

On the rating system in alpine skiing racing: Criticism and new proposals

Original

On the rating system in alpine skiing racing: Criticism and new proposals / Maisano, DOMENICO AUGUSTO FRANCESCO; Botta, Andrea; Franceschini, Fiorenzo. - In: PROCEEDINGS OF THE INSTITUTION OF MECHANICAL ENGINEERS. PART P, JOURNAL OF SPORTS ENGINEERING AND TECHNOLOGY. - ISSN 1754-3371. - STAMPA. - 230:4(2016), pp. 253-263. [10.1177/1754337115621818]

Availability:

This version is available at: 11583/2658368 since: 2016-11-30T17:08:51Z

Publisher:

SAGE Journals

Published

DOI:10.1177/1754337115621818

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

On the rating system in alpine skiing racing: criticism and new proposals

Domenico Maisano¹, Andrea Botta² and Fiorenzo Franceschini³

¹ domenico.maisano@polito.it ² andrebotta@hotmail.it ³ fiorenzo.franceschini@polito.it
Politecnico di Torino, DIGEP (Department of Management and Production Engineering),
Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

Abstract

Like most of the sports, alpine skiing has international regulations (designed by FIS, i.e., the *International Ski Federation*), aimed at coordinating competitions and rating athletes. So-called *FIS Points* represent the core of the rating system, as they allow to rate the athletes involved in competitions, for each of the five alpine skiing specialties, i.e., slalom, giant slalom, super-G, downhill and combined.

The objective of this paper is to analyze the current rating system, with special attention to the FIS Points construction and update, focusing on the potential weaknesses. Two major weaknesses emerge from this analysis: (i) the not very sound update process of FIS Points, based on the average of the best two results obtained in a specialty of interest, and (ii) the lack of a general rating of athletes based on their *eclecticism*, i.e., the ability to obtain good results in as many different specialties as possible. The second part of the paper presents some proposals for sorting out the above weaknesses. The description is supported by several practical examples, based on real and fictitious data.

Keywords: Indicators, Alpine skiing, Rating system, FIS Points.

1. Introduction

Indicators are essential tools for monitoring and evaluating complex systems in a variety of contexts [1, 2]. Common examples are socio-economic indicators, financial indicators, those used to control manufacturing processes, customer satisfaction indicators, and many others [3, 4].

In general, indicators are not “passive” observation tool, but can have a profound *normative* effect, i.e., conditioning the behaviour of the system monitored [5]. For this reason, the definition/selection of appropriate indicators is an important requirement, regardless of the context in which they are used.

In sports, individual or peer groups of athletes (e.g., teams) are generally rated through appropriate sets of indicators, which constitute the so-called *rating systems*. Let us consider, for example, the relatively complex rating systems in Formula One racing, tennis, artistic gymnastics, decathlon, etc. [1, 6, 7]. Apart from ratings systems, indicators are also used in sports for so-called *performance analysis*, which is aimed at understanding the physiological, psychological, technical and tactical demands of athletes, or even predicting the future behaviour of sporting activity [8, 9]. Performance

analysis – whose results are commonly used to support the activity of coaches, trainers, observers, bookmakers, etc. – is the subject of numerous contributions in the scientific literature, such as that by O'Donoghue [10] in tennis, Hughes and Franks [11] and Clemente et al. [12] in soccer, Hughes et al. [13] in rugby, and Barry et al. [14] in road cycling.

Returning to rating systems, they are generally developed by elected representatives of recognized international federations, which coordinate competitions [28]. Developing a rating system is a very delicate operation, with (at least) three basic requirements: (i) finding adequate way(s) to evaluate a set of abilities and skills of individual or peer groups of athletes, (ii) finding adequate way(s) to determine a score for individual and/or multiple competitions (e.g., tournaments, championships, etc.), and (iii) combining the scores obtained by each competitor in order to create a corresponding rating [29].

Despite the profound practical implications and normative effect, rating systems have rarely been analyzed from a scientific perspective. Among the few contributions in the literature, we recall the criticism and/or suggestions for improvement by Pluta et al. [16] in the field of basketball, Winchester and Stefani [17] in American football, and Mehrez et al. [18] in soccer.

The variety of sports and their great difference in terms of culture, tradition, social and economic pressure are reflected by the variety of rating systems [28]. Very interesting is the research by Stefani [15], proposing a general taxonomy of more than 150 sports and relevant rating systems. In a nutshell, sports are divided into *combat*, in which opponents are in direct physical contact (as in boxing and wrestling), *independent*, in which significant contact is not allowed (as in swimming and archery), and *object*, in which indirect contact is allowed while opponents attempt to control an object (as in basketball and football). On the other hand, rating systems are divided into *subjective*, which are usually decided subjectively, *accumulative*, in which points accrue non-decreasingly over a specific time-window, and *adjustive*, in which a rating self adjusts based on the difference between some observed result and a prediction of that result based on past performance.

Like all rating systems, those in sports competitions are often based on questionable and/or arbitrary conventions that, once established, tend to be tolerated by stakeholders (athletes, coaches, organizers, fans, etc.), without being further challenged [1]. However, the periodic adjustment of rating systems (e.g., consider the mutability of that one in Formula One racing) is evidence that they are far from being perfect and incontrovertible.

This paper focuses on the rating system of alpine skiing racing, which represents one of the key parts of the regulations designed by FIS, i.e., the *Fédération Internationale de Ski* or *International Ski Federation*, founded in 1924 and promoting the practice of various specialties of alpine skiing – i.e., slalom, giant slalom, super-G, downhill and combined [19]. The other key parts of the FIS regulations concern: (i) specifications for the technical equipment of athletes (e.g., ski length or sidecut radius limits, safety protections, etc.), (ii) specifications for the preparation of race tracks

(e.g., minimum/maximum permissible elevation, number of gates/poles, etc.), and (iii) constraints in the selection of teams of athletes for participating in sporting events (e.g., maximum number of athletes from the same country in World Cup races), etc..

The so-called *FIS Points* – hereafter abbreviated as *FP*¹ – represent the core of the rating system, as they allow to rate the athletes involved in the alpine skiing races, for each specialty. The *FPs* of the athletes involved in a generic race are also used for determining the starting order and estimating the level of difficulty of the race itself.

The objective of this paper is twofold. The first one is to analyze the current rating system, especially the *FP* construction and update process, trying to answer the following research questions: *What are the weaknesses (if any) of the rating system in use?* The second one is to make some new proposals, trying to answer the following research questions: *How can we improve (at least part of) the current rating system, so as to overcome the above weaknesses in a relatively simple way?*

The reminder of this paper is organized into four sections. Sect. 2 provides some background information on alpine skiing and the rating system in use. Sect. 3 criticizes some aspects of the rating system and provides new proposals for improvement. Description is supported by several practical examples, based on real and fictitious data. Finally, the concluding section summarizes the original contributions of the paper, practical implications, limitations and suggestions for future research.

