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

Domenico Maisano, Andrea Botta and Fiorenzo Franceschini

Abstract

Like most sports, alpine skiing has international regulations – its regulations specifically designed by the FIS, that is, the Fédération Internationale de Ski or International Ski Federation – aimed at coordinating competitions and rating athletes. FIS points, as they are called, represent the core of the rating system, as they facilitate the rating of athletes involved in competitions for each of the five alpine skiing disciplines, that is, slalom, giant slalom, super-G, downhill, and combined. The objective of this article is to analyse the current rating system, giving special attention to the FIS points’ construction and updating system and focusing on the potential weaknesses. Two major weaknesses emerge from this analysis: (1) the questionable update process of the FIS points, based on the average of the best two results obtained in a discipline of interest, and (2) the lack of a general rating of athletes based on their eclecticism, that is, the ability to obtain good results in as many different disciplines as possible. The second part of the article presents some proposals for sorting out the above weaknesses. The description therein is supported by several practical examples, based on real and hypothetical data.

Keywords

Sport indicators, alpine skiing, rating system, FIS points

Introduction

Indicators are essential tools for monitoring and evaluating complex systems in a variety of contexts. In general, indicators are not ‘passive’ observational tools, but instead can have a profound normative effect, that is, they condition the behaviour of the system monitored. 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 rating systems. Let us consider, for example, the relatively complex rating systems in Formula One racing, tennis, artistic gymnastics, decathlons, and the like. Apart from rating systems, indicators are also used in sports for performance analysis, which is aimed at understanding the physiological, psychological, technical, and tactical demands of athletes, or even for predicting the future behaviour of sporting activity. Performance analysis – whose results are commonly used to support the activity of coaches, trainers, observers, bookmakers, and so on – is the subject of numerous contributions in the scientific literature, such as that by O’Donoghue in tennis, Hughes and Franks and Clemente et al. in soccer, Hughes et al. in rugby, and Barry et al. in road cycling.

Rating systems are generally developed by elected representatives (often aided by paid administrators) of recognized international federations, which coordinate competitions. Developing a rating system is a very delicate operation, with at least three basic requirements: (1) to find adequate ways to evaluate a set of abilities and skills of individual or peer groups of athletes, (2) to find adequate ways to determine a score for individual and/or multiple competitions (e.g. tournaments, championships), and (3) to combine the scores obtained by each competitor in order to create a corresponding rating.
Despite the profound practical implications and normative effect, rating systems have rarely been analysed from a scientific perspective. Among the few contributions in the literature, there is the criticism and/or suggestions for improvement by Pluta et al.\textsuperscript{17} in the field of basketball, Winchester and Stefani\textsuperscript{18} in American football, and Mehrez et al.\textsuperscript{19} in soccer.

The variety of sports and their great differences in terms of culture, tradition, and social and economic pressures are reflected by the variety of rating systems.\textsuperscript{15} Stefani’s\textsuperscript{20} research is particularly interesting, 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 adjustable, 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.\textsuperscript{3} However, the periodic adjustment of rating systems (e.g. the mutability of the rating system in Formula One racing) is evidence that they are far from being perfect and incontrovertible.

This article focuses on the rating system of alpine skiing racing, which represents one of the key parts of the regulations designed by the FIS,\textsuperscript{21} that is, the Fédération Internationale de Ski or International Ski Federation, founded in 1924 and promoting the practice of various disciplines of alpine skiing – that is, slalom, giant slalom, super-G, downhill, and combined. The other key parts of the FIS regulations concern the following: (1) specifications for the technical equipment of athletes (e.g. ski length or sidecut radius limits, safety protections), (2) specifications for the preparation of race tracks (e.g. minimum/maximum permissible elevation, number of gates/poles), and (3) 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).

