Qeios PEER-APPROVED

v1: 2 January 2024

Research Article

A Probability-Based Algorithm for Evaluating Climbing Difficulty Grades

Peer-approved: 2 January 2024

Qeios, Vol. 6 (2024)
ISSN: 2632-3834

Quentin Ansel[1]

1. Université de Bourgogne, France

This paper describes a new mathematical model for the estimation of the grade of a climbing route. The calculation is based on the association of several route and boulder sections separated by rests. Contrary to other similar methods, this model introduces a probabilistic approach describing the uncertainty that one can have about the grade and the different feelings that climbers can have about a route grade. Mathematically, the problem takes the form of a data fitting with non-linear functions that depend on a few parameters, where the inputs and output of the model are probability distributions. Fitted parameters are optimized over a dataset containing information on well-established climbing routes. Several aspects of the model are commented on and studied. A short comparative study of some of the hardest routes in the world is also presented.

Corresponding author: Quentin Ansel, quentin.ansel@univ-fcomte.fr

I. Introduction

Climbing grades are important features of rock climbing since they aim to indicate the difficulty of each route [1]. They are also an important piece of information for climbers since they tell them how affordable a route is. Grades are also important for professional climbers searching for the most challenging climbs in the world. Different grade systems exist [1], but they are widely known for being both subjective and representing a kind of underlying truth of a route's difficulty [2]. The grades of sport and boulder lines take into account in a complex way the size, the shape, the spacing of the holds, the steepness, the length of the route, and sometimes the clipping positions [3]. All these aspects put constraints on the human ability to climb a rock. Based on their technical skills, strength, and flexibility, climbers process these data to estimate the difficulty. This is a complex task for which there is no real recipe (although machine learning techniques have been tried to predict the grade of boulder problems from the data of hold positions [4]). A climber can propose a grade based on their own experience within a given difficulty range, but the process is, in the end, based on personal feeling. The "official" grade of the route is obtained after multiple repetitions of this latter until a consensus is obtained. Mathematically, this can be seen as an average over all the personal grades. This observation leads to the idea that attributing a grade to a route can be modeled mathematically using probability theory [5]. In such a picture, the grade is a random variable that takes its value for each person who climbs the route. As a consequence, a route is attributed to a probability density of grade that tells us how likely a given difficulty level can be felt by someone. The "official" grade is then the expected value.

On another side, there has been an increasing interest in a systematic way of evaluating the grade of a route from the data of each sequence of this latter. Such an algorithm has existed informally more or less for many years, but the development of a well-defined and effective computation method is very recent [3,6-7]. To date, the most successful method is the one provided by the website `darth-grader.net` [3], called DGC in this paper, for Darth-Grader Calculator. DGC calculates a route grade based on the association of several route and boulder sections separated by rests. In many situations, it returns the correct grade (or at least a likely grade), but it has several drawbacks [8]. (i) It does not take into account the uncertainty or the probabilistic aspects. (ii) It cannot take into account an effect that can be called "a sensitivity to initial conditions," for which a small change in one or several sections can lead to a very different final result. (iii) In some specific situations, it leads to obvious false estimations. For example, a $7B$ boulder followed without rest by a $6a$ route (in the French grade system) is rated $7c+$ (hard) by DGC, while the reverse route, i.e., a $6a$ followed without rest by a $7B$, returns an $8a$ (soft) route. The $7B$ boulder being extremely harder than the $6a$ route, it must not make any significant difference in the final grade if it comes first or second.

This paper is devoted to the presentation of a mathematical model resolving the issues of DGC listed above. The algorithm is then used for a short comparative study of some of the hardest routes in the world. The program used for the numerical computation is not yet user-friendly nor available on the internet. But these further developments could be of great interest to many practitioners, both amateur and professional.

The structure of the article is as follows. First, the mathematical model is presented. The motivations behind the model are explained in detail, and the parameter fitting procedure is discussed (the system depends on a few parameters that must be determined first). Next, the comparative study of a few routes is presented. Finally, a short conclusion is made.

II. The Mathematical Model

The model is based on the idea that the grade of a route can be evaluated by decomposing the route into small sections and rests that can be graded individually. The association of all these building blocks must return the grade of the entire line. As will be explained in detail below, the association rule must take into account a few mathematical ingredients. But before going into these details, let us start the discussion with a quite widely believed fact [3], two $7a$ separated by a medium rest (i.e., a rest which is quite good, but not sufficient to recover completely from the previous physical effort), should be $7b$. In mathematical terms, one would translate this into [9]

$$7a \ M \ 7a = 7b \tag{1}$$

or more generally,

$$g_n \ M \ g_n = g_{n+2}. \tag{2}$$

where g_n denotes the grade n of the grade system, and M denotes a medium rest. Equation (2) is a recurrence relationship. If one knows the difficulty of the first grade levels, it is then possible, in principle, to span the entire grade system, to recover all the existing grades, and even the ones which have not been achieved yet. Equation (2) is interesting and easily exploitable for a human, but it is not easily used in a numerical algorithm, and most importantly, it fails to capture nuances. With this relation, it is not possible to model a situation like "these two routes are both $7a$, but this one seems quite a bit harder, but not enough to be $7a+$".

A. Energy Associated With a Climbing Grade

To incorporate thin nuances in the estimation of climbing difficulty levels, the mathematical grade system must be defined over real numbers $\mathbb{R}$ and not over integers $\mathbb{N}$. The real number associated with the difficulty of a route is denoted E, and it can be interpreted as the energy that must be provided by the climber to send the route. The analogy has, however, some limits since the difficulty does not depend on pure strength, but it also depends on technical abilities, or other physical abilities that reduce the intensity of the effort, such as flexibility. Then, the quantity E is not strictly a physical energy, and it is kept unitless. Another interpretation of the quantity E is one of the ranking points (a little bit like the 8a.nu scores [10]). Achieving a given grade returns a specific number of points. The harder the route, the higher the number of points.

