A comparison of different soft-computing techniques for the evaluation of handball goalkeepers

Eusebio Angulo¹ · Francisco P. Romero² · Julio A. López-Gómez²

Accepted: 15 October 2021 / Published online: 2 November 2021
© The Author(s) 2021

Abstract

The efficiency of handball goalkeepers is a good predictor of team ranking in tournaments, but despite this, very few studies have been carried out into the performance characteristics of elite goalkeepers. This paper provides the criteria for evaluating a handball goalkeeper and applies a variety of soft-computing methodologies for estimating their weights. More specifically, a fuzzy multi-criteria decision-making method, a metaheuristic optimisation algorithm, and statistical and domain-knowledge-based methods were used to evaluate the actions of goalkeepers during the game. Computer experiments were performed for all the proposed methodologies, using data from the 2020 European Men's Handball Championship, in order to estimate the weights of the indicators. Then, these weights were used to identify the best goalkeeper and identify and rank the top five goalkeepers as determined by the tournament organisers. The results obtained show that using the metaheuristic-based method is extremely helpful in quantifying the expert assessments, which are often challenging to express in a disaggregated form. The other two techniques offer a less optimal but more easily interpretable result for coaches and fans.

Keywords Player evaluation · Multi-criteria decision-making · Metaheuristics · Handball goalkeepers · Statistical analysis

1 Introduction

Handball is an important and popular sport, especially in Europe. The most powerful leagues, such as the German and French Premier Leagues, have very large budgets (DPA 2019). It is also one of the most widely-played sports, in multiple leagues (DHB 2019). Therefore, it is a sport with economic and human importance at the European level. Today, there is growing interest in evaluating the performance of handball teams, which involves establishing an impartial method of assessing players. In the worst case, assessments are merely subjective. The aim is to be able to make an unbi-

ased evaluation of players based on their statistics for the main actions in each match.

The most recent study, Oytun et al. (2020), evaluates different machine-learning models for predicting particular types of athletic performance in female handball players, as well as determining the significant factors influencing expected performance. The main difference between their study and this paper is that in that case, the authors applied a methodology to evaluate athletic performance in female handball players, while the present paper aims to assess the performance of handball goalkeepers.

Goalkeepers play a crucial role in handball. In the coaching community, it is well known that the goalkeeper's performance can predict the team's ranking in big events (see Hansen et al. 2017). There have, however, been relatively few research studies carried out regarding elite goalkeepers. For example, the methodology presented in Schwenkreis (2019) allows the effectiveness of a handball team, or of a single handball player, to be quantified, but it is insufficient for evaluating the performance of a handball goalkeeper in a match.

Recently, Hatzimanouil et al. (2017) conducted a full study of the shot effectiveness of each handball position with respect to the goalkeeper. Cvenić (2000) and Hatzimanouil

✉ Francisco P. Romero
FranciscoP.Romero@uclm.es

Eusebio Angulo
Eusebio.Angulo@uclm.es

Julio A. López-Gómez
JulioAlberto.Lopez@uclm.es

¹ Department of Mathematics, University of Castilla La Mancha, Paseo de la Universidad, 4, 13071 Ciudad Real, Spain

² Department of Information Systems and Technologies, University of Castilla La Mancha, Paseo de la Universidad, 4, 13071 Ciudad Real, Spain

(2019) proposed methodologies based on statistical studies for evaluating the performance of goalkeepers in handball. Cvenić (2000) applied three methods to the goalkeepers of the European Men's Handball Championship 2018 based on (i) the goalkeeper's save percentage, (ii) the goalkeeper's save percentage and time played, and (iii) further taking into account the distance from which the shots are taken. Hatzimanouil (2019), meanwhile, carried out a statistical study to classify the goalkeepers at the European Woman's Handball Championship 2018 into three categories, based on their save efficiency, time played and number of matches played. In this paper, our purpose is to establish the weights of the most outstanding actions of handball goalkeepers from different points of view during the course of a tournament. The most outstanding actions are those reported by the European Handball Federation (EHF) in their statistics for each match. Specifically, we address the following research questions: *Is it possible to obtain a weighting scheme for the actions of handball goalkeepers to:*

1. *RQ1*: identify the best goalkeeper in the tournament?
2. *RQ2*: identify the top five goalkeepers in the tournament?
3. *RQ3*: rank the best goalkeepers in the tournament correctly?

This paper uses three different approaches to answer the above research questions:

- **Multi-criteria—Group Decision-Making approach** The first of the methodologies for weighting the actions of goalkeepers is based on the application of multi-criteria decision-making (MCDM) theories, which provide a suitable solution for processing the results of the assessment. For this purpose, experts (30) were consulted in order to label the importance of each of the performance indicators linguistically. The final weighting scheme is obtained from these opinions after applying different fuzzy logic techniques of consensus and aggregation (Romero et al. 2020).
- **Metaheuristic Weighting** The weighting process can be transformed into an optimisation process in which we seek the maximum alignment of the weights obtained with the judgement of the experts in the competition (Chýna et al. 2013; Romero et al. 2021). With this aim, those weights that optimise the choice of the most valued goalkeeper or the choice and ranking of the five best goalkeepers in the competition are then found. Metaheuristic algorithms are used for this purpose.
- **Statistical Weighting** This method is based on the reference values established in Antón García (2005) widely used in handball today. Reference scales are thereby established for each type of player action over several high-level matches and championships. From these

scales, the reference weights are calculated from the resolution of a number of balance equation systems between the indicators representing the result of the same action (Goal/Save).

The features of these techniques are particularly interesting, in terms of accuracy and significance, for handball goalkeeper assessment. Firstly, each approach has been proven to establish weighting schemes in a decision-making process Pérez et al. (2018), Paul and Das (2015) and Sáez et al. (2014). Furthermore, all three allow the use of expert knowledge in order to solve the problem, which is particularly recommended in this context (Kvam 2011). Finally, they do so from three different viewpoints and via three different procedures: as a general opinion on the problem (MCDM), as a provider of relevant elements through which weights are subsequently optimised (PSO), or setting the weights of the indicators directly.

The design of the proposed approaches was different in terms of the sources of the weighting scheme. For the MCDM approach, we use the opinion of a set of experts; the metaheuristic approach is aimed at optimising the weighting scheme according to the choice of best goalkeeper in the competition, and the statistical approach is based on reference levels and the statistics of the Championship. However, to the best of our knowledge, there are no studies quantitatively assessing the impact that these approaches have on the choice of the best goalkeeper. To address this deficiency, this paper presents an assessment of the 2020 European Men's Handball Championship. The data have been gathered, analysed and compared in order to answer the research questions set out above.