2. Background information

This section is divided into two subsections, which provide a synthetic description of (i) alpine skiing and (ii) the rating system in use, from the perspective of *FP*.

2.1 Alpine skiing

Modern alpine skiing is divided into five specialties (slalom, giant slalom, super-G, downhill and combined, hereafter abbreviated as *SL*, *GS*, *SG*, *DH* and *KB* respectively). In a generic race, athletes have to run the same path, which is delimited by gates (or poles), arranged in a variety of configurations. Races related to different specialties generally differ in terms of track length, distance between gates and, consequently, speed of the skier, which is maximum in *DH* and minimum in *SL*. In *technical* specialties (i.e., *SL* and *GS*) gates are significantly closer and more angled than in *speed* specialties (i.e., *SG* and *DH*), thus requiring the athletes to run curves of smaller radius. *KB* is a sort of “hybrid” event consisting of one run of *DH* and one run of *SL*.

¹ Conventionally, the expression “*FP*” will refer to the FIS Points of an individual athlete in a specific specialty, while “*FPs*” will refer to multiple FIS Points, related to several athletes and/or specialties. FIS points are also used for other winter sports disciplines coordinated by FIS, different from alpine skiing, such as *cross-country skiing*, *snowboarding*, *ski jumping*, etc.. For simplicity, this paper will exclusively refer to alpine skiing.

From the competitive point of view, the objective of any race is to cover the race track as quickly as possible; each race can include one or more runs, depending on the specialty. Race timing starts when the athlete opens the so-called starting gate and ends when he/she crosses the finish line.

The determination of the starting sequence (or *starting list*) of athletes deserves special attention. Unlike other athletes, those in alpine skiing do not compete in the same conditions: due to the action of the athletes' ski edges, the snow around the gates (forming the race track) is progressively consumed, causing the formation of "holes", which make the run of athletes more difficult. This is especially evident in technical specialties, where curves are more numerous and angled, even more so in steeper course sections and/or in the presence of soft snow. Therefore the starting number significantly influences the final ranking of athletes in one race [20]. It is common practice to assign lower starting numbers to more competitive athletes, depending on the results achieved in the previous races; this mechanism encourages athletes to improve their performance, race after race.

The alpine skiing agonistic activity is regulated by the national federations incumbent in the localities where competitions are held; these federations in turn depend on FIS. The evaluation period for the FIS rating system is included between the second weekend of November and April 30th, for countries in the northern hemisphere (with the exception of the World Cup and European Cup races), and between July 1st and October 15th, for countries in the southern hemisphere.

In general, one alpine skiing athlete

- achieves a score in each race finished without being disqualified;
- is rated at international level, on the basis of the results achieved in the previous races²;
- obtains a starting number for future races, depending on his/her current position in the rating.

The above considerations confirm that indicators are very important in this sport, in which there are no identical race tracks and conditions, and the level of competitiveness of a race is strongly related to the competitive level of participants.

2.2 The *FP* indicator

This section describes in detail how to construct/update *FP*, which is the core of the rating system of alpine skiing. Before this, we anticipate three important features of this indicator:

1. Referring to each of the five specialties of alpine skiing, athletes are rated at international level through a relevant *FP*;
2. *FPs* are used for other practical purposes, such as: (i) determining the race *starting list*, (ii) estimating the level of difficulty of each race, based on the rating of participant athletes, and (iii) supporting the selection of teams of athletes for participating in regional/national/international FIS competitions;

² This is the reason why, according to the taxonomy by Stefani [15], the alpine skiing rating system **can be** classified as *accumulative*.

3. *FPs* of the athletes' are regularly updated taking into account the results achieved in the more recent races [21].

The block diagram in Fig. 1 summarizes the multiple role of *FP* indicators in a generic race.

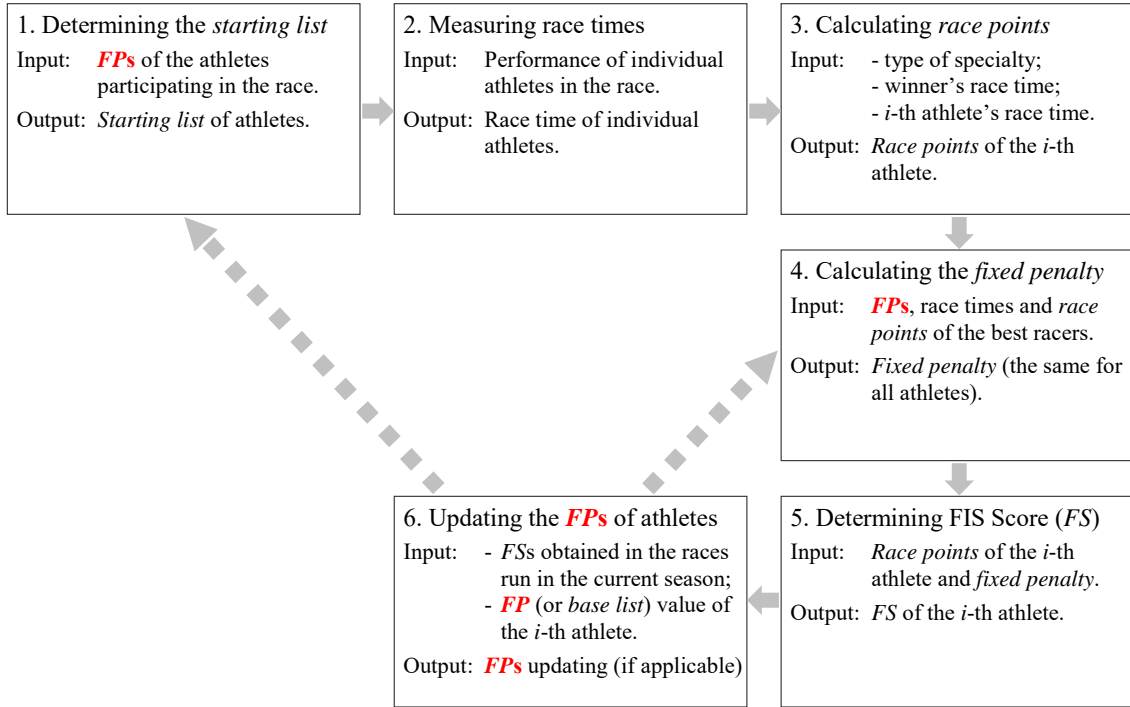


Fig. 1. Block diagram summarizing the multiple role of *FP* indicators in a generic race.

It can be noticed that the *FPs* of athletes are essential indicators for the progress of a race and the determination of the relevant results; also, race results may contribute to update the *FPs* of participants. We remark that the diagram in Fig. 1 refers to the FIS races, not necessarily the World Cup ones, which represent a special subset of the former ones. In World Cup races, athletes' *FP* rating is combined with the so-called *World Cup Start List (WCSL)*; for more information see [19]. Having said that, we now enter into a detailed description of the *FP* indicator, which is split into the following five subsections.