To provide a correspondence between a grade system and the real numbers, it is necessary to fix (in a quite arbitrary way) the energy of the first grades, and then, the reference energy of all the grades can be deduced using equation (2), which becomes $2E_n = E_{n+2}$. Assuming that the first grades are given by 2, 3, and $4a$, the energies associated with these grades are assumed to be respectively $E = 1, 2, 3$. Note that for this model, it is chosen to start the grade system with these three grades since it is very difficult to make a distinction between subdivisions (given by letters) in the degrees 2 and 3. The energy associated with each grade from 2 to $10a$ is given in figure 1. The correspondence between energy and grade as given in this figure is called _the energy of reference_. We observe an exponential increase in the energy [11] (it is not linear; notice the log

doi.org/10.32388/UYLI30.2

scale in the graph). Note that a similar conclusion has been established in Ref. [12], where the authors have performed a rigorous analysis using a Bayesian analysis. A simple curve fitting enables us to determine an approximated formula to estimate the energy of reference of a climbing grade. It is given by

$$E \approx 1.21^{2n} \tag{3}$$

with n the number corresponding to the grade g_n, with $g_1 = 2,\ \ g_2 = 3,\ \ g_3 = 4a,\ \ g_4 = 4b$,... Of course, the precise values of the energies depend on the initial ones, chosen arbitrarily, but it does not change the fact that the energy increases exponentially with the grade. The exponential increase is a key feature; the precise value of the energy is not as relevant. This observation can be understood by making an analogy with the intensity of sounds [13]. To quantify the strength of a sound, we usually use the decibel unit (dB). This is a log scale, and a small difference in dB implies a large difference in the energy carried by the sound. This unit system has been chosen because humans are not sensitive linearly to the energy received from an acoustic wave. They are only able to distinguish large differences in energy (i.e., if a sound feels a little bit stronger than another one, it does not carry a little bit more energy; it carries a lot more energy). A similar situation happens for the difficulty of climbing routes.

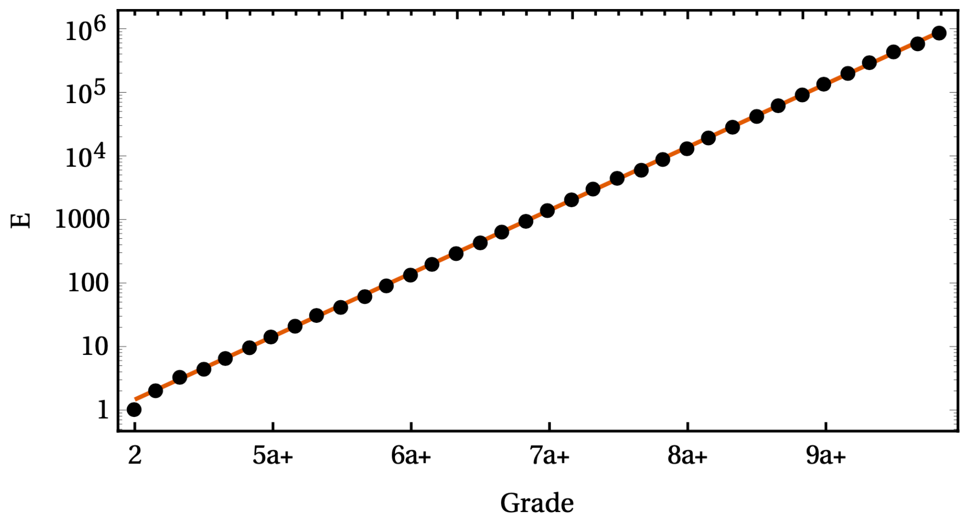

Figure 1. Energy of reference E associated with each climbing grade (French system), from level 2 to level $10a$. The curve fits as for equation $E = 1.21^{2n}$, with n the number associated with the grade g_n. The correspondence with the first grades is $g_1 = 2$, $g_2 = 3$, $g_3 = 4a$, $g_4 = 4b$,...

B. Bouldery vs. Endurance Sections

When sections are short and bouldery, it is harder to attribute a grade similar to that of an entire route. The common practice is to provide a boulder grade on short hard sections (i.e., $\lesssim 10$ moves) and to keep a route grade for long endurance sections (i.e., $\gtrsim 10$ moves) [14]. Here again, the French grade (Fontainebleau) system is employed for boulder grades. Despite its similar notation to the route grade system, the difficulty levels and the scale of the boulder system do not correspond to the route system [15]. It was initially assumed that there was a dissimilarity of one degree between boulder grades and route grades. For example, doing a $6A$ or $6B$ in Fontainebleau would be as hard as doing a $7a$ route. This correspondence remains valid for the easiest problems, but it has been modified over the years when new climbing difficulties have been introduced. For example, a $7B$ boulder corresponds to $8a$ routes, and a $8B+$ boulder corresponds to a $9a$ route. However, there still exist some discrepancies. The $8B+ \leftrightarrow 9a$ correspondence is now widely assumed for the route Hubble [16], at Raven Tor, but some very long boulders are graded $8C$, and they are assumed equivalent to $9a$ or $9a+$ routes (e.g., "la force", at Orsay's roof, "The wheel of life" in the Grampians, "Unendliche Geshichte 1+2+3" in Magic Wood) [17].

To assign an energy value to boulder grades, whose scale is already fixed by equation (2), a few assumptions are required. In the following, it is assumed that $2 \leftrightarrow 2$, $6a \leftrightarrow 4A$, $7a \leftrightarrow 6B$, $8a \leftrightarrow 7B$, $9a \leftrightarrow 8B+$, $9a+/9b \leftrightarrow 8C$, $9b/9b+ \leftrightarrow 8C+$, $9b+/9c \leftrightarrow 9A$. Here, the energy of a slashed grade corresponds to the midpoint energy between the reference energies of the two nearby grades (note that the slash will also be used with another meaning, below in the text). A linear interpolation is used for the grades in between. The last correspondences are based on Adam Ondra's opinion, as given in [15].