Consequently, it is worth highlighting the following contributions in this study: The primary contribution of this paper is the empirical analysis of three representative strategies for weighting evaluation criteria (a fuzzy approach, a metaheuristic optimisation strategy and a statistical method), whose results are compared to the tournament's best goalkeepers. A significant secondary contribution is that the results obtained will offer guidance and valuable information for data selection and will be good predictors of goalkeeper performance. Thus, our results provide significant insights into the features that distinguish an outstanding goalkeeping performance in a short tournament from those of the other players. Additional contributions of this work are the following:

- Building a dataset from *play-by-play* data from official statistics of the 2020 Men's EHF European Championship (MEN's EHF EURO 2020), with the complete set of criteria for evaluating the performance of a handball goalkeeper.

Fig. 1 Best Goalkeeper Selection Process

- Estimating the weights to evaluate handball goalkeepers, taking into account each meaningful action in a handball match.
- Putting forward a real-world application and evaluation for the sport of handball by a comparison of several weighting schemes using classical ranking metrics.

The novelty of our work lies in the fact that we address a number of difficult matters not previously considered in the literature: 1) data collection and performance indicator calculation from play-by-play data, 2) comparison of different strategies (fuzzy logic, metaheuristics, statistical) for weighting performance criteria, 3) comparison with the rankings of the best goalkeepers in a tournament.

The paper is organised as follows: Sect. 2 describes the process for evaluating the performance of handball goalkeepers. Section 3 explains the different alternatives for establishing the weights of the evaluation criteria: fuzzy, metaheuristic and statistical. Section 4 presents our Case Study for assessing the performance of goalkeepers at the 2020 European Handball Championship. Conclusions and future work are outlined in Sect. 5.

2 Handball goalkeeper evaluation process

In this section, we present the proposed model for the evaluation of handball goalkeepers in short competitions. There are three significant steps, as shown in Fig. 1. First comes the identification of criteria. In fact, the selection of suitable criteria depends on the tournament. Next, the criteria weights must be determined. Then, the third step is to implement the multi-criteria decision-making method, which consists of choosing the method to be used to select the best player.

2.1 Motivation

Choosing the “Player of the Match” is one of the most ambiguous decision-making issues in the area of player eval-

uation. The selection process is critical for identifying the crucial components of a team. This is why it has become an essential focus for every team, but there is no standard for player evaluation, which should be undertaken based on the situation. It is always challenging to choose the most valuable player because the selection criteria are changing.

Moreover, if there is one position that requires individual assessment in a sport like handball, it is the goalkeeper. Some studies in the literature quantify the influence of a goalkeeper’s performance at more than 50% of the final result. Even so, the majority of coaches think that a different set of indicators should evaluate this specific position.

The position of the goalkeeper as a player is privileged; he is always the last defender when the opposing team is trying to score a goal, and the first to start the attack when the opposing team takes a shot that is not intercepted by the defense. Goalkeepers are the only players in handball who perform most of their actions without the help of their teammates. No one can compensate for their mistakes, and they play in a space that only they can access.

For all these reasons, goalkeepers have a more significant influence than other players on the development of the game and the final result of the match. We thus need different indicators to assess their performance in each game.

2.2 Selection criteria identification

The identification of decision-making criteria is very important in player-of-the-match selection. Many approaches still limit themselves to using the count of goals/saves as the single determining factor when choosing the best player.

The evaluation of only one criterion (goals or saves) is not the most suitable approach, since many other factors must be considered in the selection process. Nowadays, it is important to structure the problem and to assess pertinent criteria explicitly before reaching a decision. A number of methods exist to solve multi-criteria problems, and at the root of many of them is the idea that most decision-making can be improved by breaking down the general evaluation of alternatives into evaluations of a number of relevant criteria.

The quantitative and qualitative criteria for evaluating handball players have been selected after an expert-based evaluation process, choosing those that seemed most appropriate for the specific evaluation according to the experts (see Fig. 2):

- A collection of basic assessment criteria is chosen as a baseline for the decision process, to determine which indicators need to be considered to model the player/performance. This set can be obtained from the literature.

Fig. 2 Player-of-the-Match Selection Process **Table 1** Selected Criteria

Indicator	Description	Positive/Negative
7mSaves	7-metre Shots Saved	+
7mMiss	7-metre Shots Missed	+
7mGoals	7-metre Goals	–
6mCSaves	6-metre Shots Saved	+
6mMiss	6-metre Shots Missed	+
6mCGoals	6-metre Goals	–
WingSaves	Wing Shots Saved	+
WingMiss	Wing Shots Missed	+
WingGoals	Wing Goals	–
BTSaves	Breakthrough Saved	+
BTGoals	Breakthrough Goals	–
FBSaves	Fast Break Saved	+
FBMiss	Fast Break Missed	+
FBGoals	Fast Break Goal	–
9mSaves	9-metre Shots Saved	+
9mMiss	9-metre Shots Missed	+
9mGoals	9-metre Goals	–
TO	Turnover	–
ST	Steal	+
Goal	Goal/Empty Goal	+
Miss	Goal Missed	–
Ass	Assist	+
Excl	2m	–
P7	Provoke 7m	+

- This basic set is extended using the criteria of a National Handball Coach. The extension achieves a higher level of granularity than the Basic Indicators Set.
- The list of criteria is made available to a large number of people who can give their opinion on it (fans, journalists, players) and contribute new indicators to the set.
- A small number of high-level experts study the results of the process to make the final decision on the criteria to be taken into account

The indicators used in this paper for evaluating the specific position of goalkeeper can be summarised as shown in Table 1:

3 Determination of criteria weights

The weights of the related criteria have to be determined with the assumption that the weights represent the level of contribution to the overall performance of the team.

3.1 Multi-criteria—group decision-making approach

The proposed method is based on that set out in previous work by the authors for handball outfield players (Romero et al. 2020). In this approach, the set of criteria obtained is given to a set of experts who are weighted according to their coaching experience. Fuzzy linguistic terms, such as most important, important, and normal, were used to determine the experts' influence weightings, which were represented by fuzzy membership functions.

Every expert, depending on his/her own experience, must express an opinion regarding the chosen criterion. In this case, each criterion is evaluated by linguistic labels belonging to one of the following predefined label sets.

$$\begin{aligned}
 S^3 &= \{s_1^3 = H - High, s_2^3 = M - Medium, s_3^3 = L - Low\} \\
 S^5 &= \{s_1^5 = VH - Very_High, s_2^5 = H - High, \\
 s_3^5 &= M - Medium, s_4^5 = L - Low, \\
 s_5^5 &= VL - Very_Low\}
 \end{aligned} \quad (1)$$

These label sets are represented by using Trapezoidal and Triangular Fuzzy Numbers (Fig. 3).

