CLIMBER PERFORMANCE RATING - A NEW APPROACH TO RATING AND RANKING IN THE CLIMBING WORLD

Dale, S.[(1)], Heywood, B.[(2)], Hochmuth, N.[(3)] and Fuchslueger, U.[(4)]

[(1)] theCrag.com, Australia. scd@thecrag.com

[(2)] theCrag.com, Australia. brendan.heywood@thecrag.com

[(3)] theCrag.com, Australia. nicky@thecrag.com

[(4)] theCrag AG, Switzerland. ulf@thecrag.com

Abstract

Quantifying personal progress, comparing individual climbing achievements and using a standardised approach for structured research were the goals for the development of the Climber Performance Rating (CPR). The authors herein describe the development of CPR, a rating system heavily grounded in real world statistics and on a sound theoretical model for climbing ability and how it relates to grades. Having performed statistical analysis of almost a million ascents recorded by climbers on the climbing platform theCrag a rating system that is both predictive and relevant to climbers of all levels was developed based on three pillars: power factor, tick shift and decay. In the progress of developing CPR major questions repeatedly discussed in the climbing world were answered. The authors are confident that with CPR, climber rating and rankings can be moved away from an arbitrary to a scientifically and statistically sound approach. By providing the CPR to the climbing community the authors are not only allowing climbers to quantify their individual progress over time and to compare themselves to their colleagues and fellow competitors on a sound basis, it also opens many more possibilities for scientific research and further statistical analysis.

Keywords: training, progress, tick-shift, power factor, performance decay

Acknowledgements: The authors thank the climbing community and especially the users of theCrag for providing the data to perform this analysis and develop the climber performance rating. Thank you also to all elite climbers and trainers that provided input and reviewed the approach.

Résumé

Quantifier les progrès personnels, comparer les ascensions individuelles et utiliser une approche standardisée pour la recherche structurée étaient les objectifs du développement du Climber Performance Rating (CPR). Les auteurs décrivent ici le développement du CPR, un système d'évaluation étroitement lié aux statistiques du monde réel et sur un modèle théorique solide des capacités en escalade, et la manière dont il se rapporte aux cotations. Après une analyse statistique de près d'un million d'ascensions enregistrées par les grimpeurs sur la plate-forme d'escalade theCrag, un système d'évaluation qui est à la fois prédictif et pertinent pour les grimpeurs de tous niveaux a été développé sur la base de trois piliers: facteur de puissance, décalage des croix et régression. Dans l'avancement du développement du CPR, plusieurs questions importantes abordées à maintes reprises dans le monde de l'escalade ont trouvé réponses. Les auteurs sont convaincus qu'avec le CPR, l'évaluation et le classement des grimpeurs passeront d'une approche arbitraire à une approche scientifiquement et statistiquement pertinente. En fournissant le CPR à la communauté de l'escalade, les auteurs permettent non seulement aux grimpeurs de quantifier leurs progrès individuels dans le temps et de se comparer à leurs collègues et concurrents sur une base saine, il ouvre également beaucoup de possibilités pour la recherche scientifique et pour d'autres analyses statistiques.

Mots clés: entraînement, progrès, décalage des croix, facteur de puissance, régression des performances

Introduction

Quantifying personal progress, comparing individual climbing achievements and using a standardised approach for structured research are key challenges for the further development of the climbing sport. Since its early days as a sport, climbing relied on grading systems to rate climbing difficulty. The climber that climbed the "hardest route" was the best climber, being able to climb a harder route than last year meant, that you were progressing, comparing performance data for training or scientific analysis was and is a major challenge.

Despite the fact that climbing is becoming increasingly mainstream and will even be in the 2020 Olympics (IFSC, 2018) there is no standardised approach to rating and ranking. The few ranking systems available (8a.nu, 2018) seem fairly arbitrary and provide no statistical or scientific evidence for their system, thus only providing very limited value to measure and monitor personal progress or act as basis for scientific studies.

In light of these challenges the authors set a goal to develop a rating system heavily grounded in real world statistics and on a sound theoretical model for climbing ability and how it relates to grades. A rating system that works for climbers of all abilities from elite climbers, to newcomers and to regular weekend climbers and reflects as accurately as possible a climbers current climbing ability while providing some predictions of abilities in unknown areas. A rating system that acts at the same time as a scientifically sound ranking system for the climbing world and as a basis for structured research in a steadily growing field for the scientific community (Draper, 2015).