The FIS points - hereafter abbreviated as FP (Conventionally, the expression FP will refer to the FIS points of an individual athlete in a specific discipline, while FPs will refer to multiple FIS points, related to several athletes and/or disciplines. FIS points are also used for other winter sports disciplines coordinated by FIS, such as cross-country skiing, snowboarding, and ski jumping. For simplicity, this article will exclusively refer to alpine skiing.) – represent the core of the rating system, as they facilitate the rating of athletes involved in the alpine skiing races for each discipline. 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 article is twofold. The first part is to analyse and criticize the current rating system, especially the FP construction and update process, in an attempt to answer the following research question: \textit{What are the weaknesses (if any) of the rating system in use?} The second part is to present some new proposals, in an attempt to answer the following research question: \textit{How can (at least part of) the current rating system be improved, so as to overcome the above weaknesses in a relatively simple way?}

The remainder of this article is organized into four sections. Section ‘Background information’ provides some background information on alpine skiing and the rating system in use. Section ‘Criticism and new proposals’ criticizes some aspects of the rating system and provides new proposals for improvement. The description section is supported by several practical examples, based on real and hypothetical data. Finally, the concluding section summarizes the original contributions of the article, practical implications, limitations, and suggestions for future research.

**Background information**

This section is divided into two sections which provide a synthetic description of (1) alpine skiing and (2) the rating system in use, from the perspective of the FP system. This description prepares for the analysis of some of the weaknesses of the rating system in use, in section ‘Criticism and new proposals’.

**Alpine skiing**

Modern alpine skiing is divided into five disciplines: slalom, giant slalom, super-G, downhill, and combined, hereafter abbreviated as \textit{SL, GS, SG, DH, and KB}, respectively. In a generic race, athletes have to run the same path, which is defined by gates (or poles), arranged in a variety of configurations. Races related to different disciplines generally differ in terms of track length, distance between gates, and, consequently, speed of the skier, which, for example, tends to be relatively large in \textit{DH} and relatively low in \textit{SL}. In \textit{technical} disciplines (i.e. \textit{SL} and \textit{GS}), gates are significantly closer and more angled than in \textit{speed} disciplines (i.e. \textit{SG} and \textit{DH}), thus requiring the athletes to run curves of smaller radii. \textit{KB} is a sort of ‘hybrid’ event consisting of one run of \textit{DH} and one run of \textit{SL}.

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 discipline. Race timing starts when the athlete opens the starting gate and ends when he or 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 ruts, which make each consecutive run of the track more difficult. This is especially evident in technical disciplines, where curves are more frequent 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 a single race.\(^2\) 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 the FIS. The evaluation period for the FIS rating system is held between the second weekend of November and 30 April for countries in the Northern Hemisphere (with the exception of the World Cup and European Cup races) and between 1 July and 15 October 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 the international level, on the basis of the results achieved in the previous races (This is the reason why, according to the taxonomy by Stefani,\(^2\) the alpine skiing rating system can be classified as accumulative);
- Obtains a starting number for future races, depending on his or 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.

**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, three important features of this indicator are anticipated:

1. Referring to each of the five disciplines of alpine skiing, athletes are rated at the international level through a relevant FP system.
2. FPs are used for other practical purposes, such as (1) determining the race starting list; (2) estimating the level of difficulty of each race, based on the rating of participant athletes; and (3) supporting the selection of teams of athletes for participating in regional/national/international FIS competitions.
3. FPs of the athletes are regularly updated, taking into account the results achieved in the more recent races.\(^2\)

The block diagram in Figure 1 summarizes the multiple roles of FP indicators in a generic race.

It can be noted 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 the adjustment of FPs of participants. It should be noted that the diagram in Figure 1 refers to the FIS races, not necessarily the World Cup races, which represent a special subset of the former. In World Cup races, an athletes’ FP rating is combined with the so-called World Cup Start List (WCSL); for more information, see FIS.\(^2\) This article will now enter into a detailed description of the FP indicator, which is split into the following five sections.

**Disciplines.** As anticipated, FP indicators facilitate the rating of athletes involved in the races of the FIS official calendar. These races are open to both professional and amateur athletes (males and females separately), who are 16 years of age or older.

FP indicators are calculated for each discipline of alpine skiing; for example, one athlete competing in all five disciplines 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 \(FP_{AB}\). For each discipline, athletes are classified in increasing order: the lower the FP, the better the ranking. More precisely, FP values are included between 0 (winner of the World Cup of that discipline) and 999.00 points (‘new entry’ athlete, who has not yet attended/completed any race in that discipline).