Once all the experts have provided the required information, if they applied different sets of linguistic labels, the information is standardised (represented on the same scale) to be used later in the aggregation processes. For this purpose, multi-granular fuzzy linguistic modelling methods (Morente-Molinera et al. 2020) can be applied. Our study adopts a model based on the concept of 2-tuple fuzzy linguistic representation (Martínez and Herrera 2012). This approach has produced excellent results in practical applications, especially in terms of efficiency (Serrano-Guerrero et al. 2020).

A high level of consensus in the expert opinions is mandatory to assure the viability of the aggregation process, i.e. we require that the experts' estimates have a common inter-

Fig. 3 Linguistic label definitions as fuzzy sets

section at some α -level cut (Hsu and Chen 1996). For this purpose, we compute a global consensus degree that is compared against a consensus threshold provided by the set of experts prior to the application of the consensus. If this degree is greater than or equal to this consensus threshold, then the consensus-reaching process is considered successful, and hence it should end. Otherwise, the experts need to discuss it further.

Once the expert opinions are validated, we use fuzzy techniques to find a compact and synthesised representation of expert opinions. The aggregation process is carried out in the following steps:

1. Each of the n experts defines the linguistic vector $V^e = \{v_1^e, v_2^e, \dots, v_n^e\}$ indicating the importance that they give to each of the chosen criteria. The union of all these linguistic vectors generates the global assessment matrix containing all the valuations from all the experts.
2. *Individual Aggregation:* The linguistic vector $V^e = \{v_1^e, v_2^e, \dots, v_n^e\}$ representing the expert evaluations of all indicators S_e is aggregated. The weights assigned to the experts W_e are used. This is done through the following steps:
 - (a) *Baseline* A fuzzy definition of each label is used (balanced or unbalanced)
 - (b) *Aggregated Valuation Matrix* We obtain the weight of each label in each criterion according to the occur-

rence frequency and the expert weighting w_{e_n} . The weighted mean operator is used to do this.

- (c) *Fuzzification* We use a fuzzy operator in order to obtain a fuzzy set for each criterion, for example the minimum t-norm, which truncates each baseline membership function according to the previously computed weighting of each label.

Then, the aggregated output will be the fuzzy set representing how relevant each indicator would be in measuring player performance.

3. *Defuzzification* The centre of area (COA) defuzzification method computes the centre of mass of the membership function of the fuzzy set (the centroid). The COA method maintains the underlying semantic ranking relation within the set of linguistic labels, i.e. given two linguistic labels $s_i, s_j \in S$ such that $s_i < s_j$ then $u_{COA}(s_i) < u_{COA}(s_j)$. Thus, the centroid of a type-1 fuzzy set A in a continuous domain X is calculated as follows:

$$U_{COA}(A) = \frac{\int_x x * \mu_A(x) dx}{\int_x \mu_A(x)}$$

and this will be the method used here to obtain a numerical output.

4. *Stratification* A stratification process was carried out to reduce the sparsity of the weighting distribution. The purpose is to group criteria with similar evaluations using a hierarchical clustering algorithm. As a result, we obtain the criteria segmented into groups according to this importance.
5. *Collective Aggregation* The aggregation process obtains the final solution according to the opinions given by the experts. These results allow us to define a weighting scheme to evaluate the performance in accordance with this set of criteria. For this purpose, we aggregate this value for each group of indicators and normalise the final result.

3.2 Metaheuristic weighting

The process of evaluating goalkeepers in a handball championship and choosing the best of them can also be addressed as an optimisation problem.

This way, given the criteria provided by experts where they establish the set of indicators for each goalkeeper in the match, two kinds of problem can be identified: (i) approximating the position of a player in the ranking of goalkeepers, (ii) approximating the total score of a goalkeeper. The first is frequently a combinatorial problem, while the second is usually a continuous problem.

This paper uses a combined approach from i) and ii), trying to obtain an optimal set of weights which allows the total

score of the players to be calculated, and an optimal ranking built. To do this, three objective functions based on three metrics described in Sect. 4.1 were used, leading to three different optimisation problems: the first tries to optimise the weights of the indicator set to obtain a ranking where the first position is occupied by the best goalkeeper. The second optimises the weights to obtain a ranking in which as many goalkeepers as possible from the top 5, as indicated by the experts in their ranking, coincide (it does not matter if the order is the same). Finally, the third problem optimises the weights of the indicator set to obtain a ranking in which the top 5 is the same as in the ranking provided by the experts.

The solution of these problems will produce interesting results: on the one hand, it will be possible to analyse the subjectivity introduced by the experts, showing which indicators had more weight in their judgement. On the other hand, experts will have feedback to assess and improve their own criteria, since they usually evaluate the players according to qualitative criteria and opinions. Finally, it will be possible to detect whether there is subjectivity when choosing the best goalkeepers of the tournament by the European Handball Federation (EHF).

In order to solve these problems, metaheuristic algorithms were employed. Metaheuristics are algorithms which are designed to solve a broad range of optimisation problems, guiding the search process and trying to explore the whole search space (see Talbi 2002; Boussaïd et al. 2013). In comparison with heuristics, they are of a higher level of abstraction, being able to incorporate and control different heuristics and adding mechanisms to escape from local optima in order to obtain the global optimum. Specifically, Particle Swarm Optimisation (PSO) was used in this study (see Shi and Eberhart 1998). It is a state-of-the-art population-based metaheuristic inspired by bird flocking. PSO starts by initialising a population of particles. They are the initial solutions or initial weights for each indicator in the indicator set which will be optimised during the PSO execution. Each solution has position and speed attributes which are first initialised. Then, the particles start to move and update their positions and speeds according to the positions that the particle has traveled and the best position in the swarm. These rules guarantee a good tradeoff between exploration and exploitation and ensure the algorithm's convergence (see Jiang et al. 2014; Samma et al. 2016). The particular expressions for updating the position and speed of each particle depends on the specific variant or PSO implementation used. The pseudocode of a canonical PSO algorithm is shown in algorithm 1.