2.2.1 Specialties

As anticipated, *FP* indicators allow to rate the athletes involved in the races of the FIS official calendar. These races are open to professional and amateur athletes (men and ladies separately), as long as they have reached the age of 16.

FP indicators are calculated for each specialty of alpine skiing; for example, one athlete competing in all five specialties of alpine skiing will obtain five distinct *FP* indicators, which, for simplicity, can be distinguished into FP_{SL} , FP_{GS} , FP_{SG} , FP_{DH} and FP_{KB} . For each specialty, athletes are classified in increasing order: the lower the *FP*, the better the ranking; precisely, *FP* values are included between 0 (winner of the World Cup of specialty) and 999.00 points ("new entry" athlete, who has not yet attended/completed any race in that specialty).

2.2.2 Starting list

Ratings based on *FP* are used for determining the starting list in FIS races, according to the following procedure:

- The first 15 athletes on the list are those with lower *FP*s in the specialty rating; the specific starting order is decided by drawing lots;
- The remaining racers are included in the list, on the basis of their *FP* values (in ascending order).

For World Cup races, the system for determining the starting list is slightly more complicated: the first thirty positions are reserved to the top-thirty athletes in the *WCSL*, and are assigned by drawing lots; the remaining positions are assigned according to the *FP* values (in ascending order) of the remaining athletes.

2.2.3 Race results

All athletes finishing a race without being disqualified receive a *FIS Score* – hereafter abbreviated as *FS* – given by:

$$FS = \text{race points} + \text{fixed penalty}. \quad (1)$$

Let us explain the two terms in the second member of Eq. 1. *Race points* are obtained through a formula, which takes into account the time gap between the *i*-th athlete of interest and the winner of the race. This score increases with increasing the time gap from the winner, whose race points are conventionally 0. The formula for calculating race points is:

$$\text{Race points} = \left(\frac{T_i}{T_0} - 1 \right) \cdot F, \quad (2)$$

being

T_0 the winner's race time;

T_i the *i*-th athlete's race time;

F is a constant term, related to the specialty of interest, which takes into account the (inverse of the) average dispersion of race times. Speed specialties (*SG* and *DH*) generally have higher F values than technical specialties (*SL* and *GS*). F values are “adjusted” annually, taking into account the results of the races of the last seasons. For the purpose of example, the values of F used for the 2015 season are: 720, 980, 1080, 1250, and 1150 for *SL*, *GS*, *SG*, *DH*, and *KB* respectively. The rationale is that taking a certain time gap in a specialty where gaps are generally large should be less penalizing than taking the same gap in a specialty where gaps are generally low. For the purpose of example: suppose that an athlete ends a *SL* race with a time of 2'14"33 (i.e., 134.33s), while the winner obtains 2'09"19 (i.e., 129.19s). Since $F = 720$ for *SL*, the resulting race points of the athlete will be $[(134.33 / 129.19) - 1] \cdot 720 = 28.65$.

Unlike race points, *fixed penalty* has the same value for all athletes attending the race of interest. This indicator takes into account the level of competitiveness of the race, according to the (real and

purported) competitive level of the participating athletes [21]. Simplifying, this indicator is calculated as

$$\text{Fixed penalty} = (A + B - C) / 10 \quad (3)$$

where

A is the sum of the *FPs* of the best 5 racers who started the race (i.e., estimation of the competitive level of the best athletes starting the race);

B is the sum of the *FPs* of the 5 racers with the best *FPs* finishing in the top-10 positions (i.e., estimation of the competitive level of the best athletes finishing the race);

C is the sum of the *race points* of the racers in *B* (i.e., estimation of the ability of the winner to outdistance the best athletes finishing the race).

In general, the higher the level of competitiveness of the race, the lower the resulting fixed penalty. More precisely, the first two terms (*A* and *B*) – both positive – tend to decrease with increasing the level of competitiveness of the best athletes starting/finishing the race (in fact, the better the FIS rating of athletes, the lower their *FP* values). The third term (*C*), which is subtracted from the sum of the previous two, tends to increase with increasing the time gap between the winner and some of the best athletes finishing the race (which is a further indication of the level of competitiveness of the race).

In some cases, the fixed penalty calculation may be slightly more complicated, due to the introduction of some corrective parameters (e.g., *correction value*, *category adder*, etc.); for details, see [21].

Having said that, a question arises: *Does FS well reflects the real performance level of athletes in a certain race?* In Sect. 3.1 we will show that, in some cases, mediocre athletes can obtain unfairly low *FS* values, even when participating in not very competitive races.

2.2.4 *FP update*

The *FPs* of each athlete are periodically updated in the winter season (approximately on a monthly basis), taking into account the *FSs* obtained in the races attended. In particular, one athlete may improve his/her *FP* in a specialty, replacing it with the mean value of the two best *FSs* obtained (in that specialty) in the new season (as long as the resulting value is lower than that of the current *FP*). This mechanism encourages athletes to participate in as many races as possible, with the aim of obtaining as low as possible *FSs*. In fact, athletes able to reduce their *FPs* are in turn likely to obtain lower starting positions and therefore better results in the future races, according to a sort of “virtuous circle”. The fact that only the best two *FSs* are considered for updating *FP* makes the FIS ranking dynamic and open to unexpected “twists” in favour of the athletes of the moment. Nevertheless, in Sect. 3.1, we will show that the current update process is not free from grey areas or paradoxes.

For the *FPs* to be relatively *responsive*, i.e., able to reflect the current performance of athletes, not that in the previous years, these indicators are subject to a kind of “aging” process. Precisely, before a new season, FIS publishes the so-called *base FIS points list*³, which contains the current *FPs* of athletes, calculated by averaging the two best *FSs* obtained in the last season and neglecting those achieved in older seasons⁴. The decision to neglect the results obtained in older seasons is dictated by the fact that for alpine skiing, like many other sports, the skill of individual athletes can change rather dramatically from year to year [15].

2.2.5 Selection of athletes

FP indicators are also used, in conjunction with the *WCSL* rankings, to determine the number of places available to athletes from the same country in international competitions (e.g., World Cup races). For example, as regards the World Cup, every country has one place⁵; additional places are distributed depending on the number of athletes ranked within the top 60 positions of the *WCSL*. In any case, no country can have more than nine athletes and no more than one athlete over the hundredth position in the specialty ranking [21].

3. Criticism and new proposals

As shown in Sect. 2, the rating system of alpine skiing is very complex since it has to deal with several factors, which make it difficult to compare the results related to different races; in particular: (i) the uniqueness of race tracks, (ii) the level of competitiveness of a race based on the competitive level of participants, (iii) the influence of the starting order on athletes’ performance, (iv) the different dispersion of race times, depending on the specialty, and (v) the variability in the performance level of individual athletes from year to year and/or from specialty to specialty.