**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 FPs in the discipline 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 30 positions are reserved for the top 30 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.

**Race results.** All athletes finishing a race without being disqualified receive an FIS Score – hereafter abbreviated as FS – given by
FS = race points + fixed penalty

The following explains the two terms in the second member of equation (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

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

where $T_0$ is the winner’s race time; $T_i$ is the $i$th athlete’s race time; and $F$ is a constant term, related to the discipline of interest, which takes into account the (inverse of the) average dispersion of race times. Speed disciplines ($SG$ and $DH$) generally have higher $F$ values than technical disciplines ($SL$ 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 as follows: 720, 980, 1250, and 1150 for $SL$, $GS$, $SG$, $DH$, and $KB$, respectively. The rationale is that taking a certain time gap in a discipline where gaps are generally large should be less penalizing than taking the same gap in a discipline where gaps are generally low. Suppose, for example, that an athlete ends an $SL$ race with a time of 134.33 s (i.e. $2'14"33$), while the winner obtains 129.19 s (i.e. $2'09"19$). 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. To simplify, this indicator is calculated as

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

where $A$ is the sum of the FPs of the best five racers who started the race (i.e. the estimation of the competitive level of the best athletes starting the race), $B$ is the sum of the FPs of the five racers with the best FPs finishing in the top 10 positions (i.e. the estimation of the competitive level of the best athletes finishing the race), and $C$ is the sum of the race points of the racers in $B$ (i.e. the 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); for details, see FIS. In light of this, a question arises: does an FS accurately reflect the real performance level of an athlete in a certain race? Section ‘Weak FP update process’ will...
show that in some cases, mediocre athletes can obtain unfairly low $FS$ values, even when participating in not very competitive races.

**FP update.** The $FP$s of each athlete are periodically updated in the winter season (approximately on a monthly basis), taking into account the $FS$s obtained in the races attended. In particular, one athlete may improve his or her $FP$ in a discipline, replacing it with the mean value of the two best $FS$s obtained in that discipline 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 $FS$s. In fact, athletes able to reduce their $FP$s are in turn likely to obtain lower starting positions and therefore better results in future races, according to a sort of ‘virtuous circle’. The fact that only the best two $FS$s 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 section ‘Weak $FP$ update process’, the authors will show that the current update process is not free from grey areas or paradoxes.

For the $FP$s to be relatively responsive, that is, able to reflect the current performance of athletes, note that in the previous years, these indicators are subject to a kind of ‘ageing’ process. Precisely, before a new season, the FIS publishes the base $FIS$ points list (The adjective ‘base’ indicates that this list contains the $FP$s of athletes at the beginning of a new season, but they could be gradually upgraded over the new season itself), which contains the current $FP$s of athletes, calculated by averaging the two best $FS$s obtained in the last season and neglecting those achieved in older seasons. (In special cases – such as athletes with less than two races attended in the last season and injured athletes – the $FP$ calculation may exceptionally consider the results achieved in older seasons, although applying some penalties; for details, see FIS.23) 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.20

**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 at least one default slot (provided that the best athlete has $FP \leq 120$);23 additional slots 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 not more than one athlete over the hundredth position in the discipline ranking.23

**Criticism and new proposals**

As shown in section ‘Background information’, the rating system of alpine skiing is very complex. It has to deal with several factors, which make it difficult to compare the results related to different races; in particular (1) the uniqueness of race tracks; (2) the level of competitiveness of a race based on the competitive level of participants; (3) the influence of the starting order on athletes’ performances; (4) the different dispersion of race times, depending on the discipline; and (5) the variability in the performance level of individual athletes from year to year and/or from discipline to discipline.

Criticisms of the FIS regulations – by athletes, coaches, organizers, and sports fans in the field of alpine skiing – have rarely been concerned with the rating system, but have focused on technical features instead, such as technical equipment of athletes and specifications for the preparation of race tracks.24,25 This may be because the $FP$ indicator ‘can be a complete mystery to most’,26 or as the current $FP$-based rating system seems to be relatively fair (The concept of *fairness* of a rating system is inherently vague. To simplify, a rating system can be considered fair if it accurately 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, that is, the ability to predict the results of future competitions, based on the present results of the rating of competitors.27 According to these comparisons, some types of rating systems are said to be better than others, although this type of evaluation is characterized by a great level of uncertainty.) – at least – not to produce glaring distortions in the evaluation of athletes. Nevertheless, two weaknesses have been identified:

1. The questionable process for updating the $FP$s of a certain discipline, based on the average of the best two $FS$s obtained during the evaluation period.
2. The lack of a global rating that allows to compare athletes according to their eclecticism.