C. Different Association Rules

The next step of the mathematical model is to introduce different association rules of energy. The underlying purpose is to incorporate in the computation the quality of a rest between two sections and the intensity of the sections (bouldery or endurance). Following the classification of DGC [3], rests are classified into good (G), medium (M), bad (B), and non-existing (N). A good rest is a rest for which an (almost) complete recovery of the physical abilities is possible. A non-existing rest is by definition the absence of rest, a bad rest is a very short break where it is possible to shake a bit and chalk the hands, but it is not sufficient for a complete recovery. A medium rest is something in between a good and a non-existing rest..

doi.org/10.32388/UYLI30.2

Using the same convention as equation (2), an arbitrary association rule is noted

$$g_n \, R \, g_k = g_l \tag{4}$$

with $R = G, M, B, N$.

The 4 types of rests provide an easy-to-use classification. However, they have some limitations, which are the same as the climbing grades. For the mathematical model, it is preferable to describe a rest with a real number, taken in the interval $I = [0, 1]$. For simplicity, such a number is also noted R in the following. A rest with $R = 0$ corresponds to a nonexistent rest, and a rest with $R = 1$ corresponds to a good rest. Medium and bad rests are defined respectively by $R = 0.5$ and $R = 0.25$.

With these materials at hand, it is now possible to translate equation (4) into an operation on energies. It takes the form of a function $E_3 = f(E_1, R, E_2)$, with $f : \mathbb{R} \times I \times \mathbb{R} \mapsto \mathbb{R}$. Here, the following form of the function is chosen:

doi.org/10.32388/UYLI30.2

$$f(E_1, R, E_2) = E_+ + \beta \left(1 + \alpha_{1,2}(R) e^{-\frac{(\ln E_2 - \ln E_1)^4}{\sigma}}\right) E_- \qquad (5)$$

$$E_+ = \max(E_2, E_1) \qquad (6)$$

$$E_- = \min(E_2, E_1) \qquad (7)$$

doi.org/10.32388/UYLI30.2

$$\alpha_{12}(R) \tag{8}$$

$$= \begin{cases} R\alpha_{G12} + (1-R)\alpha_{N12}, & \text{if } E_1 < E_2 \\ R\alpha_{G21} + (1-R)\alpha_{N21}, & \text{if } E_1 > E_2 \\ \frac{R}{2}(\alpha_{G12} + \alpha_{G21}) + \frac{1-R}{2}(\alpha_{N12} + \alpha_{N21}), & \text{if } E_1 = E_2 \end{cases}$$

$$\beta = 1 + \beta_0 e^{-\frac{(\ln E_2 - \ln E_1)^4}{\sigma}} \tag{9}$$

where β_0 is a positive constant different from zero 0 if the second section is a boulder problem. For a long section of endurance, $\beta_0 = 0$. Moreover, $\sigma > 0$ is another parameter to determine. In total, the model has 6 parameters, $\alpha_{G12}, \alpha_{G21}, \alpha_{N12}, \alpha_{N21}, \beta_0, \sigma$. They must be determined from the data of well-established routes. The derivation method of these coefficients is described in section II F.

The association function may look a little bit complicated, but it possesses the following interesting properties:

- It returns a real number that can be assimilated to the energy of a route.
- It is not commutative, e.g., $8a\ N\ 7c$ is not equal to $7c\ N\ 8a$.
- It cannot be smaller than the largest energy associated with the two sections (e.g., the energy of $8a\ G\ 7a$ cannot be smaller than the energy of $8a$).
- The easiest section contributes significantly to the total grade only if its difficulty level is not too far from the highest one (e.g., the $5b$ in $8a\ N\ 5b$ is negligible in front of the $8a$, and the result must be $8a$, but in the case of $8a\ N\ 7c$, the $7c$ is not negligible, and the final result must be higher than $8a$.)
- The quality of the rest influences the total grade by weighting the contribution of the easiest section. The hardest one remains unaffected by a rest.
- A boulder section coming in second position is usually harder to achieve. Consequently, its energy acquires an added value, but only if the boulder problem resides in the same range of difficulty as the other section.

D. Introducing probability distributions

In the previous sections, a mathematical model of routes and boulder climbing grades has been developed. An energy quantity has been assigned to any kind of climbing sequence, and association rules have been proposed to compute the energy of a full route. Real numbers have been used to open the possibility that routes or boulders with the same grade can feel more or less hard, but this aspect has not been fully exploited yet. Moreover, the model, as it is presented so far, is not fully satisfactory since it does not well describe the various opinions on the difficulty level of a line. A possible origin for the disparate viewpoints comes from the personal limitations of climbers. For example, a reachy route is harder for small people. This leads to the idea that a grade is not always fully determined, and we have a probability of obtaining a given difficulty level. The probability distribution can be spread over different levels (for example, 50% $8a$+ and 50% $8b$), but it can also be extremely localized on a single one (e.g., 3% $7c$+, 92% $8a$, and 5% $8a$+).

In the mathematical model, the probability distribution is defined by a normalized function $p(E) : \mathbb{R} \to [0,1]$. This distribution can be mapped back to the climbing grades using the energy-grade correspondence given in figure 1. The switching between two grades g_n and g_{n+1} is defined by the midpoint energy $E_{switching} = \frac{1}{2}(E(g_n) + E(g_{n+1}))$.