Methods

The first challenge in developing such a rating system lies in the accurate conversion of the many local grading systems into a harmonised system adequate for mathematical calculations. Even though there were attempts into this direction by the IRCRA (Draper, 2015), the authors decided to rely on a conversion mechanism that was developed by theCrag over the last 18 years (theCrag, 2018). This mechanism allows for the conversion of more than 20 grading systems by assigning a difficulty range for each grade for every grading system in a conversion space. This fine grained and open-ended conversion space allows for uneven difficulty jumps within a grading system (non-linear grade ranges or changes in the width of grades at a certain range) and allows also for different grading systems to have different width ranges. As such it is the basis for all performance related calculations described below and allows an easy and transparent re-conversion of the output into any local grading system a climber is familiar with and might wish to use. The second challenge in developing such a rating system is the fact that by nature only incomplete data sets are available for statistical analysis. The authors performed statistical analysis of almost a million ascents recorded by climbers on the climbing platform theCrag, taking into account that not every climber records their ascents in the same way. Some climbers may just record their successful ascents and not all their attempts, others only record the most memorable ascents, etc. In order to adjust for these facts and to eliminate other outliers the following rules were applied to the data set for analysis: ascents have to be of Ewbanks grade 20 equivalent or higher, climbers have to have recorded a minimum of 100 ascents, ascents have to be unique (no repeats), single ascents at a lower grade level are eliminated, and ascents of climbers rapidly moving through grades while finding their natural level were eliminated by applying the rule that a climber had to be a minimum of 180 days at a lower level before breaking through to the next level. In order to compensate for the limited amount of data points at the upper end of the grade scale results were verified against additional sources (99boulders, 2017) and individual feedback by selected elite climbers.

Results & Discussion

The first step in establishing a rating system that is both predictive and relevant to climbers of all levels is the formulation of the fundamental questions it has to answer. The authors identified three thereof as the foundation of any sound rating system:

- The question of the power factor, or how many routes at a particular grade should a climber do before he could expect success at the next grade level?
- The question of tick shift, or if a climber is e.g. onsighting a particular grade, what grade should they expect to red point?
- The question of decay, or how long of a break can a climber take before it affects his ability to climb at their best level?

Interestingly, all of them touch on major questions repeatedly discussed in the climbing world, obviously lacking a solid and comprehensive statistical analysis to be answered.

The power factor is based on a power algorithm, a mathematical formula where the converted ascent grade is mapped to a different mathematical plane where scores can be summed and then mapped back to grades,

$$G = log_f \sum (f^{g_n}) \qquad \sum g = 3^g + 3^g + 3^g = 3^{g+1}$$

allowing to measure sub-grade differences and progress (see Figure 1).

Figure 1: Formula for the rating grade (left) and an example for a power factor of 3 (right). G is the rating grade, g_n is the grade of an ascent, f is the power factor.

In other words, if the power factor is e.g. three then three ascents of one grade are equivalent to one ascent of the next grade. This also answers the question how much harder each grade is compared to the previous one. An analogy in the training world is Prilepin's chart (Poletayev, 2005) that applies a power law to training for weight lifting.

Climbing allows for execution of an ascent in different styles. Ascents might be of style red-point, pink-point, onsight and so on (theCrag, 2018) and also vary between climbing styles (e.g. boulder, sport climbing, traditional climbing). Tick shift is the concept that the authors apply to adjust the value of an ascent according to its ascent style per climbing style. An analogy in other sports may be the point system in ski jumping (International Ski Federation FIS, 2012) where the athlete is awarded points based on distance jumped (the grade of the route) adjusted by style points (ascent style). If a climber stops climbing or climbs at a lower level, their ability to perform at their highest level decreases over time, this is called decay. The rating system described herein aims to reflect a climber's current climbing ability. It thus can not work with an arbitrary cut off by date or period which many other rating or ranking systems have introduced but rather has to work with a time-based decay factor that is applied to a climber's rating. An analogy in other sports is for example the FIFA World Ranking (2005). Having defined the three pillars of foundation for a comprehensive Climber Performance Rating (CPR) system the authors applied observational and statistical techniques to the data set described above to determine empirical values for the power factor, the tick shifts and decay. Analysis revealed that climbers typically do 4 to 7 ascents at a lower grade before breaking through to the next grade level. Typically because it depends on the