3.3 Statistical weighting

In this section, two methodologies are established to assess the performance of a goalkeeper in a handball match. For this

Algorithm 1: Pseudocode of a Canonical Particle Swarm Optimisation Algorithm

```

1:  $f$  = Objective function to be maximised
2:  $g_{best} = []$  (Initial global best particle. Empty)
3:  $n$  = Population size

4: (Initialisation).
5: for  $i = 1 : n$  do
6:    $x_i$  = Random initialisation of the position of  $i$ th particle
7:    $v_i$  = Random initialisation of the speed of  $i$ th particle
8:    $p_{best_i} = x_i$  (Initial local best for  $i$ th particle)
9:   if ( $f(p_{best_i}) > f(g_{best})$ ) then
10:      $g_{best} = p_{best_i}$ 
11:   end if
12: end for

13: (Optimisation using PSO).
14: while stopping criterion is not satisfied do
15:   for  $i = 1 : n$  do
16:      $v_i$  = Update speed for  $i$ th particle according to the
       specific equation of the PSO version used.
17:      $x_i$  = Update position for  $i$ th particle according to the
       specific equation of the PSO version used.
18:     if ( $f(x_i) > f(p_{best_i})$ ) then
19:        $p_{best_i} = x_i$ 
20:       if ( $f(p_{best_i}) > f(g_{best})$ ) then
21:          $g_{best} = p_{best_i}$ 
22:       end if
23:     end if
24:   end for
25: end while
26: return  $g_{best}$ 

```

purpose, the value or weight of each type of significant action by the goalkeeper during a handball match is estimated. For example, each type of shot will be assessed as a goal, as a save by the goalkeeper or a miss by the shooter. Moreover, if the goalkeeper scores a goal or gives an assist it will also be evaluated among the other actions.

First of all, the official European Statistics established by the European Handball Federation (EHF) includes 25 types of action. To be able to use both methodologies based on the statistics of the goalkeepers, we have to make the following simplification:

1. There are 7 goalkeeper actions which ultimately have little weight due to the infrequency with which they happen in a handball match. These are: Turnover, Assist, Received 7 meters, Missed shot to goal, Exclusion, Goalkeeper's goal and Steal. Looking at all the goalkeepers of the European Championship and their statistics in all the matches played, between the 7 actions there were only 79 events. If we compare this to the 4611 events covering the other actions, they are less significant due to their limited occurrence in matches. We will give these 7 parameters a constant reference value based on the study of Antón García (2005), adding this weight when they are positive actions, for example giving an assist, and subtract

it when they are negative actions, for example missing a ball.

2. In addition, when a player shoots it must be a goal, a save or a miss (throwing wide or high of the goal). To simplify matters, consider the possible outcomes to be only goal or no goal, that is, a save and a missed shot (throwing wide) are the same. It is assumed that a shot misses the goal because of the good performance of the goalkeeper, and also, the goalkeeper's team does not concede the goal. With this second simplification we go from calculating the weights for 6 (types of shots) * 3 (possibilities goal, no goal or miss) = 18 variables to 6 (types of shots) * 2 (possibilities goal or no goal) = 12 variables. In addition, there are far fewer missed shots than saved shots or goals, for example 7-meter shots were only missed (rather than saved) 28 times out of 338 total.

The first methodology will be grounded on the reference values established by Antón García (2005), which is an important study of handball, where the reference values for each type of action by handball players during various high-performance matches and championships are analysed.

After the simplifications explained above, we have 12 variables left to estimate their weights, which can be grouped in pairs, since they are directly related, each type of shot being either a goal or not a goal. For each type of shot, we solve the following system of equations:

$$\begin{aligned}\beta_1 X + (1 - \beta_1)Y &= K_1 \\ \beta_2 X + (1 - \beta_2)Y &= K_2\end{aligned}\quad (2)$$

where X is the weight of each non-goal shot, Y the weight for each goal shot, and β_1 the normal reference value of a non-goal shot probability of the goalkeeper for this type of shot. β_2 is the reference value of the non-goal shot probability of the goalkeeper has for this type of shot with a very negative performance (β_1 and β_2 are calculated from Antón García (2005)). The scoring will be normalised over all the values $[-5, 5]$. Therefore, good performances will have a value $K_1 = 2$ and a normal performance $K_2 = 0$.

This system of equations is repeated with its reference values (β_1 and β_2) for each type of shot, giving all the weights for actions based on the mean and variance of the global data from Handball European 2020 for the indicator set.

The second methodology has the same simplifications as the first, but the way the β_1 , β_2 , K_1 and K_2 are calculated varies. In this case, they are calculated based on the total data of goalkeepers of the Handball European Championships held in 2020. Specifically, the average probability of goals over the total number of shots is calculated as $\bar{x}$ and the standard deviations are calculated as σ for each type of shot, using the statistics from all the matches of the European

Championship. The same formula as in (2) is applied, but in this second methodology:

- $\beta_1 = \bar{x} + \sigma$.
- $\beta_2 = \bar{x} - \sigma$.
- $K_1 = 5$, represents the score of a great performance.
- $K_2 = -5$, represents the score of a very negative performance.

With these premises, both parameters are recalculated for each type of shot and the weight is calculated in each case, for shots that are goals and not goals.

4 Experiment

For the implementation of the experiments, we analyse the best goalkeepers at the Handball European Championship held in 2020.

4.1 Evaluation metrics

This subsection describes the different metrics used in the paper to assess the results of the decision-making methods.

- Mean Reciprocal Rank (MRR): It tries to measure "Where is the first relevant item?", in this case, "Where is the Most Valuable Goalkeeper?". More formally, MRR calculates the reciprocal of the rank at which the first relevant document was retrieved (see Craswell 2018). The Most Valuable Goalkeeper in the dataset used is Gonzalo Pérez de Vargas. Thus, given a dataset D where *best* is the player ranked first, the aim is to obtain a set of weights for the indicators from the indicator set in order to build a ranking where *best* ranks first. Formally, it is computed as the inverse of the position of the best player in the ranking built according to Equation (3).

$$MRR = \frac{1}{n} \sum_{i=1}^n \frac{1}{rank_{best}} \quad (3)$$

where n is the number of times the ranking is built or the number of repetitions of the experiment and $rank_{best}$ is the position occupied in the ranking by the best player. The *MRR* function is bounded between 0 and 1. Therefore, when the experiment is formulated as an optimisation problem, the aim is the following: given a dataset D , the best player *best* and a fixed number of experimental repetitions n , the aim is to obtain a set of weights which maximises MRR. This means the best player will rank first in the ranking built by the optimisation algorithm in

each repetition. This problem is formulated by means of Eq. 4.