With the introduction of probabilities, the energy of a route becomes a random variable. As a consequence, during the computation, we have to specify a probability distribution of energy for each climbing section and rest. Many choices of input probability distribution can be decided. For the climbing sections, they are all chosen to be uniform on the interval $[\ln(E_1), \ln(E_2)]$, where E_1 and E_2 are respectively the minimum and the maximum energy allowed on the uncertainty interval. A uniform probability distribution on an interval A is noted by u_A. The input probability for the energy E is therefore noted $u_{[\ln(E_1),\ln(E_2)]}(E)$. Logarithms are utilized to return approximately a uniform distribution on climbing grades rather than on energy. The values of E_1 and E_2 are by default the switching energies between the nearest grades, but they can also be any energy. The interval of uncertainty can be chosen arbitrarily large. This is useful for sections very difficult to estimate (i.e., the interval can cover more than a single difficulty level).

In addition to the probability distribution on energies, a probability distribution $p(R) : I \to [0,1]$ is assigned to the rests. They are also assumed to be uniform on a subinterval $[R_1, R_2] \in I$, i.e., $p(R) = u_{[R_1,R_2]}(R)$.

Now that input probability distributions are specified, the question is how the probability distribution of a full route can be computed. For each section and rest,

doi.org/10.32388/UYLI30.2

random values of energies and rests are generated numerically. With these values, the energy of the full route is computed recursively using equation (5). The process is repeated many times (typically several thousand repetitions). The final energy of the route is stored in memory for each repetition. Each value of the final energy is different, and they are realizations of a random variable corresponding to the energy of the full route. Then, the probability distribution of the line is reconstructed with the histogram of the computed energies. Note that, contrary to the initial probability distributions, the final one is not necessarily uniform.

E. Summary of the algorithm and illustration through an example

Now that all the ingredients of the mathematical model have been introduced, it is possible to summarize the calculation steps more clearly.

1. Define the probability distribution $u_{[\ln(E_1^{(n)}),\ln(E_2^{(n)})]}$ for each climbing section of the routes, as well as the probability distribution $u_{[R_1^{(n)},R_2^{(n)}]}$ for each rest of the route. The indices (n) refer to their order of appearance in the full route. The definition of these probabilities consists of specifying $E_1^{(n)}$, $E_2^{(n)}$, $R_1^{(n)}$, and $R_2^{(n)}$.
2. For $i=1$ to $i=N$, with N chosen sufficiently large (quite good values are $N=2000$ or $N=5000$), do the following steps:

- Generate random energies $E_i^{(n)}$, and random rest $R_i^{(n)}$, using the probability laws $u_{[\ln(E_1^{(n)}),\ln(E_2^{(n)})]}$ and $u_{[R_1^{(n)},R_2^{(n)}]}$.
- Compute recursively the energy of the route using $E_i^{(n)}$, $R_i^{(n)}$, and equation (5). The recursion is performed as follows. First, compute $f(E_i^{(1)},R_i^{(2)},E_i^{(3)})$, then compute, $f(f(E_i^{(1)},R_i^{(2)},E_i^{(3)}),R_i^{(4)},E_i^{(5)})$,... until all the sections and rests of the route have been used. The result of this computation corresponds to the final energy of the route, and it is noted $E_{i,\text{final}}$.

3. Compute a histogram from the data of the N values $E_{i,\text{final}}$. Normalize the histogram to get a probability distribution.

To illustrate the algorithm, the example of a hypothetical route is considered. The line is defined by:

$$g = 7a\ M\ 7a\ M/G\ 6B/6B+. \tag{10}$$

The use of slashes differs from the usual convention and its first appearance in this paper. It means that the probability distribution is not centered on a single grade, but the interval used for the probability distribution is determined by the reference energy of the two grades.

The distributions associated with each section of the route are the following:

- For a $7a$ section, $p_{7a}(E) = u_{[\ln 733.5, \ln 1075]}(E)$, with $733.5 = \frac{1}{2}(E_{6c+} + E_{7a})$ and $733.5 = \frac{1}{2}(E_{7a+} + E_{7a})$
- For a medium rest, $p_M(R) = \delta(R-0.5)$, with δ the Dirac distribution.
- For a medium/good rest, $p_{M/G}(R) = u_{[0.5,1]}(R)$.
- For a $6B/6B+$ section, $p_{6B/6B+}(E) = u_{[\ln 872, \ln 2166]}(E)$, with $872 = E_{6B}$ and $2166 = E_{6B+}$.

The probability distribution of the full route is computed in two steps, each step being given by the computation of the map f. The first application of f returns us a distribution $p_{intermediate} = f(p_{7a}, p_M, p_{7a})$, which corresponds to the distribution associated with $7a\ M\ 7a$. The second use of f provides us the final probability distribution $p_{final} = f(p_{intermediate}, p_{M/G}, p_{6B/6B+})$. Of course, this last distribution corresponds to equation (10).

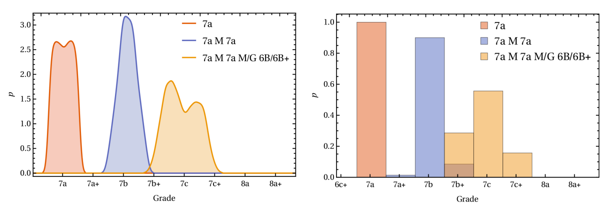

Figure 2. (Left) smooth probability distribution $p_{7a}(E)$, $p_{intermediate}(E)$, and $p_{final}(E)$. To simplify the reading, the abscissa is in a log scale, and reference energy values are replaced by their corresponding grade. (Right) probability $p(g_n)$ to find a given grade for the three cases of the left panel (see equation (11) for its definition).

The probability distributions $p_{7a}(E)$, $p_{intermediate}(E)$, and $p_{final}(E)$ are plotted in figure 2. Several observations can be made. First, the shape of the distribution is not conserved. The initial distribution is (almost) flat; the small variations are numerical artifacts induced by the finite sampling of the distribution. The second distribution is closer to a normal law, and the third one has two distinct peaks. The accumulation of uncertainties at each section of the route is responsible for a large uncertainty in the final difficulty level. For most people, it is likely a 7c route, but 7b+ and 7c+ are also possible. In fact, almost fifty percent of the climbers may disagree on the final grade. In the end, the 7c may remain the "official" grade, but no clear consensus may be achieved. Interestingly, we recover the property illustrated in equation (2). This is a key observation for the consistency of the model. Indeed, equation (2) was used to define the reference energies, but it is not used anymore in the definition of f (equation (5)). It is therefore non-trivial that f conserves the relation (2).