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

**Weak $FP$ update process**

According to the existing rating system, any athlete obtaining two $FS$s, whose average is lower than the current $FP$, may improve it; this applies to each discipline. 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,
the authors identified at least three questionable aspects, as described below.

**Differences between disciplines.** Since athletes’ predispositions to participate in different races can vary greatly from discipline to discipline, it is not reasonable that the FPs are updated using the average of the best two FSs, for each of the five disciplines. For example, athletes focused on technical disciplines generally attend more races than those focused on speed disciplines or KB. Of course, this facilitates the ascent in the relevant discipline ratings of the former athletes, rather than the latter ones.

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

**Small data sample.** Since FP should reflect the real competitive level of one athlete in a discipline 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 Figure 2, in which two fictitious athletes attend eight races each (six of which are shared), obtaining eight relevant FSs. The discipline of interest is SL and it is assumed that these races are held in the same season.

Despite the fact that Athlete 2 systematically beats Athlete 1 in the six races 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 has a new FP about 1.8 points lower than that of Athlete 2. This advantage may lead to the climbing of 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 only the best two FSs for determining the new FP.

**Risk of opportunistic behaviour.** The questionable update process of FP may result 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 or her full potential – can lead the other

---

**Table 1.** Average number of annual races finished by athletes, in the five FIS disciplines.

<table>
<thead>
<tr>
<th>Discipline</th>
<th>SL</th>
<th>GS</th>
<th>SG</th>
<th>DH</th>
<th>KB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. no. of annual races</td>
<td>13.5</td>
<td>16.7</td>
<td>6.7</td>
<td>4.6</td>
<td>2.0</td>
</tr>
</tbody>
</table>


For each discipline, data are obtained using a sample of 400 athletes (200 males and 200 females), randomly selected among those included in the relevant FIS ratings, at the end of the 2015 season.

---

**Figure 2.** FSs obtained by two fictitious athletes in 10 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 or her FP rating more than Athlete 2.
athletes to obtain abnormally low $FS$s. Figure 3 summarizes this mechanism.

Let us consider the example in Table 2, in which an $SL$ race is attended by 11 relatively mediocre athletes (i.e. A2 to A12) and just one top athlete (i.e. A1), in which case (a) wins the race expressing his or her full potential, while case (b) wins the race while not utilizing his or her full potential. Race times of athletes A2–A12 are the same in both cases. For every athlete, the $FP$ is also reported, which is necessary to determine the fixed penalty of the race. 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 equation (3)). Nevertheless, the slight increase in the fixed penalty does not fully compensate for 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 classic situation is that in which a top athlete helps some less established fellow athletes, by participating in one or more minor races at the end of the season. Returning to the example in Figure 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.

To overcome the weaknesses of the $FP$ update process and make it more robust, the authors suggest the following:

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

The graph in Figure 4 shows the suggested new values of $Y$. They are kept to 2 for $DH$ and $KB$, as these disciplines are on average less attended than the other ones (see the corresponding $X$ values in Figure 4). For the remaining disciplines, the new $Y$ values (i.e. 5, 6, and 3 for $SL$, $GS$, and $SG$, respectively) are determined in an attempt to keep them proportional to $X$. (A similar criterion is adopted in canoe slalom, that is, a sport

Table 2.

