, 2, ..., 29$  contain triangular numbers. In order to calculate a crisp value for each alternative website, we used Equation (6) as a defuzzification method. This resulted in a crisp value for each alternative website. Then we ranked the websites taking first the one with the greater value. This process was done for the values estimated by SAW and WPM. The results are presented in tables  $F_{SAW} = [f_i], i = 1, 2, ..., 29$  and  $F_{WPM} = [f_i], i = 1, 2, ..., 29$  that presented above. The final ranking of the websites based on the values of the websites presented in tables  $F_{SAW}$  and  $F_{WPM}$  (Tables 3 and 4).

**Table 3.** The final ranking of the websites based on the values calculated by simple additive weighting (SAW) and presented in  $F_{SAW}$ .

1	National Gallery of Greece	2.27
2	Benaki Museum	2.24
3	Metropolitan Museum	2.23
4	Hermitage Museum	2.20
5	Byzantine & Christian Museum in Athens	2.17
6	Museo Del Prado	2.05

(7)

7	Vatican Museum	2.05
8	Archaeological Museum of Thessaloniki	2.02
9	Victoria & Albert Museum	1.99
10	Boston Museum of Fine Arts	1.98
11	Guggenheim Museum	1.97
12	MoMa	1.97
13	De Young Museum of Fine Arts	1.96
14	Tokyo National Museum	1.93
15	Smithsonian museum	1.90
16	British Museum	1.87
17	Tate Modern	1.87
18	Australian Museum	1.82
19	Brooklyn museum	1.78
20	Rijksmuseum	1.77
21	Oriental Institute Museum	1.75
22	NTNU university museum	1.73
23	GETTY	1.72
24	University of Michigan Museum of Art	1.70
25	Museum of Byzantine Culture in Thessaloniki	1.64
26	Museum of Islamic Art - Doha	1.50
27	Barberini – Corsini Gallery – Roma	1.39
28	National Museum New Delhi	1.35
29	Galleria Nazionale d'Arte Moderna	1.23

Table 3. Cont.

**Table 4.** The final ranking of the websites based on the values calculated by weighted product model (WPM) and presented in  $F_{WPM}$ .

1	National Gallery of Greece	301.72
2	Benaki Museum	299.66
3	Hermitage Museum	279.63
4	Byzantine & Christian Museum in Athens	273.95
5	Metropolitan Museum	266.48
6	Museo Del Prado	234.58
7	Vatican Museum	229.94
8	Victoria & Albert Museum	221.92
9	Archaeological Museum of Thessaloniki	216.56
10	Guggenheim Museum	212.86
11	Tokyo National Museum	202.90
12	De Young Museum of Fine Arts	199.05
13	Boston Museum of Fine Arts	198.13
14	MoMa	195.04
15	Smithsonian museum	177.39

16	British Museum	171.28
17	Tate Modern	168.62
18	Rijksmuseum	167.52
19	Australian Museum	162.33
20	Brooklyn museum	152.29
21	Oriental Institute Museum	147.52
22	NTNU university museum	146.17
23	GETTY	136.56
24	University of Michigan Museum of Art	134.10
25	Museum of Byzantine Culture in Thessaloniki	121.36
26	Museum of Islamic Art - Doha	95.39
27	Barberini – Corsini Gallery – Roma	84.10
28	National Museum New Delhi	71.32
29	Galleria Nazionale d'Arte Moderna	61.76

#### Table 4. Cont.

### 6. Comparison of Methods

In order to compare the two methods presented in this paper, we conducted an experiment using AHP. The validation process aimed at comparing the two different combinations AHP-FSAW and AHP-FWPM. The main reason for selecting AHP for comparing these two combinations is the AHP method supports comparisons in pairs. The criteria used for comparing these two proposed methods are acquired by [29,30] and adjusted to our case:

- Completeness: This criterion shows whether the framework is complete.
- Accuracy: This criterion shows whether the method's processes are accurate.
- Efficiency: The efficiency is calculated by taking into account the positive comments of users and the total number of comments about the framework.
- Satisfaction-method: Reveals the satisfaction with the method.
- Satisfaction-result: Reveals the satisfaction with the result achieved.

For the implementation of the validation process, we used a group of evaluators that consisted of nine expert users (two web designers, one software engineer, three curators, and two archaeologists). The values of Table 5 have resulted after finding the geometric mean of all values of the nine tables collected by the nine evaluators.

	Completeness	Accuracy	Efficiency	Satisfaction- Method	Satisfaction- Result
Completeness	1.00	0.49	0.44	0.34	0.36
Accuracy	2.03	1.00	0.60	0.31	2.03
Efficiency	2.29	1.66	1.00	0.31	1.61
Satisfaction-method	2.92	3.19	3.19	1.00	1.88
Satisfaction- result	2.79	0.49	0.62	0.53	1.00

Table 5. Pairwise comparison of the criteria for the comparison of the two methods.