The continuous probability distribution $p(E)$ provides precise information on the difficulty of the line, but it does not directly answer the following question: what is the probability of getting a given grade? In other words, we would like to pass from a continuous description with the variable E to a discrete description with the variable g_n, expressed in a usual grade system. The mapping is performed with the integral:

$$p(g_n) = \int_{(E_{n-1}+E_n)/2}^{(E_{n+1}+E_n)/2} p(E)\, dE. \qquad (11)$$

These discrete probabilities are given in the right panel of figure 2.

F. Computation of the Model Parameters

The computation of the free parameters of the model is a nontrivial task. They can be determined by fitting the model with experimental data, i.e., data of well-established routes.

A collection of 33 well-characterized routes is used as a training dataset for the model (see appendix). The number of elements in the dataset is small compared to typical machine learning problems [18–19], but here, the number of parameters of the model is small, and it does not seem necessary to use datasets with thousands or millions of elements [20]. The dataset entries are chosen to represent as much as possible all the possible kinds of associations. The parameter fitting of α_{G12}, α_{G21}, α_{N12}, α_{N21}, and β_0 is performed by minimizing the cost function:

$$C = \frac{1}{N_{route}} \sum_{r=1}^{N_{route}} |\ln(E_{r,target}) - \ln(\langle E_{r,final}\rangle)| \qquad (12)$$

with, $N_{route} = 33$ the number of routes in the training dataset, $E_{r,target}$ as the target energy associated with the grade of the route r of the dataset. This energy is assumed to be known without uncertainty, and the grade/energy correspondence is the one given in figure 1. $\langle E_{r,final}\rangle$ denotes the estimated mean value of the final energy of the route r, as computed by the algorithm. This energy depends directly on the series of climbing sequences and rests of the line (which are specified in the dataset), and it also depends on the parameters to fit. Note that C can be evaluated with good accuracy using only a few random realizations of $\langle E_{r,final}\rangle$ (on the order of 50 compared to an order of 5000 to evaluate the probability distribution correctly). The cost function F is then minimized with the algorithm JAYA [21]. The result of the optimization is: $\alpha_{G12} = 0.0016$, $\alpha_{G21} = 0.2873$, $\alpha_{N12} = 0.8485$, $\alpha_{N21} = 0.3828$, and $\beta_0 = 0.3754$.

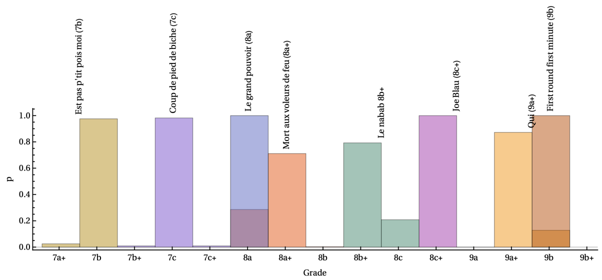

Figure 3. Probabilities $p(g_n)$ predicted by the algorithm. The target (official) grades of the routes are given after their names in brackets. As we can observe, the probabilities are always highly concentrated around the target grade.

Note that 4 decimals are taken into account to obtain a fine result. However, extensive numerical tests have shown that the parameter sensitivity is quite low. Restricting the parameters to one or two decimals can provide quite good results.

The parameter σ is not determined like the other parameters. It is fixed by hand to $\sigma = 20$. The dataset does not allow us to determine it very accurately, but having an accurate value does not seem strictly necessary. It is chosen large enough so that only nearby grades can influence each other, but not too much, to limit the influence of a small grade on a large one.

With this kind of mathematical model, which is aimed at predicting new data from the basis of known ones (i.e., like in machine-learning techniques), another dataset called the validation set is used to verify the model's predictive power. In typical machine learning systems, the training dataset is around 80% of all the data, and the validation dataset is the remaining 20%

doi.org/10.32388/UYLI30.2

[18-19]. Here, the validation set is composed of 8 routes with different grades (see appendix for details). The probabilities $p(g_n)$ of these routes, as predicted by the algorithm, are given in figure 3. We observe a very good correspondence between the target grades and the highest-grade probability predicted by the algorithm, hence validating the relevance of the model.

G. Further comments on the assumptions of the model

To construct the model, several assumptions are required. They have already been more or less discussed in the text, but they are commented on again with more details.

a. Initialization values of energies

The energy scale requires initial values to be generated. They are chosen arbitrarily. The choice of initial value can potentially have an impact on the final result. However, we observe that an asymptotic behavior is reached quickly (see Fig. 1), and thus discrepancies between different initializations should be mostly visible on the easiest grades. Furthermore, the fact that the model parameters are optimized for a given energy scale tends to limit the dependence of the model on the energy values. The most important thing is the relation between the different levels, not their precise values. Moreover, the fact that probability distributions are used over a very wide range of energy tends to suppress issues related to fine-tuning of the energy scale. Indeed, with the probability approach, it is not necessary to get "the true" grade (which is an inadequate notion), but a plausible grade distribution, given the knowledge of prior data. In addition, the probability distributions are typically large in the space of variables, so small variations of these latter tend to be negligible.

b. Construction of the energy scale based on route data

The route energy scale is constructed recursively using an empirical approach, and the boulder energy scale is mapped non-linearly onto the route scale. This approach is motivated by an empirical relation of recurrence. The reverse procedure, which would consist of first defining the boulder energy scale, is less obvious since boulders are very short and it can be harder to decompose the problem into well-identified sections (although it is not impossible). Moreover, there is no real proof that the recurrence relation holds exactly true at any climbing level (of routes). It may be interesting to make the relation grade-dependent. Nonetheless, as already pointed out, the most important thing is not the energy scale by itself, but the association rule, which is fitted from experimental data. Again, this procedure tends to limit modelization artifacts.