$$\text{Maximise } MRR(D, best, n)_{rank_{best} \in \mathbb{Z}} \quad (4)$$

- Mean Average Precision (MAP): this is a more strict metric than the MRR. It evaluates the complete list of recommended items up to a specific cut-off N . In this way, MAP will compute the elements of the top N which are found, heavily weighting the error made at the top of the list and gradually decreasing the importance of the error as they go to the lower items on the list. More formally, MAP is the arithmetic mean of the average precision values for an information retrieval system over a set of N query topics (see Sakai (2007) and Beitzel et al. (2018)). The drawback of this metric is that it does not consider the recommended list as an ordered list. This paper uses a simplified version of MAP. To do that, MAP is computed using only the successes, which means, the elements of the top N identified. The errors made are not weighted and computed. Thus, MAP is computing according to Eq. (5).

$$MAP = \frac{1}{N} \sum_{i=1}^N \frac{i}{rank_i} \quad \text{if } i = rank_i \quad (5)$$

Thus, MAP will be a bounded function between 0 and 1 where the values of MAP are continuous values which computes the number of top N goalkeepers detected out of the total of N . In the experiments carried out, $N = 5$ has been fixed.

When the problem is formulated as an optimisation problem, the formalisation is the following: given a dataset D and a cut-off N , the aim is to obtain a set of weights which maximises MAP formula according to Eq. (6).

$$\text{Maximise } MAP(D, N)_{rank_i \in \mathbb{Z}} \quad (6)$$

- A modified Manhattan distance (MMD): Inspired by Ekstrøm et al. (2019), a metric is defined with the purpose of comparing two rankings based on the distance between them. For each item, the absolute error has an interpretation as the Manhattan distance of the individual rankings from the ground-truth ranking. Therefore, it evaluates whether the top N elements of the ranking provided by the experts in the dataset D is exactly the same and is in the same order as the ranking built. In the experiments carried out, $N = 5$ is fixed. This is a bounded metric between 0 and 1 which provides a continuous value in order to obtain more accurate information about the ranking and the deviation from the ranking pro-

Table 2 Selected GoalKeepers

Rank	GoalKeeper	Team
1	Gonzalo Perez de Vargas	Spain
2	Torbjoern Bergerud	Norway
3	Marin Sego	Croatia
4	Roland Mikler	Hungary
5	Alfredo Quintana	Portugal

vided by the experts. This metric is computed according to Eq. (7).

$$MMD = \frac{1}{N} \sum_{i=1}^N \frac{1}{|i - rank_i| + 1} \quad (7)$$

where N is the specific cut-off to compute the metric, i is the player who ranks i , and $rank_i$ is the position held by the i element in the ranking built. The metric computes the distance between the element i in the ranking provided by the experts and in the ranking built. It then uses the inverse function and accumulates the results of the top N elements.

When the problem is addressed as an optimisation problem, the formulation is the following: given a dataset D , and a cut-off N , the aim is to obtain a set of weights which maximises MMD according to Eq. (8).

$$\text{Maximise } MMD(D, N)_{rank_i \in \mathbb{Z}} \quad (8)$$

4.2 Alternative selection

For the implementation of the experiment, we have chosen the five goalkeepers selected as the best goalkeepers in the competition (see Table 2). The selection is based on their performances during the tournament (see Table 3). The Most Valuable Goalkeeper was calculated 40 percent from the votes received by fans; a panel of EHF experts decided the remaining 60 percent.

4.3 Weighting results

This section presents the processes for calculating the weighting schemes for each evaluation criteria using the alternatives set out above and the results obtained.

4.3.1 Fuzzy decision-making

The set of criteria defined to evaluate a goalkeeper is given by 30 Spanish handball coaches who are weighted according to their coaching experience. As mentioned above, fuzzy

Table 3 Excerpt of GoalKeepers Statistics in 2020 EHF Championship

	MP	7mS	7mG	6mCS	6mCG	FBS	FBG	9mS	9mG	NSS	NSG	TO	ST	Goal	Miss	Ass
M. Sego	9	1	11	13	31	5	17	24	43	33	75	2	1			
Gonzalo P.V.	9	7	14	5	18	2	12	30	42	26	63			2		
R. Mikler	7	6	11	12	26	1	17	30	41	26	78		1			1
T. Bergerud	9	4	17	14	38	3	13	45	55	38	92					1
A. Quintana	8	3	11	15	50	5	11	29	44	47	119	3	2	1	1	

linguistic terms, such as most important, important, and normal, were used to determine the coaches' experience. For example, a National Handball coach was assigned the *most important* influence, while a Level 1 coach was assigned an *important* influence weighting.

Each expert can decide individually on the level of importance that should be given to each criterion value. For this purpose, they used the following linguistic label set:

$$S^3 = \{s_1^3 = H - High, s_2^3 = M - Medium, s_3^3 = L - Low\}$$

The results of the survey can be seen in Fig. 4 (left) (dark blue: very important, light blue: important, white, not very important).

Using the fuzzy definition of the linguistic labels (see Fig. 3), the fuzzification of each of the criterion assessments can be carried out. Once it is verified, there is enough consensus among the opinions collected from the experts, an aggregation process must be carried out to obtain a weight for each criterion. For this purpose, linear combination of each element according to the weight of the criteria is applied.

Then, after a defuzzification process, we obtain a weighting for each criterion represented as a weight vector (See Table 7 MCDM-R). This approach does not guarantee a normalisation of the criteria weights, making comparison with other methods difficult. Thus, to ensure that the weighting scheme is consistent and coherent with different techniques, we perform a normalisation process (See Table 7 MCDM-N).

In addition, with the aim of reducing the sparsity of the weightings obtained, we carry out a clustering/stratification process as can be seen in Fig. 4 (right).

Consequently, different groups of indicators, with their corresponding weightings are found. As a result, we aggregate this value for each group of indicators and normalise the final result. See the weighting scheme obtained in Table 7 MCDM-G. Table 8 shows the results obtained according to the evaluation metrics using these weighting schemes. A priori, a weighting scheme based on expert opinion could be an excellent strategy to evaluate goalkeeper performance. In contrast, the results obtained show that this procedure is not especially useful for identifying the best goalkeepers in a

tournament. The main problem is related to ignoring the frequency of actions in the process of evaluating the goalkeeper. This weakness is detrimental to the evaluation of those goalkeepers who play longer during the tournament, who are the most eligible as the best goalkeepers in the competition and who may accumulate more negative actions considered important by the experts. Nevertheless, the results obtained allow us to order the top 5 best goalkeepers almost correctly (see Table 6). This way, it can be seen as a useful technique in order to have a ranking of goalkeepers previously chosen by other criteria.