In order to calculate the weights of the criteria, we have used the PriEst and were the following  $w_{completeness} = 0.08$ ,  $w_{accuracy} = 0.08$ ,  $w_{efficiency} = 0.08$ ,  $w_{satisfaction-method} = 0.08$ ,  $w_{satisfaction-result} = 0.15$ .

After completing the process of evaluating the criteria, the evaluators participated in the second phase of the validation and comparison of the two methods: AHP-FSAW and AHP-FWPM. First, the implementation of the two methods was presented to the nine experts and then the results of each method. The evaluators were then asked to complete pairwise comparison two by two tables for each criterion. The values of the tables were processed by the PriEst and calculated a value for each method:  $AHP_{AHP-FSAW} = 0.631$ ,  $AHP_{AHP-FWPM} = 0.369$ .

Taking into account these AHP values, the method AHP-FSAW was considered better than the AHP-FWPM method. The results may not be very different, but the human experts thought that the results produced by the AHP-FSAW were three times better than the results of AHP-FWPM. The two methods did not differentiate in completeness and efficiency but AHP-FSAW was also considered slightly better in accuracy and in the satisfaction of the technique.

#### 7. Discussion

AHP has the ability to model expert opinion and, therefore, is considered ideal for the calculation of the weights of the criteria. However, AHP is a time-consuming technique because of the mathematical calculations and the number of pairwise comparisons. The complexity increases as the number of alternatives and criteria increases or changes [31]. Since complexity rises with the increase in websites, the number of alternatives that can be compared is limited. This is one of the main reasons for selecting to combine AHP with another method.

For this reason, AHP was combined in turn with fuzzy SAW and fuzzy WPM. The complexity of these methods does not increase with the rate of AHP when the number of alternative websites increases. Both fuzzy SAW and fuzzy WPM use linguistic terms, which are easier for users to comprehend and, therefore, the implementation of the methods is considered easier. Furthermore, in the case of an evaluation, in which several evaluators are involved that do not have experience in implementing, fuzzy SAW or fuzzy WPM seems more appropriate. However, fuzzy SAW and fuzzy WPM does not provide a specified way for calculating the weights of criteria as AHP does. Taking into account all this information, AHP can successfully be combined with fuzzy SAW or fuzzy WPM.

The results of the first part of the evaluation revealed that the most important criterion of the first level is usability, followed by content. Within the sub-criteria of content, the quality of content was considered the most important criterion. Regarding usability, the sub-criteria structure/navigation and easy to use/simplicity are considered almost equally important. As far as functionality is concerned, the existence of multimedia is considered the most important criterion.

Regarding the results of the websites' evaluation, fuzzy SAW and fuzzy WPM had similar results but not identical. Both methods revealed that the best website was considered to be the website of the conservation lab of the National Gallery of Greece. The particular website provided rich content related to the activities of the department, the different departments, the equipment, and the staff. Its content is enriched with multimedia. The user interface is well designed and generally, the website is well structured and usable. The website of the Benaki Museum in Athens was rated as second by both methods. However, one may be concerned with the fact that two Greek websites were rated first. Although the language is a factor that may have influenced the evaluators, one can also observe that other Greek sites have been ranked in the last five.

The methods also gave the same rating in the last nine websites. Two of the last ranked websites of museums' conservation labs are the websites of the National Museum of New Delhi and the Galleria Nazionale d'Arte Moderna. Their content was poor, and it did not provide information about the staff, the facilities, and the equipment of the conservation lab. Furthermore, the websites had only a few photos and no other multimedia that could improve the interaction with the user. Finally, both websites did not appear to be updated until the time of the evaluation.

The other websites received different ratings by fuzzy SAW and fuzzy WPM but their position in the final ranking did not vary considerably. For example, the websites of the conservation labs of the Hermitage Museum and the Byzantine & Christian Museum in Athens were rated fourth and fifth according to fuzzy SAW and fifth and sixth according to fuzzy WPM. Similarly, the Archaeological Museum of Thessaloniki was rated eighth using fuzzy SAW and ninth using fuzzy WPM.

The comparison, however, of the two methods revealed that the combination of AHP with SAW provided the best final ordered list of the websites that were evaluated, and the human experts preferred this combination taking into account its accuracy and the implementation of the technique.

## 8. Conclusions

The evaluation of websites of cultural content that are targeted to a variety of users [20,32,33] is a very complicated procedure that focuses on the examination of several different criteria. The main contribution of this paper is on showing how this process may be implemented and comparing different methods for processing the results of the evaluation. More specifically, the paper provides the criteria for evaluating a website of cultural heritage and presents the application of AHP for estimating the weights of the criteria. Then it compares two different multi-criteria decision-making models, fuzzy SAW and fuzzy WPM, for being combined with AHP and processing the results of the evaluation.