<table>
<thead>
<tr>
<th>Athlete</th>
<th>Rank</th>
<th>$FP$</th>
<th>(a) Winner expressing his or her full potential</th>
<th>(b) Winner ‘going slower’</th>
<th>$\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Time (s)</td>
<td>Gap (s)</td>
<td>Race points</td>
</tr>
<tr>
<td>A1</td>
<td>1</td>
<td>4.33</td>
<td>211.10</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>A2</td>
<td>2</td>
<td>25.42</td>
<td>214.10</td>
<td>3.00</td>
<td>10.23</td>
</tr>
<tr>
<td>A3</td>
<td>3</td>
<td>22.64</td>
<td>214.30</td>
<td>3.20</td>
<td>10.91</td>
</tr>
<tr>
<td>A4</td>
<td>4</td>
<td>28.52</td>
<td>214.35</td>
<td>3.25</td>
<td>11.08</td>
</tr>
<tr>
<td>A5</td>
<td>5</td>
<td>26.81</td>
<td>214.50</td>
<td>3.40</td>
<td>11.60</td>
</tr>
<tr>
<td>A6</td>
<td>6</td>
<td>30.31</td>
<td>214.51</td>
<td>3.41</td>
<td>11.63</td>
</tr>
<tr>
<td>A7</td>
<td>7</td>
<td>20.97</td>
<td>214.55</td>
<td>3.45</td>
<td>11.77</td>
</tr>
<tr>
<td>A8</td>
<td>8</td>
<td>23.37</td>
<td>214.70</td>
<td>3.60</td>
<td>12.28</td>
</tr>
<tr>
<td>A9</td>
<td>9</td>
<td>21.78</td>
<td>214.78</td>
<td>3.68</td>
<td>12.55</td>
</tr>
<tr>
<td>A10</td>
<td>10</td>
<td>32.78</td>
<td>214.81</td>
<td>3.71</td>
<td>12.65</td>
</tr>
<tr>
<td>A11</td>
<td>11</td>
<td>30.01</td>
<td>215.00</td>
<td>3.90</td>
<td>13.30</td>
</tr>
<tr>
<td>A12</td>
<td>12</td>
<td>21.80</td>
<td>DNF</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

FP: FIS points; $FS$: FIS Score; DNF: did not finish; N/A: not available.

Two scenarios are considered: (a) the winner expresses his or her full potential and (b) the winner ‘goes slower’ than his or her full potential. $F = 720$ since it is considered an $SL$ race in the season 2015.

Both for cases (a) and (b), race points and fixed penalty are calculated using equations (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.

$\Delta$ (in the last column) is calculated as $FS(b) - FS(a)$.\n
Downloaded from pip.sagepub.com at Gazi University on January 2, 2016
discipline with some similarities with respect to alpine skiing. Not surprisingly, the curve of the suggested $Y$ values tends to copy that of the relevant $X$ values ($R^2 = 0.98$). Although the $Y$ values for $SL$ and $GS$ may seem rather large, it is noted that they are roughly 30% of $X$; therefore, it will not be so difficult for athletes to attend a number of races $\geq Y$. (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.)

Returning to the example in Figure 2 and applying the suggested $Y$ value (i.e. $Y = 5$, for an $SL$ race), the best $Y$ $FS$s mean of Athlete 1 would be 50.99 (higher than the current value, that is, 50), while that of Athlete 2 would be 45.27 (lower than the current value, that is, 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.

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$^{15}$ – 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.

### Lack of a global rating

The current rating system includes five independent indicators, that is, $FP_{SL}$, $FP_{GS}$, $FP_{SG}$, $FP_{DH}$, and $FP_{KB}$, which facilitate the comparison of (male or female) athletes, discipline by discipline. These independent $FP$s depict the degree of performance of athletes at the local level, but not at the global level. In other words, they do not allow the selection of more eclectic athletes, that is, those able to obtain good results in as many different disciplines as possible. The authors believe that a global evaluation would provide a more 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 by cumulating the results obtained throughout the season in all the five disciplines 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 discipline is not necessarily related to the ability to excel in another one. Speed disciplines (i.e. $SG$ and $DH$) tend to reward the athletes’ qualities of fluidity and power, while technical disciplines (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, the $FP$s of the first 1000 athletes of each discipline were considered, according to the official lists published at the end of the 2015 season.$^{28}$ Obviously, among the athletes considered, some are more focused on a limited number of disciplines and other ones are more versatile. Curiously, the number of athletes participating in the ‘intersection’ of all five disciplines (i.e. $SL \cap GS \cap SG \cap DH \cap KB$) is just 286 (see the qualitative representation in Figure 5).