4.3.2 Metaheuristic weighting

The process of finding the optimal weights for the indicators from the indicators set using metaheuristic algorithms is simple, except for two issues that must be defined a priori: the first is related to the codification of the individuals or particles in the case of PSO. The second is concerned with the fitness function which will be used to run the metaheuristic algorithm.

This way, each individual is codified as a real value vector with as many elements as there are indicators in the set of indicators. Each indicator of each individual is initialised according to the bounds defined for a given indicator. Therefore, the algorithm will have an input matrix of $m \times n$ dimension, where m is the number of individuals in the population (30 individuals have been used) considered by the algorithm and n is the number of indicators. Regarding the fitness function the algorithm will optimise, it must be chosen according to the nature of the problem (combinatorial, continuous) and the implementation of the algorithm. Metrics described in Subsect. 4.1 have been used leading to three different optimisation problems. In order to validate the results obtained, 100 repetitions of each problem were carried out.

The results obtained in the metaheuristic weighting experiments are shown in Table 7. This shows the different attributes or indicators from the indicator set, as well as its value for each metric. Mean and median values for each of the described metrics have been reported to provide more accurate information. Furthermore, the values of the metrics for each set of weights shown in Table 7 were computed. Table 8

Fig. 4 Expert Opinions (left) and Clustering (right)

Table 4 Parameters for methodology 1

Indicators	β_1	β_2	K_1	K_2
7m saves and goals	0.25	0.17	2	0
6m saves and goals	0.43	0.35	2	0
Wing saves and goals	0.40	0.32	2	0
BT saves and goals	0.43	0.35	2	0
FB saves and goals	0.21	0.13	2	0
9m saves and goals	0.68	0.60	2	0

shows these results. Although *MeanManh* and *MedManh* are not capable of identifying the best goalkeeper in the tournament, they offer good results to identify and rank the top 5 goalkeepers.

4.3.3 Statistical weighting

The process of calculating weights for the indicators is explained in Subsect. 3.3. After simplification, the weights are estimated by solving Eq. 2. With this resolution, the weight of each *goal* and *no goal* is obtained for each type of shot.

Tables 4 and 5 show the values (β_1 , β_2 , K_1 and K_2) for the methodology 1 and methodology 2, respectively. Methodology 1 based on the reference values indicated Antón García (2005) and methodology 2 based on the mean and variance of the global data from Handball European 2020 for indicators set.

The results obtained in the statistical weighting experiment are shown in Table 7. The results are consistent if we take into account the possible game-changing actions for a goalkeeper in a match and the value of these actions. The

Table 5 Parameters for methodology 2

Indicators	β_1	β_2	K_1	K_2
7m saves and goals	0.455	0.149	5	−5
6m saves and goals	0.464	0.270	5	−5
Wing saves and goals	0.503	0.284	5	−5
BT saves and goals	0.464	0.270	5	−5
FB saves and goals	0.354	0.094	5	−5
9m saves and goals	0.627	0.461	5	−5

main problem of this methodology is that it does not take into account the number of matches played by each goalkeeper. Since at high performance the players are more effective and this means the goalkeepers with most games have more goals conceded and their score is thus penalised.

4.4 Results and discussion

Table 6 shows the ranking results obtained for each technique and their relative order (in parentheses) among the top 5 selected. Table 7 shows the weights calculated per action for every technique implemented. The evaluation of these results according to the performance metrics is shown in Table 8. The most suitable strategy for answering the three research questions is to use a metaheuristic optimisation algorithm in order to optimise the distance-based performance metric.

The results obtained through the use of a decision-making mechanism based on fuzzy logic are not suitable for any of the three metrics. This is because the proposed mechanism is based on the importance of the actions without considering their frequency. Thus, a negative action that is very frequent will hurt goalkeepers who play many games. Nev-

Table 6 Selected GoalKeepers Ranks

GoalKeeper	Team	EHF	MCDM	Meta	Stats
Gonzalo Perez de Vargas	Spain	1	8 (1)	6 (5)	5 (3)
Torbjoern Bergerud	Norway	2	16 (3)	2 (1)	11 (5)
Marin Sego	Croatia	3	12 (2)	3 (2)	2 (1)
Roland Mikler	Hungary	4	19 (4)	4 (3)	3 (2)
Alfredo Quintana	Portugal	5	29 (5)	5 (4)	7 (4)

ertheless, it is capable of ordering the top five goalkeepers correctly (see Table 6), and so it can be considered a proper technique for sorting alternatives previously chosen through other techniques, based on the frequency of actions. Regarding the weights obtained by this approach in comparison to the best fit (which is obtained by using metaheuristic algorithms) shown in Table 7, it should be noted that five features are not weighted (*TO*, *ST*, *Miss*, *Excl* and *P7*) because they are not considered relevant by the experts. In addition, *6mCSaves*, *WingGoals*, *BTSaves*, *BTGoals*, *FBGoals* and *9mGoals* features show a large deviation from the weights obtained by using metaheuristic algorithms because the fuzzy approach ignores the frequency of the actions.

The best results, according to the selected performance metrics, are obtained by using the metaheuristic-based weighting method. However, from the point of view of handball experts, it is not reasonable that *WingSaves* and *9mSaves* are more valuable for a goalkeeper than *7mSaves*, *FBsaves* and *6mSaves*. This is because this methodology optimises by fitting the 5 best goalkeepers in the tournament, and it is likely that the variables that make the difference are in the types of shot a goalkeeper has more chance of saving. Moreover, this is a “blind” approach which tries to optimise a performance metric, and therefore the results are not always easily interpretable, as the algorithm can over-fit the data to maximise the metric. Despite this, metaheuristic weighting provides the best result in the three metrics considered. Finally, we should not forget that the choice of the 5 best goalkeeper in the tournament is subjective and is established by the organisation.

Concerning the results obtained with statistical weighting, they are not the best among the metrics used, but neither are they bad, because they place 3 goalkeepers within the best 5 (see Table 6). As for the weights obtained for each indicator from the point of view of handball experts, they are the most reasonable since they are based on the study carried out by Antón García (2005) and on the European Handball Federation’s own statistics. Comparing the weights obtained by the statistical approach to the weights obtained by metaheuristic algorithms, it should be noted that the variability of the statistical weights is very low, in both positive and negative

features, which makes it difficult to differentiate properly between two goalkeepers.