Criticisms of the FIS regulations – by athletes, coaches, organizers, sports fans in the field of alpine skiing – have rarely concerned the rating system, but technical features instead, such as technical equipment of athletes, specifications for the preparation of race tracks, etc. [22, 23]. This may be because the *FP* indicator “can be a complete mystery to most” [24], or as the current *FP*-based rating system seems to be relatively fair⁶ or – at least – not to produce glaring distortions in the evaluation of athletes. Nevertheless, we have identified two weaknesses:

³ The adjective “base” indicates that this list contain the *FPs* of athletes at the beginning of a new season, but they could be gradually upgraded over the new season itself.

⁴ In special cases – such as athletes with less than two races attended in the last season, injured athletes, etc. – the *FP* calculation may exceptionally consider the results achieved in older seasons, although applying some penalties; for details, see (FIS, 2015b).

⁵ Provided that the best athlete has $FP \leq 120$ [21].

⁶ The concept of *fairness* of a rating system is inherently vague. Simplifying, a rating systems can be considered fair if it well reflects the “true” value of competitors, expressed during some competitions. Unfortunately, this value is not measurable in an incontrovertible way. Despite this inevitable limitation, the literature contains some comparisons between different typologies of rating systems, according to their *predictive power*, i.e., the ability to predict the results of future competitions, based on the present results of the rating of competitors [31]. According to these comparisons,

1. Not very sound process for updating the *FPs* of a certain specialty, based on the average of the best two *FSs* obtained during the evaluation period;
2. The lack of a global rating that allows to compare athletes according to their eclecticism.

The following two subsections go into these weaknesses separately and present some proposals for improvement.

3.1 Weak *FP* update process

According to the existing rating system, any athlete obtaining two *FSs*, whose average is lower than the current *FP*, may improve it; this applies to each specialty. The rationale behind this criterion is probably to encourage a rapid turnover of athletes, so that younger and fitter athletes are able to climb the FIS standings relatively quickly. Despite this (purported) advantage, we identified (at least) three questionable aspects, as described below.

1. *Differences between specialties.* Since the athletes' propensity to participate in different races can vary greatly from specialty to specialty, it is not reasonable that the *FPs* are updated using the average of the best two *FSs*, for each of the five specialties. For example, athletes focussed on technical specialties generally attend more races than those focussed on speed specialties or *KB*. Of course, this facilitates the ascent of athletes focused in technical specialties in the relevant specialty ratings.

To quantify these differences, let us consider the data in Tab. 1, which concern the estimate of the average number of races attended annually by FIS athletes in each specialty. These estimates are obtained by randomly selecting 400 athletes (200 men and 200 ladies) from the FIS rating of each specialty and counting the average number of races concluded during the season 2015.

Tab. 1. Average number of annual races finished by athletes, in the five FIS specialties. For each specialty, data are obtained using a sample of 400 athletes (200 men and 200 ladies), randomly selected among those included in the relevant FIS ratings, at the end of the 2015 season.

Specialty	<i>SL</i>	<i>GS</i>	<i>SG</i>	<i>DH</i>	<i>KB</i>
Avg. no. of annual races	13.5	16.7	6.7	4.6	2.0

2. *Small data sample.* Since *FP* should reflect the real competitive level of one athlete in a specialty of interest, the *FP* update process based on the average of a sample of just two *FSs* could lead to dubious estimates. To better clarify this concept, let us consider the example in Fig. 2, in which two fictitious athletes attend 8 races each (6 of which are shared), obtaining 8 relevant *FSs*. The specialty of interest is *SL* and it is assumed that these races are held in the same season.

Despite Athlete 2 systematically beats Athlete 1 in the 6 races that they share (i.e., race 1, 2, 3, 5, 7, and 8), Athlete 1 has a lower mean value of the best-two *FSs*, thanks to a very low *FS* in race

6 (unattended by Athlete 2). At the end of the 10 races, both athletes improve their initial *FP* (i.e., 50 for both of them), although Athlete 1 with a new *FP* about 1.8 points lower than that of Athlete 2. We remark that this advantage may lead to climb more than 100 positions in the *SL* rating and may persist for the rest of the current season (and even the subsequent one). This paradox is due to the use of just the best-two *FS*s for determining the new *FP*.

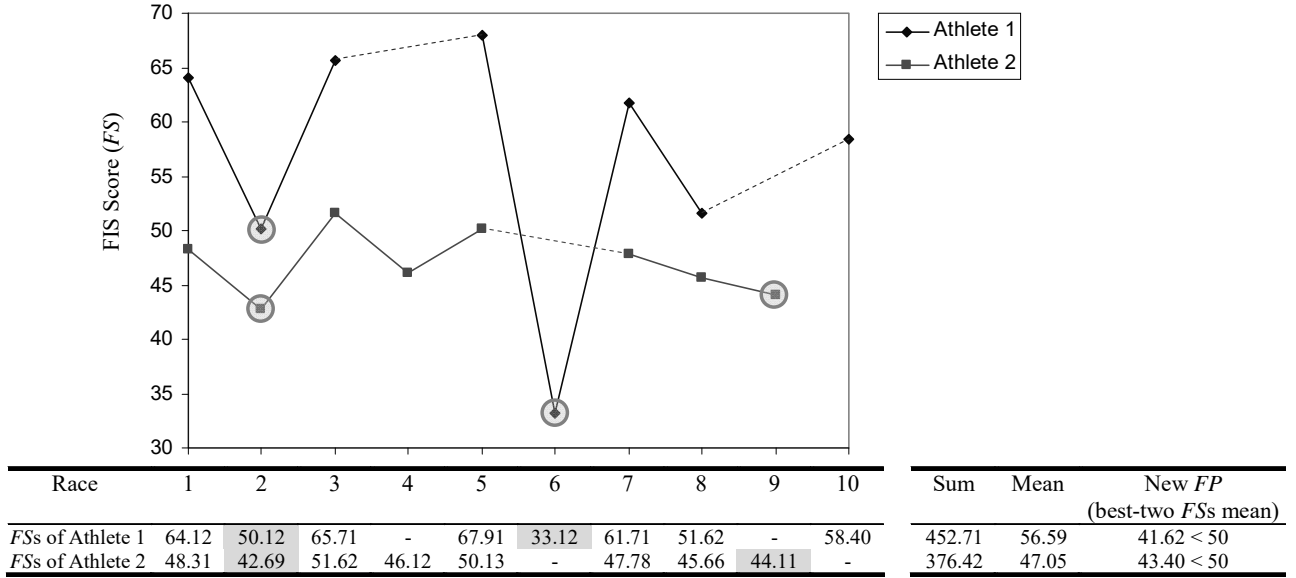


Fig. 2. *FS*s obtained by two fictitious athletes in ten *SL* races. The two best results of each athlete are highlighted in grey (both in the table and graph). The initial *FP* is 50 for both the athletes. According to the current *FP* update process, Athlete 1 improves his/her *FP* rating more than Athlete 2.