Next, pairs of disciplines are considered and the number of athletes in the intersection is identified; these numbers, which of course belong to $[0, 1000]$, give a coarse indication of the degree of affinity between pairs of disciplines (e.g. $SL$ and $GS$, $SL$ and $SG$). A more refined indication is provided by the Pearson correlation coefficients relating to $FP$ values of the athletes in the intersection between pairs of disciplines. 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 discipline goes hand in hand with that in the second discipline), values close to −1 indicate a strong negative correlation, and values close to 0 indicate an absence of correlation.$^{29}$ For example, the intersection of the athletes in $SL$ and $GS$ identifies 709 athletes and a Pearson correlation coefficient of 0.37. Analysis results are shown in Table 3.

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

Now comes the proposal of constructing a global rating, based on a new indicator \( FP^{(G)} \), defined as the arithmetic mean of the best three \( FP \)s of each athlete

\[
FP^{(G)} = \frac{1}{3} \sum_{i \in S} FP_i
\]

\( S \subseteq \{ SL, GS, SG, DH, KB \} \) being the subset of the three disciplines with lowest \( FP \) values.

It is noted that since \( FP_i \in [0,999] \) with decreasing preference, it follows that \( FP^{(G)} \in [0,999] \) with decreasing preference too. (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

\[
FP^{(G)} = 999 - FP^{(G)}
\]

where \( FP^{(G)} \in [0,999] \) with an increasing preference. The structure of \( FP^{(G)} \) is justified by the following considerations:

1. Since there are two technical disciplines (i.e. \( SL \) and \( GS \)), two speed disciplines (i.e. \( SG \) and \( DH \)), and a hybrid discipline (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) disciplines. 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) \( FP \)s does not penalize those athletes that, for some reason, have not been able to obtain good results in the totality of the disciplines.

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

In light of the athletes’ \( FP \)s in the individual disciplines, \( FP^{(G)} \) 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 30 athletes in Table 4, \( R^2 \approx 75\% \) is obtained. 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.

Obviously, \( FP^{(G)} \) can be updated with the same frequency of the \( FP \)s related to the individual disciplines. For the purpose of example, \( FP^{(G)} \) was applied to FIS male athletes, at the end of the 2015 season, obtaining the results in Table 4 (limited to the top 30 positions).

### Conclusion

This article analysed (part of) the rating system adopted for FIS competitions in alpine skiing. The first part of the article 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 proposals for improvement were suggested. The first one concerns the update process of \( FP \), which is questionable and, even worse, can sometimes be cheated. The proposal to vary the number \( Y \) of the best \( FS \)s used for updating \( FP \) depending on the discipline would make the process fairer. The second aspect concerns the lack of a global rating system rewarding the more eclectic athletes, that is, those able to achieve positive results both in technical and speed disciplines. 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 (1) the calculation of the fixed penalty, (2) the determination of the starting list, and (3) the conversion of the World Cup scores into \( FP \) values – is left for future development of this research. Finally, the authors plan to make a structured comparison between the FIS rating system and those related to other sports.

---

**Table 3. Table of correlation between the \( FP \)s of the athletes competing in different disciplines.**

<table>
<thead>
<tr>
<th>SL</th>
<th>GS</th>
<th>SG</th>
<th>DH</th>
<th>KB</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL</td>
<td>–</td>
<td>0.37 (709)</td>
<td>–0.02 (496)</td>
<td>–0.12 (403)</td>
</tr>
<tr>
<td>GS</td>
<td>–</td>
<td>0.26 (653)</td>
<td>0.04 (329)</td>
<td>0.19 (594)</td>
</tr>
<tr>
<td>SG</td>
<td>–</td>
<td>0.50 (783)</td>
<td>–0.35 (720)</td>
<td>0.48 (754)</td>
</tr>
<tr>
<td>DH</td>
<td>–</td>
<td>–</td>
<td>0.35 (720)</td>
<td></td>
</tr>
<tr>
<td>KB</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

\( SL \): slalom; \( GS \): giant slalom; \( SG \): super-G; \( DH \): downhill; \( KB \): combined.