Finally, with respect to the results obtained by the different approaches in the three metrics considered (see Table 8) the results can be grouped by metric. Regarding reciprocal rank, the fuzzy approach obtained the worst result with a maximum value of 0.125, followed by the statistical approach which returned 0.200 as the maximum value. Finally, the metaheuristic approach was able to identify the best goalkeeper in the tournament, returning a value of 1. For the MAP metric, the fuzzy approach gave the worst value, 0.200, followed by the statistical approach which returned a value of 0.600. The best value was again provided by the metaheuristic algorithm, which achieved a value of 0.800. In this case, the difference between the statistical and metaheuristic algorithms was not so great. Finally, concerning the modified Manhattan distance, the fuzzy and statistical approaches gave near-zero values, while the metaheuristic approach obtained 0.828 as the maximum value. Therefore, the data shown in Table 8 confirm that metaheuristic algorithms are the best quantitative approach for evaluating the performance of handball goalkeepers.

5 Conclusions

The evaluation of handball goalkeepers is a highly complex procedure that focuses on examining several different criteria. This paper shows how to implement this process and compares several soft-computing techniques to process the evaluation results.

To this end, the evaluation scenario based on the MEN’s EHF EURO 2020 is designed to compare several soft-computing techniques in a real tournament. This experiment shows that the metaheuristic-based method outperformed the other approaches in terms of identifying and ranking the best handball goalkeeper. This implies that the criteria weights to be used to identify and rank the best goalkeepers can be obtained by using this algorithm. The method also allows the expert evaluations, which are often difficult to obtain in a disaggregated form, to be estimated and quantified. On the other hand, the use of statistical techniques allows the identification and sorting of the best goalkeepers in the tournament to

Table 7 Weights according to the method

Indicator	MDCM-R	MCDM-N	MCDM-G	MeAvPre	MedAvPre	MeanReRank	MedReRank	MeanManh	MedManh	%J.Anton	%Euro
7mSaves	1.48	1.96	1.50	1.45	1.62	1.39	1.59	1.06	1.10	0.21	0.23
7mMiss	1.04	1.38	1.00	0.87	0.75	1.01	1.03	0.96	1.03	0.21	0.23
7mGoals	-0.75	-1.00	-0.75	-0.53	-0.29	-0.56	-0.28	-0.66	-0.42	-0.04	-0.10
6mCSaves	1.48	1.96	1.50	0.52	0.39	0.63	0.46	1.43	1.70	0.16	0.33
6mMiss	0.88	1.17	1.00	0.89	0.85	0.74	0.48	1.28	1.38	0.16	0.33
6mCGoals	-0.86	-1.15	-0.75	-0.81	-0.81	-1.59	-1.79	-0.47	-0.43	-0.09	-0.19
WingSaves	1.30	1.73	1.50	1.34	1.40	1.57	1.80	1.10	1.09	0.17	0.28
WingMiss	0.93	1.23	1.00	1.48	1.63	0.87	0.79	1.44	1.61	0.17	0.28
WingGoals	1.11	-1.47	-1.50	-0.22	-0.06	-1.21	-1.22	-0.15	-0.08	-0.08	-0.18
BTSaves	1.48	1.96	1.50	0.95	0.95	0.99	0.92	0.89	0.79	0.16	0.33
BTGoals	-0.86	-1.15	-0.75	-0.31	-0.12	-0.17	0.00	-1.32	-1.48	-0.09	-0.19
FBSaves	1.51	2.00	1.50	1.03	1.07	0.91	0.78	0.99	0.99	0.22	0.30
FBMiss	1.06	1.41	1.00	1.41	1.52	0.93	0.89	0.96	0.99	0.22	0.30
FBGoals	-0.60	-0.80	-0.75	-1.65	-1.79	-0.95	-0.87	-0.17	-0.04	-0.03	-0.09
9mSaves	1.06	1.41	1.50	1.69	1.80	1.61	1.82	1.77	1.90	0.10	0.28
9mMiss	0.63	0.84	0.75	0.56	0.54	1.05	1.16	1.49	1.65	0.10	0.28
9mGoals	-1.06	-1.72	-1.50	-0.13	-0.03	-0.26	-0.09	-0.36	-0.22	-0.15	-0.33
TO	0.00	0.00	0.00	0.90	0.87	0.92	0.90	0.82	0.71	-0.20	-0.20
ST	0.00	0.00	0.00	1.03	1.06	0.85	0.71	1.08	1.24	0.20	0.20
Goal	1.25	1.49	1.65	1.01	0.92	1.01	0.98	0.86	0.79	0.50	0.50
Miss	0.00	0.00	0.00	-0.85	-0.62	-1.05	-1.09	-1.00	-0.95	-0.20	-0.20
Ass	1.50	2.00	2.00	0.70	0.68	1.09	1.17	1.10	1.09	0.20	0.20
Excl	0.00	0.00	0.00	-0.97	-0.88	-1.16	-1.34	-1.04	-1.10	-0.30	-0.30
P7	0.00	0.00	0.00	-1.08	-1.13	-1.01	-0.95	-0.98	-1.04	-0.30	-0.30

Table 8 Value of the metrics for the weights obtained in each methodology

Metric	MDCM-R	MCDM-N	MCDM-G	MeanAvPre	MedAvPre	MeanReRank	MedReRank	MeanManh	MedManh	%Ref-J.Anton	%Ref-Euro
Rec.Rank	0.125	0.062	0.083	0.250	0.330	1.000	1.000	0.142	0.166	0.200	0.083
MAP	0.200	0.000	0.000	1.000	1.000	0.400	0.600	0.800	0.800	0.600	0.200
Mod. Manh.	0.012	0.012	0.012	0.050	0.050	0.066	0.100	0.828	0.828	0.040	0.015

be carried out with acceptable results. Furthermore, it offers more consistent weights from the point of view of the game, which is advantageous to the general evaluation of the performance of handball goalkeepers.

In contrast, the use of a fuzzy technique based on expert opinions produced poor results in terms of choosing the best goalkeepers in the tournament. The main weakness of this technique is to ignore the frequency of the actions when assigning a specific weight in the evaluation process. This factor is detrimental to some goalkeepers in comparison to others.

In future work, this method can be explored or optimised along the following lines. Firstly, the application of a new soft-computing technique by applying a fuzzy aggregation operator to combine expert opinion with the statistical characteristics of each criterion. Secondly, it is possible to develop an adaptive metaheuristic algorithm that considers the insights obtained by the statistical techniques. The algorithm developed will therefore be “less blind”, since it will seek to optimise the objective function defined, and will do so taking into account the observations of the statistical approach, which provide practical information about the reality of each game.