c. Correspondence between rest quality and percentage of rest

Following DGC convention, 4 types of rests are used as simple user inputs for the algorithm. However, the core of the model uses a continuity of different rests, and similarly to climbing grades, a correspondence between the discrete and the continuous descriptions must be imposed. Again, this contains a part of arbitrariness, but the precise values are motivated by empirical assumptions so that the model can reproduce qualitatively the desired output. The use of $R = 0$ and 1 for respectively N and G rests are obviously the most reasonable assumptions. The value $R = 0.5$ for a M rest is related to the definition of medium: something halfway between two sizes. The value $R = 0.25$ for a B rest follows the same idea as the medium case, but now the extremities are given by N and M rests. Similarly to grades, the use of probability distributions and the use of a fitted model tend to suppress the artifacts induced by these arbitrary definitions.

d. Definition of the parameterized association rules

The definition of the association rule depends on some arbitrary assumptions (e.g., the use of Gaussian functions). Other functions may produce very similar outputs. This choice is motivated by the fact that f must be a map $\mathbb{R} \times [0, 1] \times \mathbb{R} \longmapsto \mathbb{R}$, it must be non-commutative, it is strictly increasing with respect to the first and last arguments, it must be sensitive mostly around the largest grade, depending on the rest and the nature of the section, and the output is weighted to describe diverse situations when sections are combined together. The adjusted parameters are used to reduce the part of arbitrariness. This may be enhanced with models using additional parameters, but this would require a larger dataset and a more detailed description of the routes in the databases.

e. Uniform distributions

Uniform distributions are used as a default choice of probability distribution. This is a basic assumption in cases where no real information on a true probability distribution is available. It is possible to replace the uniform distributions with more realistic ones if they are known.

III. Application: comparison of some of the world's hardest routes

To finish this paper, the algorithm is applied in a short comparative analysis of some of the hardest routes in the world (at the moment of writing this paper).

The routes are Silence (9c) [17], DNA (9c) [22], B.I.G. (9c) [23], Sleeping Lion (9b+) [24], and Excalibur (9b+) [25]. The lines have been visited by several climbers, but since their first ascent, none has been repeated yet [26]. The first ascents were made respectively by Adam Ondra (2017), Seb Bouin (2022), Jackob Schubert (2023), Chris Sharma (2023), and Stefano Ghisolfi (2023). For each of these climbers, these routes may be considered their hardest achievement so far.

The grade density probabilities are computed with the following sequences:

- Silence: $8b\ G\ 8c/8C+\ \ B/M\ 8A/8B\ B\ 7C+\ \ G\ 7B$,
- DNA: $8c/8c+\ \ M/G\ 8A/8B\ B\ 8A+/8B\ N/B\ 8c+$,
- B.I.G: $9b\ N/B\ 8A/8B\ N\ 8a/8a+$,
- Sleeping Lion: $7B/7B+\ \ G\ 8A\ M/G\ 8A\ B/M\ 8A/8A+\ \ N/B\ 8A$
- Excalibur: $8B\ B/M\ 8C$.

It should be noted that some sequences of Silence and DNA are slightly different from the ones initially given after the first ascent, due to additional opinions given by other climbers. For Silence, the first crux may be slightly harder than $8C$ since it has been tried by at least 3 other climbers, but only A. Ondra has done it yet. Concerning the second crux of the route, it may be easier than initially expected. For DNA, after the comments on J. Schubert, the boulder problems can be harder than the grades given by S. Bouin, especially the first one, which has a reachy move.

The probability density $p(E)$ of each route is given in figure 3. The grade probabilities $p(g_n)$ are given in table III. We observe that all the routes are predicted in the range $9b+$ / $9c$, but only Silence can be considered as a pure $9c$ since it has a 100% probability on this grade. The second hardest route is DNA but with a $9c$ probability around 55%. All the other routes are more likely to be $9b+$. However, several comments on these results can be made.

First of all, Sleeping Lion is very close to DNA; the difficulty also seems clearly above $B.I.G.$ Then, Chris Sharma could have announced $9c$ for this route. The frontier is very close. If we consider that the two middle rests are respectively M and B, the $9c$ probability goes to 69%. If the $8A$ boulder problems are all replaced by $7C+/8A$ boulders and the $8A/8A+$ boulder is replaced by a $8A$ boulder, the grade falls to $9b+$ with a probability of 99%.

As a second comment, the density probabilities of B.I.G and Excalibur are quite similar. They are also very large, and this complicates the choice of a grade. For B.I.G, the arguments were clearly asserted by Ondra and Schubert: the route is very long, and the crux feels very hard coming from the ground. Moreover, the route remains more or less humid all the time, and this is a source of additional difficulty. The $9c$ possibility is small, but for these two climbers, the route feels harder than all the other $9b+$ they have made. For Excalibur, this is quite the opposite. The route is short, and so is the duration of the effort. The key to success relies on the middle rest. For a medium rest, the grade is $9b+$ at 72% (still, it remains a hard one), but for a very bad or nonexistent rest (let us say, there is no rest at all), the route is $9c$ at only 60%.

To conclude this small application of the algorithm, it enables us to render quite accurately the true difficulty of a route. A clear consensus on the grade of Silence may be easily achieved, but this may not be the case for DNA, B.I.G., and Sleeping Lion. Their grades could be adjusted by the next repeaters. With these lines very difficult to grade, the maths do not always return the final conclusion. One also has to answer the question, "Which grade do you want for this route?" As A. Ondra said about B.I.G [27]: "I think this amazing route deserves a nice grade, too." This is not the first time that such a thing has happened! Maybe almost all climbing crags have amazing routes with nice grades?