Pearson correlation coefficients and (in brackets) the number of athletes in the intersection between pairs of disciplines are reported.
Table 4. Results of the calculation of \( FP^{(G)} \) (in bold) for male athletes, at the end of the 2015 season (limited to the top 30 positions).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Competitor</th>
<th>Country</th>
<th>( FP_{SL} )</th>
<th>( FP_{SG} )</th>
<th>( FP_{PG} )</th>
<th>( FP_{DH} )</th>
<th>( FP_{KB} )</th>
<th>( FP^{(G)} )</th>
<th>World Cup (rank and points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HIRSCHER Marcel</td>
<td>AUT</td>
<td>0.37</td>
<td>0.00</td>
<td>5.83</td>
<td>76.49</td>
<td>0.00</td>
<td>0.12</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>JANSPRUD Kjetil</td>
<td>NOR</td>
<td>38.73</td>
<td>5.47</td>
<td>0.00</td>
<td>0.00</td>
<td>2.53</td>
<td>0.84</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>PINTJURAILT Alexis</td>
<td>FRA</td>
<td>4.23</td>
<td>1.81</td>
<td>4.44</td>
<td>37.35</td>
<td>0.32</td>
<td>2.12</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>JANKA Carlo</td>
<td>SUI</td>
<td>289.85</td>
<td>4.71</td>
<td>4.47</td>
<td>4.61</td>
<td>0.00</td>
<td>3.03</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>MAYER Matthias</td>
<td>AUT</td>
<td>123.96</td>
<td>13.51</td>
<td>3.03</td>
<td>2.32</td>
<td>3.78</td>
<td>3.04</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>NEUREUTHER Felix</td>
<td>GER</td>
<td>0.00</td>
<td>3.85</td>
<td>94.02</td>
<td>154.98</td>
<td>6.00</td>
<td>3.28</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>LIGETY Ted</td>
<td>USA</td>
<td>11.68</td>
<td>1.69</td>
<td>5.95</td>
<td>10.75</td>
<td>2.67</td>
<td>3.44</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>PARIS Dominik</td>
<td>ITA</td>
<td>204.06</td>
<td>186.03</td>
<td>2.53</td>
<td>2.55</td>
<td>5.28</td>
<td>3.45</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>MUFFAT-JEANDET Vic.</td>
<td>FRA</td>
<td>4.48</td>
<td>3.85</td>
<td>48.73</td>
<td>49.38</td>
<td>1.96</td>
<td>3.54</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>BAUMANN Romed</td>
<td>AUT</td>
<td>121.83</td>
<td>11.07</td>
<td>4.94</td>
<td>3.79</td>
<td>3.14</td>
<td>3.96</td>
<td>14</td>
</tr>
<tr>
<td>11</td>
<td>FEUZ Bear</td>
<td>SUI</td>
<td>86.18</td>
<td>21.18</td>
<td>5.60</td>
<td>2.77</td>
<td>5.28</td>
<td>4.55</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>THEAUX Adrien</td>
<td>FRA</td>
<td>83.37</td>
<td>62.48</td>
<td>3.83</td>
<td>4.66</td>
<td>5.88</td>
<td>4.79</td>
<td>22</td>
</tr>
<tr>
<td>13</td>
<td>FRANZ Max</td>
<td>AUT</td>
<td>170.85</td>
<td>68.13</td>
<td>4.51</td>
<td>3.97</td>
<td>5.92</td>
<td>4.80</td>
<td>15</td>
</tr>
<tr>
<td>14</td>