Finally, the results obtained can be used as part of an evaluation model for handball goalkeepers. The application of this model could help those who make decisions about team line-ups, who sign new players, who choose national team goalkeepers, or need to choose better goalkeepers for matches, among many other applications.

Acknowledgements This work has been partially supported by the State Research Agency (AEI) of the Spanish Ministry of Science and Innovation under grants PID2019-104735RB-C42 and PID2020-112967GB-C32. Moreover, part of this research was included in the project BIDEA, which aimed at establishing a new framework for the quantification of the handball player performance to allow for a more advanced understanding of the game.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

Code availability The source code of the current work is available from the corresponding author on reasonable request.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Antón García J (2005) El modelo de juego en alto nivel en el balonmano de 2005: características, índices y escalas de rendimiento como bases para un modelo de preparación eficaz. *Área de balonmano* 37:1–9
- Beitzel SM, Jensen EC, Frieder O (2018) *Encyclopedia of Database Systems*. MAP, pp 2200–2201. Springer New York, New York, NY
- Boussaid I, Lepagnot J, Siarry P (2013) A survey on optimization meta-heuristics. *Inform Sci* 237:82–117. <https://doi.org/10.1016/J.INS.2013.02.041>
- Chýna V, Kuncová M, Sekni J (2013) Estimation of weights in multi-criteria decision-making optimization models. In: *Proceedings of 30th international conference mathematical methods in economics*, pp 355–360
- Craswell N (2018) *Encyclopedia of Database Systems*. Mean Reciprocal Rank, pp 2217–2217. Springer New York, New York, NY
- Cvenić J (2000) Contribution to methodology of efficiency evaluation of handball goalkeepers. *J Human Sport Exerc* 14(Proc5):S2480–S2486
- DHB (2019) Unser markenleitbild. URL <https://dhh.de/der-dhb/verband.html>
- DPA (2019) Start der ehf-champions league: Thq kiel und sg flensburg mssen vollgas geben. URL <https://www.shz.de/14920136>
- Ekstrøm CT, Gerds TA, Jensen AK (2019) Sequential rank agreement methods for comparison of ranked lists. *Biostatistics* 20(4):582–598
- Hansen C, Sanz-Lopez F, Whiteley R, Popovic N, Ahmed HA, Cardinale M (2017) Performance analysis of male handball goalkeepers at the world handball championship 2015. *Biol Sport* 34(4):393
- Hatzimanouil D (2019) Goalkeepers' rating, evaluation and classification, according to the number of games, participation time and effectiveness at the women's european handball championship in 2018. *Int J Perform Anal Sport* 19(4):595–607
- Hatzimanouil D, Giatsis G, Kepesidou M, Kanioglou A, Loizos N (2017) Shot effectiveness by playing position with regard to goalkeeper's efficiency in team handball. *J Phys Educ Sport* 17(2):656–662
- Hsu H-M, Chen C-T (1996) Aggregation of fuzzy opinions under group decision making. *Fuzzy Sets Syst* 79(3):279–285. [https://doi.org/10.1016/0165-0114\(95\)00185-9](https://doi.org/10.1016/0165-0114(95)00185-9)
- Jiang S, Ji Z, Shen Y (2014) A novel hybrid particle swarm optimization and gravitational search algorithm for solving economic emission load dispatch problems with various practical constraints. *Int J Electr Power Energy Syst* 55:628–644
- Kvam PH (2011) Comparing hall of fame baseball players using most valuable player ranks. *J Quant Anal Sports* 7(3)
- Martínez L, Herrera F (2012) An overview on the 2-tuple linguistic model for computing with words in decision making: extensions, applications and challenges. *Inform Sci* 207:1–18
- Morente-Molinera J, Wu X, Morfeq A, Al-Hmouz R, Herrera-Viedma E (2020) A novel multi-criteria group decision-making method for heterogeneous and dynamic contexts using multi-granular fuzzy linguistic modelling and consensus measures. *Inform Fus* 53:240–250. <https://doi.org/10.1016/j.inffus.2019.06.028>
- Oytun M, Tinazci C, Sekeroglu B, Acikada C, Yavuz HU (2020) Performance prediction and evaluation in female handball players using machine learning models. *IEEE Access* 8:116321–116335. <https://doi.org/10.1109/ACCESS.2020.3004182>
- Paul S, Das S (2015) Simultaneous feature selection and weighting - an evolutionary multi-objective optimization approach. *Pattern Recog Lett* 65:51–59
- Pérez I, Cabrerizo F, Alonso S, Dong Y, Chiclana F, Herrera-Viedma E (2018) On dynamic consensus processes in group decision making problems. *Inform Sci* 459:20–35
- Romero FP, Angulo E, Serrano-Guerrero J, Olivas JA (2020) A fuzzy framework to evaluate players' performance in handball. *Int J Comput Intell Syst* 13(1):549–558
- Romero FP, Lozano-Murcia C, López-Gómez JA, Sanchez-Herrera EA, Sanchez-Lopez E (2021) A data-driven approach to predicting the most valuable player in a game. *Comput Math Methods*. <https://doi.org/10.1002/cmm4.1155>
- Sakai T (2007) Alternatives to Bpref. In: *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR'07* pp 71–78
- Samma H, Lim CP, Saleh JM (2016) A new Reinforcement Learning-based Memetic Particle Swarm Optimizer. *Appl Soft Comput* 43:276–297
- Schwenkreis F (2019) A graded concept of an information model for evaluating performance in team handball. In: Hammoudi S, Quix C, Bernardino J (eds) *Proceedings of the 8th International Conference on Data Science, Technology and Applications, DATA 2019, Prague, Czech Republic, July 26–28, 2019*, pp 196–202. *SciTePress*. <https://doi.org/10.5220/0007920001960202>
- Serrano-Guerrero J, Chiclana F, Olivas JA, Romero FP, Homapour E (2020) A TIOWA fuzzy linguistic aggregation methodology for searching feature-based opinions. *Knowl Based Syst*. <https://doi.org/10.1016/j.knosys.2019.105131>
- Shi Y, Eberhart RC (1998) A Modified Particle Swarm Optimizer. In: *Proceedings of IEEE International Conference on Evolutionary Computation*, pp 69–73. IEEE Computer Society, Washington, DC, USA
- Sáez JA, Derrac J, Luengo J, Herrera F (2014) Statistical computation of feature weighting schemes through data estimation for nearest neighbor classifiers. *Pattern Recog* 47(12):3941–3948
- Talbi EG (2002) A Taxonomy of Hybrid Metaheuristics. *J Heurist* 8:541–564. <https://doi.org/10.1023/A:1016540724870>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.