3. *Risk of opportunistic behaviour.* The not very sound update process of *FP* may results in even worse consequences; precisely, in “minor” races (i.e., races attended by athletes of medium-low competitive level), the presence of just a single top athlete (i.e., with a relatively low *FP*) who wins the race – although skiing intentionally slower than his/her full potential – can lead the other athletes to obtain abnormally low *FS*s. Fig. 3 summarizes this mechanism.

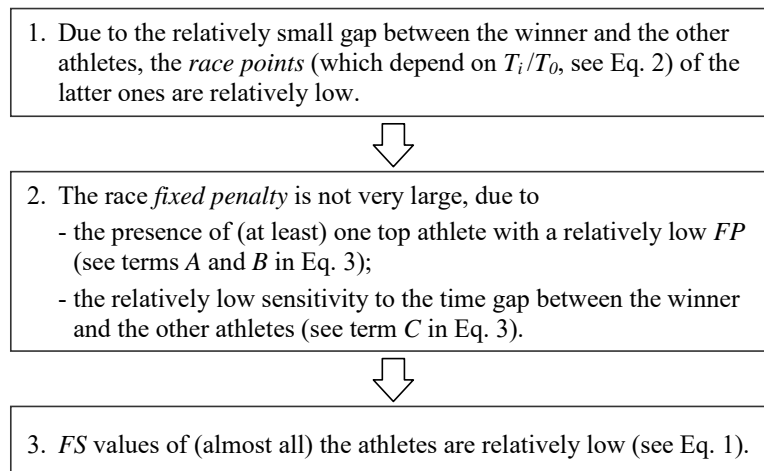


Fig. 3. Summary of the opportunistic behaviour that leads to the production of relatively low *FS*s for athletes.

Let us consider the example in Tab. 2, in which a *SL* race is attended by 11 relatively mediocre athletes (i.e., A2 to A12) and just one top athlete (i.e., A1) that in case (a) wins the race expressing his/her full potential, while in case (b) wins the race “going slower”. Race times of athletes A2 to A12 are the same in both the cases. For every athlete, it is also reported the *FP*, which is necessary to determine the race *fixed penalty*. The fixed penalty in case (b) is slightly higher than that in case (a), given that the time gaps between the winner and the other athletes are lower (see term *C* in Eq. 3). Nevertheless, the slight increase in the fixed penalty does not fully compensate the decrease in the race points of athletes. As a consequence, the *FS*s of athletes decrease considerably from case (a) to case (b).

The authors remark that this opportunistic behaviour is not pure imagination. A classical situation is that of a top athlete who helps his/her less established fellow athletes, agreeing to participate in one or more minor races, perhaps at the end of the season. Returning to the example in Fig. 2, it could be imagined that the outstanding score of Athlete 1 in race 6 (and the clear benefits that it generates for the following races) is the result of this kind of deplorable agreement.

Tab. 2. Example of a fictitious race of *SL*, attended by a top athlete (i.e., A1, the winner) and 11 relatively mediocre athletes (A2 to A12). Two scenarios are considered: (a) the winner expresses his/her full potential and (b) the winner “goes slower” than his/her full potential.

Athlete	Rank	<i>FP</i>	(a) Winner expressing his/her full potential				(b) Winner “going slower”				Δ
			Time [s]	Gap [s]	Race points	<i>FS</i>	Time [s]	Gap [s]	Race points	<i>FS</i>	
A1	1	4.33	211.10	0.00	0.00	13.71	213.60	0.00	0.00	17.14	+3.43
A2	2	25.42	214.10	3.00	10.23	23.94	214.10	0.50	1.69	18.82	-5.12
A3	3	22.64	214.30	3.20	10.91	24.62	214.30	0.70	2.36	19.50	-5.13
A4	4	28.52	214.35	3.25	11.08	24.79	214.35	0.75	2.53	19.66	-5.13
A5	5	26.81	214.50	3.40	11.60	25.31	214.50	0.90	3.03	20.17	-5.14
A6	6	30.31	214.51	3.41	11.63	25.34	214.51	0.91	3.07	20.20	-5.14
A7	7	20.97	214.55	3.45	11.77	25.48	214.55	0.95	3.20	20.34	-5.14
A8	8	23.37	214.70	3.60	12.28	25.99	214.70	1.10	3.71	20.84	-5.14
A9	9	21.78	214.78	3.68	12.55	26.26	214.78	1.18	3.98	21.11	-5.15
A10	10	32.78	214.81	3.71	12.65	26.36	214.81	1.21	4.08	21.21	-5.15
A11	11	30.01	215.00	3.90	13.30	27.01	215.00	1.40	4.72	21.86	-5.16
A12	12	21.8	DNF*	N/A	N/A	N/A	DNF*	N/A	N/A	N/A	N/A
			<i>A</i>	<i>B</i>	<i>C</i>	<i>Fixed Penalty</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>Fixed Penalty</i>	
			91.52	93.09	47.51	13.71	91.52	93.09	13.25	17.14	

- $F = 720$ since it is considered a *SL* race in the season 2015.
- Both for case (a) and (b), *race points* and *fixed penalty* are calculated using Eqs. 2 and 3 respectively. The best-five *FP*s of the athletes starting the race are those of A1, A7, A9, A12 and A3, while the best-five *FP*s of athletes finishing in the top-10 positions are those of A1, A7, A9, A3 and A8.
- Δ (in the last column) is calculated as $FS(b) - FS(a)$.
- i.e., “did not finish”.

To overcome the weaknesses of the *FP* update process and make it more sound, we suggest to:

- increase the number (hereafter abbreviated as *Y*) of the best *FS*s to be averaged for updating *FP*;
- keep *Y* roughly proportional to the average number of races attended annually by athletes (hereafter abbreviated as *X*), in a certain specialty.

The graph in Fig. 4 shows the suggested new values of Y . They are kept to 2 for DH and KB , as these specialties are on average less attended than the other ones (see the corresponding X values in Fig. 4). For the remaining specialties, the new Y values (i.e., 5, 6, and 3 for SL , GS , and SG respectively) are determined trying to keep them proportional to X^7 ; not surprisingly, the curve of the suggested Y values tends to copy that of the relevant X values ($R^2 \approx 0.98$). Although the Y values for SL and GS may seem rather large, we remark that they are roughly 30% of X , therefore it will not be so difficult for athletes to attend a number of races $\geq Y^8$.

Returning to the example in Fig. 2 and applying the suggested Y value (i.e., $Y = 5$, for a SL race), the best- Y FS s mean of Athlete 1 would be 50.99 (higher than the current value, i.e., 50), while that of Athlete 2 would be 45.27 (lower than the current one, i.e., 50). This result is probably fairer than that shown before, since it better reflects the manifested superiority of Athlete 1 to Athlete 2, which has also emerged from multiple direct confrontations.