AHP proved very effective in estimating the weights of criteria as it provided a well-defined procedure. Then AHP was combined with fuzzy SAW and provided a view for the electronic presence of the websites of the museums' conservation labs. In an effort to compare fuzzy SAW with fuzzy WPM, we also combined AHP with fuzzy WPM. The combination of these methods provided a different view of the electronic presence and this view compared with the one produced with the application of fuzzy SAW. Both of these methods can successfully process the results of the main evaluation experiment and be combined with AHP to provide better results. Additionally, these methods use linguistic terms which makes the evaluation process easier for both expert and non-expert users.

Researchers of the field of the evaluation of cultural heritage (e.g., [19]) believe that the reliability of the results of an evaluation experiment mainly depends on the expertise of the evaluators and propose the use of a double expert system (with both usability and domain experts) to increase the reliability of the results. This latter suggestion is taken into account when forming the group of human experts presented in this paper in order to ensure the reliability of the evaluation's results.

Taking into account the results of the evaluation, both fuzzy SAW and fuzzy WPM provided a similar view of the electronic image of the conservation labs of museums. However, we also performed a comparison test of the two methods using a different group of the human experts of the initial evaluation experiment. Nevertheless, some of the experts participated in both groups of evaluators. The comparison of the two methods gave precedence in the AHP-SAW combination with regard to the results, the accuracy, and the implementation of the method.

This combination could also be successfully be easily used in a different group of websites of cultural content, but it could be also used in the evaluation of the website of a completely different domain as the set of criteria is general and not really focused on cultural content. The group of evaluators should change to involve experts in a field related to the domain of the evaluated websites, but the rest of the steps would be similarly implemented.

In order to confirm the claim of the appropriateness of these methods in the evaluation experiments of websites of a different domain, it is among our future plans to apply these methods in other domains, such as e-government and e-health. Furthermore, we aim at comparing the results of these two models with other multi-criteria decision-making models in order to find out which model provides the best results.