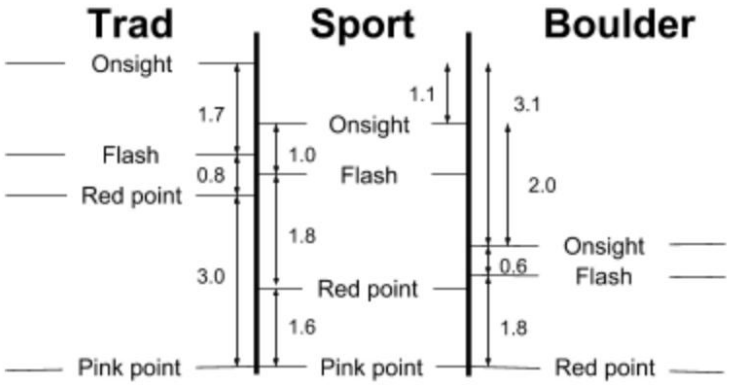
grading system as different grading systems might have different width. Sensitivity analysis revealed that ranking results are not particular sensitive to small differences in the power factor, however smaller power factors do produce noticeable quantisation effects due to the assumption that all routes of a particular grade are assumed to be exactly the same difficulty. For this reason the authors decided to choose a power factor of seven at the upper end of the statistical analysis. This means that seven ascents of a lower grade equal one ascent at the next higher grade[1].

It is commonly accepted that an onsight ascent is harder than a red-point and a red-point again harder than a pink-point ascent, but by how much? Climbers typically have a consistent grade shift between their ability to onsight and red-point a route. For example look at the shift between the green and the red dotted line in Jakob

Schubert's graph below (Figure 2).

Figure 2: Sport climbing timeline for Jakob Schubert. The green dotted line represents the onsight performance whereas the red dotted line represents the red-point performance. The delta between the two corresponds to the tick shift. The x-axis is calendar years, the solid black line represents the CPR.

The authors chose to use red-point as the reference tick type and performed statistical analysis using the CPR curves of specific tick types for climbers, comparing the average daily difference of the shift across each grade. Some additional filtering was applied to yield the results shown in Figure 3.



1: Note that the exact number may vary between grades and grading systems.

Figure 3: Tick-shifts for traditional climbing, sport climbing and bouldering. The numbers represent the delta between the respective tick types in Ewbanks grades for trad and sport climbing and V-grades for bouldering.

The CPR system applies these tick shifts for each ascent's tick type to get an adjusted rating. This means that for example Adam Ondra's recent first flash of Super Crackinette (9a+, 4734 CPR points) (Levy, 2018) has only a little less value in terms of CPR than his pink point ascent of Silence (9c, 4814 CPR points) (Carpenter, 2017).
As outlined above, the authors chose to apply a continuous decay of a climber's rating. Due to the lack of meaningful data on this topic the authors chose to infer the decay rate by looking at climbers that have plateaued, meaning climbers that continue climbing but do not progress anymore. Using this process and talking to selected trainers a decay rate of about half a grade Ewbanks per year was identified. The value of decay is certainly one that requires more investigation and input by the climbing community.
Having established the CPR system based on the three pillars described above, the authors are confident that climber rating and rankings can be moved away from an arbitrary to a scientifically and statistically sound approach. By providing the CPR to the climbing community on the climbing platform theCrag the authors are not only allowing climbers to quantify their individual progress over time and to compare themselves to their colleagues and fellow competitors on a sound basis; CPR also opens many more possibilities such as supporting scientific research, identifying wrong grading of selected routes, comparing the grading of different climbing areas on a statistical basis or evaluating a climber's grade shift from indoor to outdoor climbing. The authors encourage the scientific community to give feedback on CPR and thus contribute to help shaping the future of climbing and the scientific research about it.

References

IFSC: A large and fast-growing worldwide climbing community. (n.d.). Retrieved March 21, 2018 from https://www.ifsc-climbing.org/index.php/about-ifsc/what-is-the-ifsc/key-figures
8a.nu: Scorecard and ranking explanation. (n.d.). Retrieved March 21, 2018, from https://www.8a.nu/user/
Draper: Draper, Nick, et al., Sports Technology, 8:3-4, 88-94
theCrag: Grades. (n.d.). Retrieved March 21, 2018, https://www.thecrag.com/en/article/grades
99boulders: Retrieved March 17, 2017 from https://www.99boulders.com/hardest-boulder-problems
theCrag: Tick types. (n.d.). Retrieved March 21, 2018, from https://www.thecrag.com/en/article/ticktypes
Petr Poletayev: Olymp: 3 – 4:21-24:2005, Translated by Andrew Charniga, Jr. Sportivny Press (http://www.sportivnypress.com/documents/75.html)
FIFA: "FIFA/Coca-Cola World Rankings". FIFA. Archived from the original on 8 March 2005. Retrieved 23 June 2008.
International Ski Federation FIS. (September 2012). THE INTERNATIONAL SKI COMPETITION RULES
Carpenter: Hayden Carpenter, September 11th, 2017, Rock and Ice
Levy: Michael Levy, February 12th, 2018, Rock and Ice