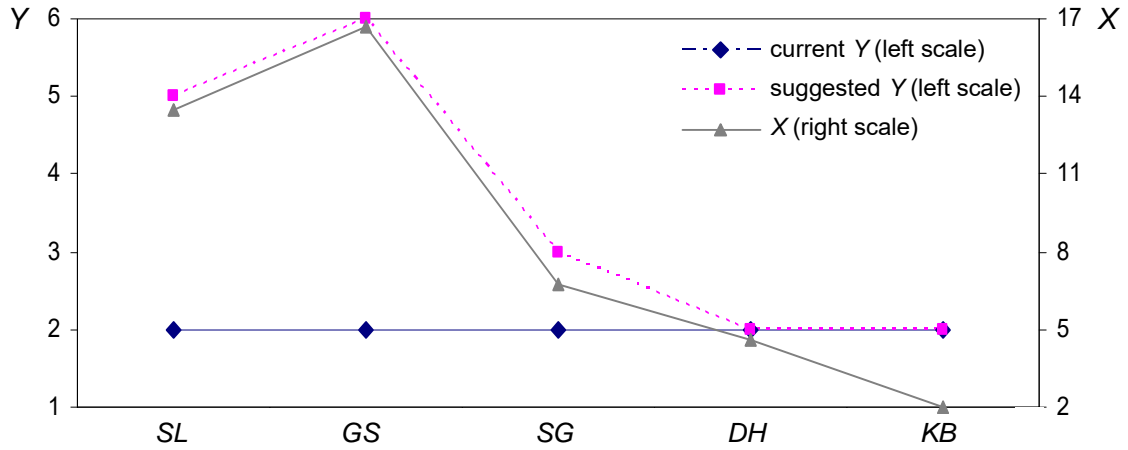


Fig. 4. Comparing the current Y values (i.e., all equal to 2) and those suggested, for updating the FP s of athletes.

Using the best Y values (instead of the best two) in the FP update process could make the FIS ratings less dynamic and uncertain. Nevertheless, the uncertainty in the outcome of a race – which is a key ingredient to keep the interest in a sport [28] – would be preserved, thanks to other measures already in use, such as drawing lots for deciding the starting order of the first 15/30 athletes of a race.

3.2 Lack of a global rating

The current rating system includes five independent indicators, i.e., FP_{SL} , FP_{GS} , FP_{SG} , FP_{DH} and FP_{KB} , which allow to compare (men or ladies) athletes, specialty by specialty. These independent FP s depict the degree of performance of athletes at *local* level, but not at *global* one. In other words, they do not allow to select the more *eclectic* athletes, i.e., those able to obtain good results in as many different specialties as possible. We believe that a global evaluation would provide a more

⁷ A similar criterion is adopted in *canoe slalom*, i.e., a sport discipline with some similarities with respect to alpine skiing [24].

⁸ Alternatively, the evaluation period can be slightly extended (e.g., from 12 to 18 months), to allow athletes to collect an adequate number of FS s.

comprehensive picture of the degree of performance of an athlete in alpine skiing. Not surprisingly, the World Cup overall rating, which is aimed at crowning the best athlete of the season, is obtained cumulating the results obtained throughout the season in all the five specialties of alpine skiing; unfortunately, this rating is limited to the (relatively few) athletes competing in World Cup races. The need for a global rating for the totality of the athletes is also corroborated by the fact that the ability of an athlete to excel in one specialty is not necessarily related to the ability to excel in another one. Speed specialties (i.e., *SG* and *DH*) tend to reward the athletes' qualities of fluidity and power, while technical specialties (i.e., *SL* and *GS*) tend to reward the qualities of speed and precision in the movements. *KB*, which includes two (or more) runs "borrowed" from *SL* and *DH*, tends to reward a sort of mixture of these qualities. These general considerations can be confirmed by a correlation analysis. Precisely, we considered the *FPs* of the first 1,000 athletes of each specialty, according to the official lists published at the end of season 2015 [26]. Obviously, among the athletes considered, some are more focused on a limited number of specialties and other ones are more versatile. As a curiosity, the athletes in the "intersection" of all five specialties (i.e., $SL \cap GS \cap SG \cap DH \cap KB$) are just 286 (see the qualitative representation in Fig. 5).

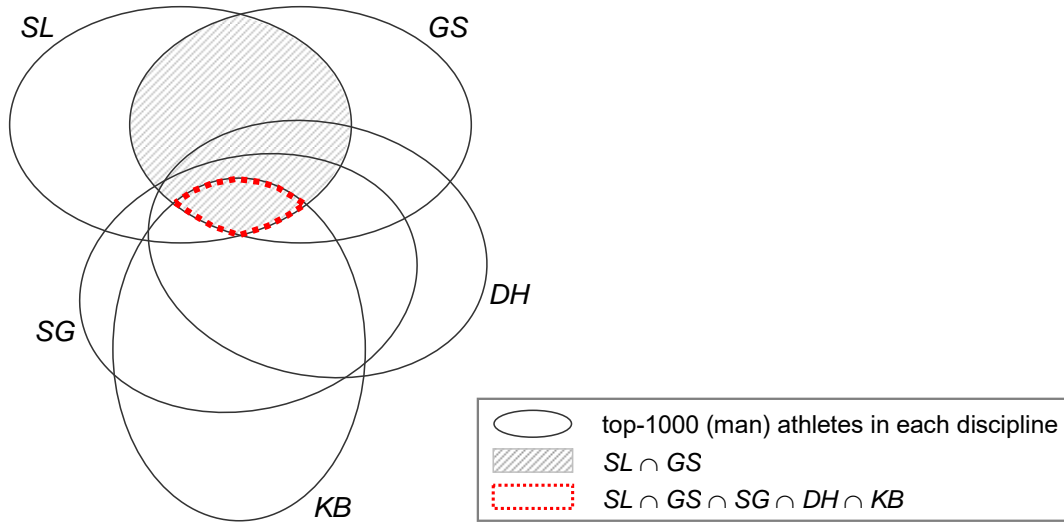


Fig. 5. Qualitative representation of the intersection between the top-1000 (man) athletes of each specialty of alpine skiing.

Next, we consider pairs of specialties and identify the number of athletes in the intersection; these numbers, which of course belong to $[0, 1000]$, give a coarse indication of the degree of affinity between pairs of specialties (e.g., *SL* and *GS*, *SL* and *SG*, etc.). A more refined indication is that provided by the Pearson correlation coefficients relating to *FP* values of the athletes in the intersection between pairs of specialties. This coefficient is included within $[-1, 1]$; values close to 1 indicate a strong positive correlation (i.e., the competitive level of athletes in the first specialty goes hand in hand with that in the second specialty), values close to -1 indicate strong negative correlation, while values close to 0 indicate absence of correlation [27]. For example, the

intersection of the athletes in *SL* and *GS* identifies 709 athletes and a Pearson correlation coefficient 0.37. Analysis results are shown in Tab. 3.