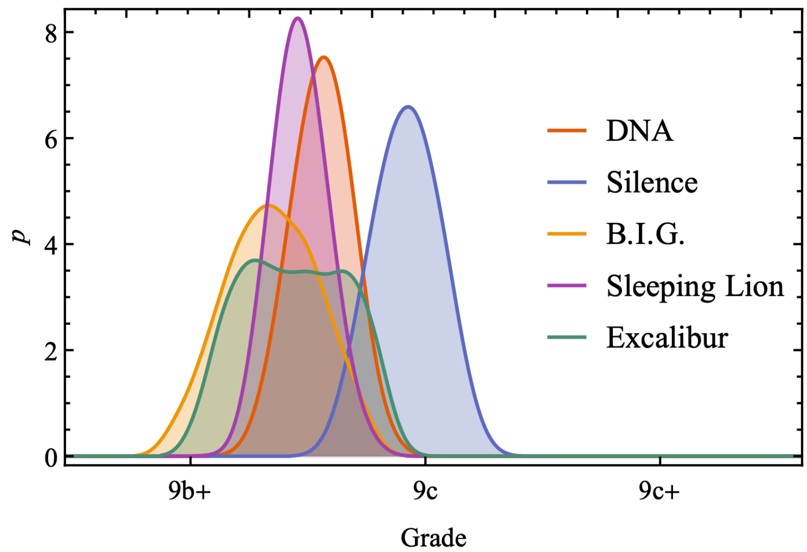

Figure 4. Probability $p(E)$ to obtain a given grade for DNA, Silence, B.IG., Sleeping Lion, and Excalibur. To simplify the reading, the abscissa is in a log scale, and reference energy values are replaced by their corresponding grades.

	$9b+$	$9c$
DNA	45%	55%
Silence	0%	100%
B.I.G.	84%	16%
Sleeping Lion	77%	23%
Excalibur	65%	35%

Table I. Probability $p(g_n)$ to obtain a given grade for DNA, Silence, B.IG., Sleeping Lion, and Excalibur.

IV. Conclusion

In this paper, a mathematical model for the estimation of climbing route grades was presented. The calculation is based on the association of several route and boulder sections separated by rests. The model incorporates several key ingredients. First, the level of difficulty and the quality of a rest are quantified using real numbers. They are interpreted respectively by the amount of energy that must be provided to achieve the route (or in some sense the number of rewarding points for sending a given difficulty), and a percentage of the quality of the rest. Based on real-life observations, a correspondence between grade/energy and "type of rest"/"percentage of rest" is made. Interestingly, the energy scales exponentially with the grade. A second important feature is the association rule that allows us to combine the energy of sections separated by rests into the energy associated with the full line. Using the grade/energy correspondence, the grade of the full section can be deduced. The association rule depends on a few parameters derived by a fitting procedure. This latter is performed utilizing an optimization algorithm that computes the best set of parameters reproducing the difficulty level of a dataset of well-characterized routes. The fitted model is then checked using a second dataset. The last important point is the incorporation of probability densities in the model. The use of probabilities is motivated by the various difficulty levels that climbers may witness on the same line. In this setting, the grade attributed by each climber is a random variable, and the "official" grade, the grade given in a guidebook, for instance, is the expected value.

As an illustrative example, the algorithm was applied to compare a few of the most difficult routes in the world: Silence, DNA, B.I.G., Sleeping Lion, and Excalibur. This short comparative study emphasizes very different situations: cases where there is no possible uncertainty

doi.org/10.32388/UYLI30.2

about the grade (everybody will agree on it), and cases where a consensus is far more difficult. Several motivations used to choose one grade rather than another have also been reviewed.

The algorithm has been extensively tested on numerous routes, which are not necessarily included in databases (if they could not be considered as "references" for the grade), and the algorithm has always returned a plausible result. In some way, the use of probabilities reduces the potential failures since it always returns the most likely climbing grades given some prior knowledge. The prior knowledge is constituted of both the optimized model and the accuracy of the route description. It is not guaranteed that the algorithm always returns the most accurate description of the route. Indeed, the predictions inherit all the usual limitations of data fitting. Nevertheless, this model can be further improved to take into account additional features. For example, the addition of very physical sections may not have the same impact as the addition of more technical sections. Moreover, the fitting parameters may also be improved with a bigger dataset. Another possibility is to replace equation (5) with a neural network. A neural network usually has a very large number of free parameters that must be fixed with the training procedure. This large number of degrees of freedom in the model can enable us to reduce many assumptions (such as the form of eq. (5)). Basically, it would learn the association rule from nothing else than the dataset. This may be very interesting, but it would require a very huge dataset and heavy numerical computations for the training.

As a final remark, the algorithm can also be used to predict the grade of boulder problems if the moves or the sections are graded individually. This kind of computation is straightforward; it is only sufficient to replace the route energies of reference with the boulder energies of references in the computation of the histogram. As an example, the algorithm predicts a 100% probability on $9A$ for _Burden of Dreams_, which is the correct grade according to the community (for this computation, the following sequence is used: $8B+\ N\ 7C\ N\ 7C+\ N\ 7A/B\ 8A$).

Appendix A: Datasets

The dataset used for the parameter fitting is given in Table II. The dataset used to validate the fitted parameters is given in Table III. The dataset is restricted to routes that are commonly assumed to be a reference for the grade (the popularity of the route can be either local or international). Some of the dataset entries are adapted from [3]. In such a case, the name of the route is written in italics.