<td>KRIECHMAYR Vincent</td>
<td>AUT</td>
<td>36.85</td>
<td>9.60</td>
<td>4.41</td>
<td>5.53</td>
<td>4.92</td>
<td>4.95</td>
<td>24</td>
</tr>
<tr>
<td>15</td>
<td>BANK Ondrej</td>
<td>CZE</td>
<td>17.56</td>
<td>7.24</td>
<td>13.08</td>
<td>5.67</td>
<td>3.14</td>
<td>5.35</td>
<td>41</td>
</tr>
<tr>
<td>16</td>
<td>VILETTA Sandro</td>
<td>SUI</td>
<td>56.12</td>
<td>17.91</td>
<td>5.99</td>
<td>8.50</td>
<td>24.16</td>
<td>4.88</td>
<td>45</td>
</tr>
<tr>
<td>17</td>
<td>CAVIEZEL Mauro</td>
<td>SUI</td>
<td>18.40</td>
<td>15.04</td>
<td>5.44</td>
<td>8.50</td>
<td>3.49</td>
<td>5.81</td>
<td>42</td>
</tr>
<tr>
<td>18</td>
<td>ZAMYA Adam</td>
<td>SVK</td>
<td>6.00</td>
<td>6.96</td>
<td>37.63</td>
<td>41.52</td>
<td>4.56</td>
<td>5.84</td>
<td>46</td>
</tr>
<tr>
<td>19</td>
<td>KOSTELIC Ivica</td>
<td>CRO</td>
<td>5.85</td>
<td>16.83</td>
<td>11.85</td>
<td>9.97</td>
<td>1.71</td>
<td>5.84</td>
<td>50</td>
</tr>
<tr>
<td>20</td>
<td>INNERHOFER Christof</td>
<td>ITA</td>
<td>42.42</td>
<td>27.90</td>
<td>5.60</td>
<td>5.99</td>
<td>5.99</td>
<td>5.86</td>
<td>46</td>
</tr>
<tr>
<td>21</td>
<td>WEBRECHT Andrew</td>
<td>USA</td>
<td>282.25</td>
<td>18.34</td>
<td>4.72</td>
<td>8.25</td>
<td>5.95</td>
<td>6.31</td>
<td>40</td>
</tr>
<tr>
<td>22</td>
<td>COOK Dustin</td>
<td>CAN</td>
<td>118.94</td>
<td>5.83</td>
<td>3.09</td>
<td>10.62</td>
<td>36.05</td>
<td>6.61</td>
<td>30</td>
</tr>
<tr>
<td>23</td>
<td>MILLER Bode</td>
<td>USA</td>
<td>21.55</td>
<td>8.83</td>
<td>6.00</td>
<td>6.00</td>
<td>11.52</td>
<td>6.94</td>
<td>N/A</td>
</tr>
<tr>
<td>24</td>
<td>REICHELT Hannes</td>
<td>AUT</td>
<td>999.00</td>
<td>17.73</td>
<td>2.83</td>
<td>0.98</td>
<td>21.76</td>
<td>7.18</td>
<td>6</td>
</tr>
<tr>
<td>25</td>
<td>MERMILLOD BONDIN</td>
<td>FRA</td>
<td>10.32</td>
<td>16.91</td>
<td>7.46</td>
<td>24.16</td>
<td>4.35</td>
<td>7.38</td>
<td>38</td>
</tr>
<tr>
<td>26</td>
<td>SVINDAL Aksel Lund</td>
<td>NOR</td>
<td>53.64</td>
<td>10.75</td>
<td>6.00</td>
<td>6.00</td>
<td>10.25</td>
<td>7.42</td>
<td>N/A</td>
</tr>
<tr>
<td>27</td>
<td>LURSIER Justin</td>
<td>SUI</td>
<td>9.87</td>
<td>7.00</td>
<td>18.14</td>
<td>46.08</td>
<td>5.88</td>
<td>7.58</td>
<td>96</td>
</tr>
<tr>
<td>28</td>
<td>KILDE Aleksander A</td>
<td>NOR</td>
<td>28.20</td>
<td>12.20</td>
<td>5.89</td>
<td>15.71</td>
<td>5.28</td>
<td>7.79</td>
<td>75</td>
</tr>
<tr>
<td>29</td>
<td>GOLDBERG Jared</td>
<td>USA</td>
<td>21.53</td>
<td>10.16</td>
<td>9.00</td>
<td>9.29</td>
<td>5.35</td>
<td>7.88</td>
<td>72</td>
</tr>
<tr>
<td>30</td>
<td>FILL Peter</td>
<td>ITA</td>
<td>94.35</td>
<td>28.44</td>
<td>5.50</td>
<td>5.24</td>
<td>13.16</td>
<td>7.97</td>
<td>34</td>
</tr>
</tbody>
</table>

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