Tab. 3. Table of correlation between the *FPs* of the athletes competing in different specialties; Pearson correlation coefficient and (in brackets) the number of athletes in the intersection between pairs of specialties are reported.

	<i>SL</i>	<i>GS</i>	<i>SG</i>	<i>DH</i>	<i>KB</i>
<i>SL</i>	-	0.37 (709)	-0.02 (496)	-0.12 (403)	0.03 (875)
<i>GS</i>		-	0.26 (653)	0.04 (529)	0.19 (594)
<i>SG</i>			-	0.50 (783)	0.48 (754)
<i>DH</i>				-	0.35 (720)
<i>KB</i>					-

It is worth noting that the resulting Pearson coefficients are generally low (in absolute value), which indicates no correlation or weak correlation. This result confirms that the FIS ratings related to the various specialties are quite independent from each other, especially those relating to speed specialties with respect to those relating to technical specialties.

Let us now come to the proposal of constructing a global rating, based on a new indicator ($FP^{(G)}$), defined as the arithmetic mean of the best-three *FPs* of each athlete:

$$FP^{(G)} = \frac{1}{3} \sum_{i \in S} FP_i, \quad (3)$$

$S \subseteq \{SL, GS, SG, DH, KB\}$ being the subset of the three specialties with lowest *FP* values.

We remark that since $FP_i \in [0, 999]$ with decreasing preference, it follows that $FP^{(G)} \in [0, 999]$ with decreasing preference too⁹. The structure of $FP^{(G)}$ is justified by the following considerations:

1. Since there are two technical specialties (i.e., *SL* and *GS*), two speed specialties (i.e., *SG* and *DH*) and a hybrid specialty (i.e., *KB*), $FP^{(G)}$ penalizes pure technical or pure speed athletes, who are likely to excel at no more than two (out of five) specialties. This choice seems consistent with the spirit of the new indicator, which should reflect eclecticism.
2. At the same time, the use of only three (out of five) *FPs* does not penalize those athletes that, for some reason, have not been able to obtain good results in the totality of the specialties.

We observe that the introduction of $FP^{(G)}$ would not entail any change in the current *FP*-based rating system. In other words, $FP^{(G)}$ is simply a new aggregate indicator (which aggregates the five *FP* specialty indicators already in use), which would provide a useful synthesis assessment.

Obviously, $FP^{(G)}$ can be updated with the same frequency of the *FPs* related to the individual specialties. For the purpose of example, we applied $FP^{(G)}$ to FIS men athletes, at the end of the 2015 season, obtaining the results in Tab. 4 (limited to the top 30 positions).

⁹ Alternatively, the global rating may be expressed by the variant $FP^{(G)'}$, with increasing preference – i.e., top-ranks are represented by large numbers, in analogy with the World Cup overall rating – by introducing the simple transformation: $FP^{(G)'} = 999 - FP^{(G)}$, being $FP^{(G)'} \in [0, 999]$ with an increasing preference.

Tab. 4. Results of the calculation of $FP^{(G)}$ (bolded) for men athletes, at the end of the 2015 season (limited to the top 30 positions).

Rank	Competitor	Country	FP_{SL}	FP_{GS}	FP_{SG}	FP_{DH}	FP_{KB}	$FP^{(G)}$	World Cup (rank and points)
1	HIRSCHER Marcel	AUT	0.37	0.00	5.83	76.49	0.00	0.12	1 (1448)
2	JANSRUD Kjetil	NOR	38.73	5.47	0.00	0.00	2.53	0.84	2 (1288)
3	PINTURAUULT Alexis	FRA	4.23	1.81	4.44	37.35	0.32	2.12	3 (1006)
4	JANKA Carlo	SUI	289.85	4.71	4.47	4.61	0.00	3.03	10 (643)
5	MAYER Matthias	AUT	123.96	13.51	3.03	2.32	3.78	3.04	9 (717)
6	NEUREUTHER Felix	GER	0.00	3.85	94.02	154.98	6.00	3.28	4 (838)
7	LIGETY Ted	USA	11.68	1.69	5.95	10.75	2.67	3.44	11 (560)
8	PARIS Dominik	ITA	204.06	186.03	2.53	2.55	5.28	3.45	7 (745)
9	MUFFAT-JEANDET Vic.	FRA	4.81	3.85	48.73	49.38	1.96	3.54	12 (551)
10	BAUMANN Romed	AUT	121.83	11.07	4.94	3.79	3.14	3.96	14 (461)
11	FEUZ Beat	SUI	86.18	21.18	5.60	2.77	5.28	4.55	19 (405)
12	THEAUX Adrien	FRA	83.37	62.48	3.83	4.66	5.88	4.79	22 (365)
13	FRANZ Max	AUT	170.85	68.13	4.51	3.97	5.92	4.80	15 (457)
14	KRIECHMAYR Vincent	AUT	36.85	9.60	4.41	5.53	4.92	4.95	24 (356)
15	BANK Ondrej	CZE	17.56	7.24	13.08	5.67	3.14	5.35	41 (199)
16	VILETTA Sandro	SUI	56.12	17.91	5.99	5.77	5.17	5.64	45 (174)
17	CAVIEZEL Mauro	SUI	18.40	15.34	5.44	8.50	3.49	5.81	42 (197)
18	ZAMPA Adam	SVK	6.00	6.96	37.83	41.52	4.56	5.84	66 (95)
19	KOSTELIC Ivica	CRO	5.85	16.83	11.85	9.97	1.71	5.84	50 (155)
20	INNERHOFER Christof	ITA	42.42	27.90	5.60	5.99	5.99	5.86	46 (167)
21	WEIBRECHT Andrew	USA	282.24	18.34	4.72	8.25	5.95	6.31	40 (203)
22	COOK Dustin	CAN	118.94	5.83	3.09	10.62	36.05	6.51	30 (271)
23	MILLER Bode	USA	21.55	8.83	6.00	6.00	11.52	6.94	N/A N/A
24	REICHELT Hannes	AUT	999.00	17.73	2.83	0.98	21.76	7.18	6 (760)
25	MERMILLOD BLONDIN	FRA	10.32	16.91	7.46	24.16	4.35	7.38	78 (76)
26	SVINDAL Aksel Lund	NOR	53.64	10.75	6.00	6.00	10.25	7.42	N/A N/A
27	MURISIER Justin	SUI	9.87	7.00	18.14	46.08	5.88	7.58	96 (46)
28	KILDE Aleksander A...	NOR	28.20	12.20	5.89	15.71	5.28	7.79	75 (80)
29	GOLDBERG Jared	USA	21.53	10.16	9.00	9.29	5.35	7.88	72 (82)
30	FILL Peter	ITA	94.35	28.44	5.50	5.24	13.16	7.97	34 (234)

In light of the athletes' FP s in the individual specialties, $FP^{(G)}$ seems to provide quite reasonable results. Moreover, these results have a rather strong correlation with the World Cup results; regarding the $FP^{(G)}$ and World Cup points of the thirty athletes in Tab. 4, we obtain $R^2 \approx 75\%$). This confirms that $FP^{(G)}$ can be used for extending the eclecticism evaluation, from the (few) best athletes of the world (i.e., those attending World Cup races) to the (many) remaining ones.