**Author Contributions:** Conceptualization, K.K.; methodology, K.K.; investigation, A.B., K.K., C.K.; data curation, A.B., K.K.; writing—review and editing, K.K., A.B.; project administration, K.K., C.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- Jesús, G.M.; Nuria Recuero, V.; Blasco, L.F. The influence of museums' websites on users' intentions. *ESIC Market. Econ. Bus. J.* 2017, 48, 369–392.
- Cunnliffe, D.; Kritou, E.; Tudhope, D. Usability evaluation for museum web sites. *Mus. Manag. Curatorsh.* 2001, 19, 229–252. [CrossRef]
- 3. van Welie, M.; Klaasse, B. *Evaluating Museum Websites using Design Patterns;* Technical Report Number: IR-IMSE-001; Vrije Universiteit: Amsterdam, The Netherlands, 2004.
- 4. Di Blas, N.; Guermand, M.P.; Orsini, C.; Paolini, P. Evaluating the Features of Museum Websites. In Proceedings of the Museums and the Web 2002: Selected Papers from an International Conference 2002, Boston, MA, USA, 17–20 April 2002.
- 5. Fotakis, T.; Economides, A.A. Art, science/technology and history museums on the web. *Int. J. Digit. Cult. Electron. Tour.* **2008**, *1*, 37–63. [CrossRef]
- 6. Pallas, J.; Economides, A.A. Evaluation of art museums' web sites worldwide. *Inf. Serv. Use* **2008**, *28*, 45–57. [CrossRef]
- 7. Saaty, T. The Analytic Hierarchy Process; McGraw-Hill: New York, NY, USA, 1980.
- 8. Mulubrhan, F.; Akmar Mokhtar, A.; Muhammad, M. Comparative Analysis between Fuzzy and Traditional Analytical Hierarchy Process. *MATEC Web Conf.* **2014**, *13*, 01006. [CrossRef]
- 9. Kabassi, K. Evaluating Museum Websites Using a Combination of Decision-Making Theories. *J. Herit. Tour.* **2019**. [CrossRef]
- 10. Kabassi, K.; Amelio, A.; Komianos, V.; Oikonomou, K. Evaluating Museum Virtual Tours: The case study of Italy. *Information* **2019**, *10*, 351. [CrossRef]
- Chou, S.-Y.; Chang, Y.-H.; Shen, C.-Y. A fuzzy simple additive weighting system under group decision-making for facility location selection with objective/subjective attributes. *Eur. J. Oper. Res.* 2008, 189, 132–145. [CrossRef]
- 12. Triantaphyllou, E. *Multi-Criteria Decision Making: A Comparative Study;* Kluwer Academic Publishers (now Springer): Dordrecht, The Netherlands, 2000; p. 320.
- 13. Senapati, T.; Yager, R.R. Fermatean fuzzy weighted averaging/geometric operators and its application in multi-criteria decision making methods. *Eng. Appl. Artif. Intell.* **2019**, *85*, 112–121. [CrossRef]
- 14. Kabassi, K. Evaluating Websites of Museums: State of the Art. J. Cult. Herit. 2017, 24, 184–196. [CrossRef]
- 15. Davoli, P.; Mazzoni, F.; Corradini, E. Quality Assessment of Cultural Web Sites with Fuzzy Operators. *J. Comput. Inf. Syst.* **2005**, *46*, 44–57.
- 16. Kabassi, K.; Maravelakis, E.; Konstantaras, A. Heuristics and Fuzzy Multi-criteria Decision Making for Evaluating Museum Virtual Tours. *Int. J. Incl. Mus.* **2018**, *11*, 1–21. [CrossRef]
- 17. Kabassi, K.; Botonis, A.; Karydis, C. Evaluating the Websites of the Museums' Conservation Labs: The Hidden Heroes. In Proceedings of the 9th International Conference on Information, Intelligence, Systems and Applications, Zakynthos, Greece, 23–25 July 2018.
- 18. Hwang, C.L.; Yoon, K. *Multiple Attribute Decision Making: Methods and Applications*; Springer: New York, NY, USA, 1981.
- 19. Karoulis, A.; Sylaiou, S.; White, M. Usability Evaluation of a Virtual Museum Interface. *Informatica* **2006**, 17, 363–380.
- 20. Sweetnam, M.S.; Agosti, M.; Orio, N.; Ponchia, C.; Steiner, C.M.; Hillemann, E.-C.; Siochrú, M.; Lawless, S. User Needs for Enhanced Engagement with Cultural Heritage Collections. In Proceedings of the Second International Conference TPDL, Paphos, Cyprus, 23–27 September 2012.
- 21. Chen, C.T. Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets Syst.* **2000**, *114*, 1–9. [CrossRef]
- 22. Zhao, R.; Govind, R. Algebraic characteristics of extended fuzzy numbers. *Inf. Sci.* **1991**, 54, 103–130. [CrossRef]
- 23. Yager, R.R.; Filev, D.P. Essentials of Fuzzy Modeling and Control; Wiley: New York, NY, USA, 1994.
- 24. Tsaur, S.H.; Tzeng, G.H.; Wang, G.C. Evaluating tourist risks from fuzzy perspectives. *Ann. Tour. Res.* **1997**, 24, 796–812.
- 25. Tang, M.T.; Tzeng, G.H.; Wang, S.W. A hierarchy fuzzy MCDM method for studying electronic marketing strategies in the information service industry. *J. Int. Inf. Manag.* **1999**, *8*, 1–22.

- 26. Yao, J.S.; Wu, K. Ranking fuzzy numbers based on decomposition principle and signed distance. *Fuzzy Sets Syst.* **2000**, *116*, 275–288. [CrossRef]
- 27. Klic, G.I.; Yan, B. Fuzzy Sets and Fuzzy Logic Theory and Applications; Prentice Hall: London, UK, 1995.
- 28. Yao, J.S.; Chiang, J. Inventory without backorder with fuzzy total cost and fuzzy storing cost defuzzified by centroid and signed distance. *Eur. J. Oper. Res.* **2003**, *148*, 401–409. [CrossRef]
- 29. De Angeli, A.; Matera, M.; Costabile, M.F.; Garzotto, F.; Paolini, P. On the Advantages of a Systematic Inspection for Evaluating Hypermedia Usability. *Int. J. Hum. Comp. Inter.* **2003**, *15*, 315–335. [CrossRef]
- 30. Matera, M.; Costabile, M.F.; Garzotto, F.; Paolini, P. SUE inspection: An effective method for systematic usability evaluation of hypermedia. *IEEE Trans. Syst. Man Cybern. Part A* **2002**, *32*, 93–103. [CrossRef]
- 31. Jadhav, S.; Sonar, R. Framework for evaluation and selection of the software packages: A hybrid knowledge based system approach. *J. Syst. Softw.* **2011**, *84*, 1394–1407. [CrossRef]
- 32. Wubs, H.; Huysmans, F. Click to the Past. The Netherlend Institute for Social Research. Available online: http://www.scp.nl/english/Publications/Summaries\_by\_year/Summaries\_2006/Click\_to\_the\_past (accessed on 23 February 2020).
- 33. Purday, J. Think culture: Europeana.eu from concept to construction. *Electr. Libr.* 2009, 33, 170–180.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).