Route Name	Location	Target Grade	Sequences
Silence	Flatenger	9c	8b G 8C/C+ B/M 8A+/8B B 7C+ G 7B
Change	Flatenger	9b+	7C M/G 8B/+ G 7C N 7a/b G 9a
Bibliographie	Céüse	9b+	8b+ G 8A+/B B 9a
La Dura Dura	Oliana	9b+	7C N 7C N 8B+ M 8c/c+ M 8a
Zvěřniec	Holštejn	9b+	9a+ G 8B+ N/B 7A
Vasil Vasil	Sloup	9b+	8b N 8B+ N 7A
Taurus	Byci Skala	9b	8C+ N 8b
Chaxi Raxi	Oliana	9b	8B+ G 9a/a+
Move	Flatenger	9b+	9a B/G 8B+
Beyond Integral	Pic St. loup	9b	9a+ G 8A+
Biographie	Céüse	9a+	8c+ M 7C+/8A M 7c
Papichulo	Oliana	9a+	8b/b+ G 8c/c+ B 8b+/c
Les Yeux Plus Gros que l'Antre	Russan	9a+	8c G 9a
Pornographie (with kneepad)	Céüse	8c+	8c B 8b
Mange ta Soupe	St. Pancrasse	8c	8a+ N 7B+
Le minimum	Buoux	8c	7c N 7C M 7c
Eliot Le Pilote	St. Ange	8b+	8a+ G 7B
Check Up Gros	Parmelan	8b+	7B+ B 8a+
La Rage de Vivre	buoux	8b+	8b G 8a+
Clafoutis	Mouxy	8a+	7c+ M 7c+
Mo..Mo..Motus	La Baume du Syratu	8a	7b+ G 7c+
Gras Mouillé	Romeyer	8a	7c M 7c
Appel au Roi	Barmaud	7c+	6C N 7a+/b
Les samares	Barmaud	7c	7a/a+ G 6B G 6B G 6c
Écureuil Plongeur	Barmaud	7c	7b M 7b
La Chose	Céüse	7c	6C+ G 7a
Télémaque Contrôle	Saint Léger	7b+	7a+/b G 7a+
Galaxy	Céüse	7b+	6C G 7a+
Le Concasseur	Tamée	7b+	6c G 6B N 7a
Miaou	Mouxy	7b	6c+ B 7a
Une Ténébreuse Affaire	Barmaud	7a	6c G 6c
Œstrogène	La Brême	6b+	6b G 6a+
Les Aigles	La brême	5c	5a+ G 5b+

Table II. Dataset used for the parameter fitting.

Route Name	Location	Target Grade	Sequences
First Round First Minute	Margalef	9b	7C+ N/B 8A+ N/B 7B+/C
Qui	Geisterschmiedwand	9a+	8c M 8A+
Joe Blau	Oliana	8c+	8a+ G 8b+ M 8b
Le Nabab	Saint Léger	8b+	8a M 8a M 8a G 7b+
Mort aux voleurs de feux	Le Quint	8a+	6a+ N 7B G 6B
Le Grand Pouvoir	Le Quint	8a	7b N/B 6C/C+
Coup de Pied de Biche	Rurey	7c	6B+/C G 7b
Est pas P'tit Pois Moi	Rurey	7b	6a G 7a+ M 6c

Table III. Dataset used to validate the fitted parameters.

Endnotes

[1] N. Draper, 14 climbing grades, The Science of Climbing and Mountaineering, 227 (2016).

[2] D. Scarff, Estimation of climbing route difficulty using whole-history rating, arXiv preprint arXiv:2001.05388 (2020).

[3] DARTH GRADER - the grade calculator ().

[4] A. Dobles, J. C. Sarmiento, and P. Satterthwaite, Machine learning methods for climbing route classification, Web link: http://cs229.stanford.edu/proj2017/finalreports/5232206.pdf (2017).

[5] Y. Caumel, Probabilités et processus stochastiques (Springer Science & Business Media).

[6] L. Kempen, A fair grade: assessing difficulty of climbing routes through machine learning, Formal methods and tools, University of Twente (2018).

[7] M. Andric, I. Ivanova, and F. Ricci, Climbing route

difficulty grade prediction and explanation, in IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (2021) pp. 285–292.

[8] At the moment of writing this paper, October 2023.

[9] The notation is similar to the one introduced in [3], but the underlying mathematical structure is different.

[10] Ranking · sportclimbing - 8a.nu ().

[11] This is a key point that differs significantly from the approach of DGC. In the technical notes of the website, it is written that grades follow a linear relation. This interpretation is wrong. Relation (2) does not lead to something linear. However, the calculation method of the website is correct in the sense that relation (2) is obtained by the calculator.

[12] A. Drummond and A. Popinga, Bayesian inference of the climbing grade scale, arXiv preprint arXiv:2111.08140 (2021).

[13] E. Hecht, Physics, Physics Series (Brooks/Cole).

[14] This decomposition using both route and boulder grades is already used in [3], however, a mapping chosen here is slightly different.

[15] Route/boulder grades discrepancy according 8a and ondra - 8a.nu news ().

[16] Buster martin smashes 10th ascent of hubble | climber magazine ().

[17] index - escalade9 ().

[18] C. Bishop, Pattern Recognition and Machine Learning, Information Science and Statistics (Springer-Verlag).

[19] S. Marsland, Machine Learning: An Algorithmic Perspective (CRC Press) google-Books-ID: n66O8a4SWGEC.

[20] Different tests with different sizes of databases have been performed, and increasing the size of the training dataset would not change the result significantly; it would only adjust the parameters slightly.

[21] R. Venkata Rao, Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems, 7, 19.

[22] Grimper, Seb bouin enchaîne DNA, possible second 9c de l'histoire...

[23] Jakob schubert (@jakob.schubert) • photos et vidéos instagram ().

[24] D. Miller, Exclusive interview: Chris sharma makes 5.15c FA in siurana.

[25] P. Délas, Stefano ghisolfi libère excalibur 9b+ ! - stefano ghisolfi frees excalibur 9b+ !

[26] 8a.nu: Global climbing news ().

[27] B.\. M. www.benes michl.cz, My reaction on grading b.i.g. 9c by jakob schubert.

Supplementary data: available at https://doi.org/10.32388/UYLI30.2

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.