4. Conclusions

This paper analyzed (part of) the rating system adopted for FIS competitions in alpine skiing. The first part of the paper provided a simplified description of the construction/use of FP s, which represents the core of the rating system. Subsequently, two weaknesses of the current rating system were highlighted and some improvement proposals suggested. The first one concerns the update process of FP , which is not very sound, and, even worse, potentially “gamable”. The proposal to vary the number (Y) of the best FS s used for updating FP depending on the specialty would make the process fairer. The second aspect concerns the lack of a global rating system rewarding the more eclectic athletes, i.e., those able to achieve positive results both in technical and speed specialties. The proposed indicator ($FP^{(G)}$) seems to reach the goal in a very simple way, using data already available. One advantage of this indicator is to make a generic FIS athlete comparable with the

athletes classified in the World Cup overall rating.

A limitation of this study is that it focused on only a few specific aspects of the current rating system. The analysis of other potentially questionable aspects – such as (i) the calculation of the *fixed penalty*, (ii) the determination of the *starting list* and (iii) the conversion of the World Cup scores into *FP* values – is left for future development of this research. Finally, we plan to make a structured comparison between the FIS rating system and those related to other sports.

References

- [1] Franceschini, F., Galetto, M., Maisano, D. (2007) Management by Measurement: Designing Key Indicators and Performance Measurement Systems. Springer, Berlin.
- [2] Parmenter, D. (2010) Key performance indicators (KPI): developing, implementing, and using winning KPIs. John Wiley & Sons, Hoboken, New Jersey.
- [3] Vyas, S., Kumaranayake, L. (2006) Constructing socio-economic status indices: how to use principal components analysis. Health policy and planning, 21(6): 459-468.
- [4] Nudurupati, S.S., Bititci, U.S., Kumar, V., Chan, F.T. (2011) State of the art literature review on performance measurement. Computers & Industrial Engineering, 60(2): 279-290.
- [5] Hauser, J., Katz, G. (1998) Metrics: you are what you measure!. European Management Journal, 16(5): 517-528.
- [6] Bouyssou, D. (2000) Evaluation and decision models: a critical perspective (Vol. 32). Springer Science & Business Media.
- [7] Lins, M.P.E., Gomes, E.G., de Mello, J.C.C.S., de Mello, A.J.R.S. (2003) Olympic ranking based on a zero sum gains DEA model. European Journal of Operational Research, 148(2): 312-322.
- [8] Hughes, M.D., Bartlett, R.M. (2002) The use of performance indicators in performance analysis. Journal of Sports Sciences, 20(10): 739-754.
- [9] Lago-Peñas, C., Lago-Ballesteros, J., Rey, E. (2011) Differences in performance indicators between winning and losing teams in the UEFA Champions League. Journal of Human Kinetics, 27: 135-146.
- [10] O'Donoghue, P. (2002) Performance models of ladies' and men's singles tennis at the Australian Open. International Journal of Performance Analysis in Sport, 2(1): 73-84.
- [11] Hughes, M., Franks, I. (2005) Analysis of passing sequences, shots and goals in soccer. Journal of Sports Sciences, 23(5): 509-514.
- [12] Clemente, F.M., Martins, F.M., Couceiro, M.S., Mendes, R.S., Figueiredo, A.J. (2015) Developing a tactical metric to estimate the defensive area of soccer teams: The defensive play area. To appear in Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, doi: 10.1177/1754337115583198.
- [13] Hughes, M., Hughes, M., Williams, J., James, N., Vuckovic, G., Locke, D. (2012) Performance indicators in rugby union. Journal of Human Sport and Exercise, 7(2): 383-401.
- [14] Barry, N., Burton, D., Sheridan, J., Thompson, M., Brown, N.A. (2015) Aerodynamic performance and riding posture in road cycling and triathlon. Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 229(1):28-38.
- [15] Stefani, R. (2011) The methodology of officially recognized international sports rating systems. Journal of Quantitative Analysis in Sports, 7(4): 1-22.
- [16] Pluta, B., Andrzejewski, M., Lira, J. (2014) The effects of rule changes on basketball game results in the men's european basketball championships. Human Movement, 15(4): 204-208.
- [17] Winchester, N., Stefani, R.T. (2013) An innovative approach to National Football League standings using bonus points. Applied Economics, 45(1): 123-134.
- [18] Mehrez, A., Pliskin, J.S., Mercer, A. (1987) A new points system for soccer leagues: Have expectations been realised? European Journal of Operational Research 28(2): 154-157.
- [19] FIS (2015a) International Ski Federation (FIS) official website. URL: www.fis-ski.com [Accessed 15 May 2015].
- [20] Lešnik, B., Podovšovnik Axelsson, E., Supej, M. (2013) Influence of the start number on elite alpine

skiing competitors' results. *Kinesiologia Slovenica*, 19(2): 17-27.

- [21] FIS (2015b) FIS Points Rules 2014/15 (version 01.07.2014). Available from www.fis-ski.com [Accessed 15 May 2015].
- [22] Ligety, T. (2012) Great Letter to FIS by David Dodge (January 3rd, 2012). Available from <http://www.tedligety.com/blog/great-letter-to-fis-by-david-dodge/> [Accessed 15 May 2015].
- [23] Gilgien M., Crivelli P., Spörri J., Kröll J., Müller E. (2015) Characterization of Course and Terrain and Their Effect on Skier Speed in World Cup Alpine Ski Racing. *PLoS ONE* 10(3): e0118119.
- [24] Bullen, R. (2012) A brief summary on how FIS points work. URL: <http://www.robertbullen.co.uk/guide-to-fis-points.html> [Accessed 15 May 2015].
- [25] ICF (2015) Canoe Slalom Competition Rules. Available from <http://www.canoeicf.com> [Accessed 15 May 2015].
- [26] FIS (2015c) Official results FIS Alpine World Championships 2015. Available from www.fis-ski.com [Accessed 15 May 2015].
- [27] Ross, S.M. (2009) Introduction to probability and statistics for engineers and scientists. Academic Press.
- [28] Vamplew, W. (2007) Playing with the rules: Influences on the development of regulation in sport. *The international journal of the history of sport*, 24(7): 843-871.
- [29] Stefani, R. T. (1999). A taxonomy of sports rating systems. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 29(1): 116-120.
- [30] Foster, K. (2012). Is there a global sports law?. In *Lex Sportiva: What is Sports Law?* (part of the series *ASSER International Sports Law Series*), pp. 35-52, R.C.R Siekmann and J. Soueck eds., Springer, Berlin (co-published with TMC Asser Press).
- [31] Stefani, R. (1998). Predicting outcomes. *Statistics in sport*, pp. 249-275, Oxford University Press, London.