Research Article

Christoph Bartneck* and Elena Moltchanova

Fair World Para Masters Point System For Swimming

https://doi.org/10.1515/jqas-2023-0051
Received May 18, 2023; accepted December 8, 2023

Abstract: A fair and inclusive competition depends on a scoring system that takes all relevant factors into account. We analysed the current World Para Point System for swimming and identified several theoretical and practical disadvantages. We propose and test a Fair World Para Point System that not only improves the algorithm, but also extends it to accommodate for the age of the athlete. It also provides a method to break point ties. This will enable para masters swimmers for the first time to compete fairly with each other. We also develop and publish tools that enable event organisers to directly use the Fair World Para Point System.

Keywords: para, swimming, point, masters, age, disability

1 Introduction

A fair competition is essential to all sports [32]. Given the wide range of disabilities present in para athletes, classification systems have been introduced to enable athletes with different impairments to compete with each other. These classification systems are inherently difficult to define [23] and differ across sports. For para swimmers, a point system has been introduced that is based on disability classifications [18, 35]. Swimmers are grouped into ten physical impairment sport classes (S1-S10), three visual impairment classes (S11-13) and one intellectual impairment class (S14). The lower the number, the higher the activity limitations. The sport classes S1-S10 cover a wide range of disabilities and limitations.

A swimmer in the S1 class has “a significant loss of muscle power or control in legs, arms and hands. Some athletes also have limited trunk control. This may be caused by tetraplegia, for example. Swimmers in this class usually use a wheelchair in daily life.”[1] Swimmers in the S3 class typically have amputations of both arms and legs or has no use of their legs and/or trunk or has a sever lack of control in all limbs. Swimmers in the S10 class have minimal physical impairments, such as the loss of a hand.

The fairness of the system is reliant on the definitions for these classes and much debate on the classification criteria is present [9, 37]. Burkett et al. [5] show that the relationship between the impairment and swimming performance is inconsistent. Alternatives to the current practice have been proposed based on extreme value theory [14] and partial least squares regression [17]. In 2018 the classification system has been revised, but its impact is still somewhat unclear [28]. Still, as long as there are enough athletes in one class they can at least compete against each other in a relatively fair way within that class. Although the classification systems and their impact on the sport is the source of some controversy, we cannot adequately identify and address these issues in this study. It would require a dedicated project to fully explore this relationship.

1 https://www.paralympic.org/sites/default/files/document/150915170806821_2015_09_15%2BExplanatory%2Bguide%2BClassification_summer%2BFINAL%2B_5.pdf

***Corresponding author: Christoph Bartneck,** University of Canterbury, Department of Computer Science and Software Engineering
Elena Moltchanova, University of Canterbury, Department of Maths and Statistics

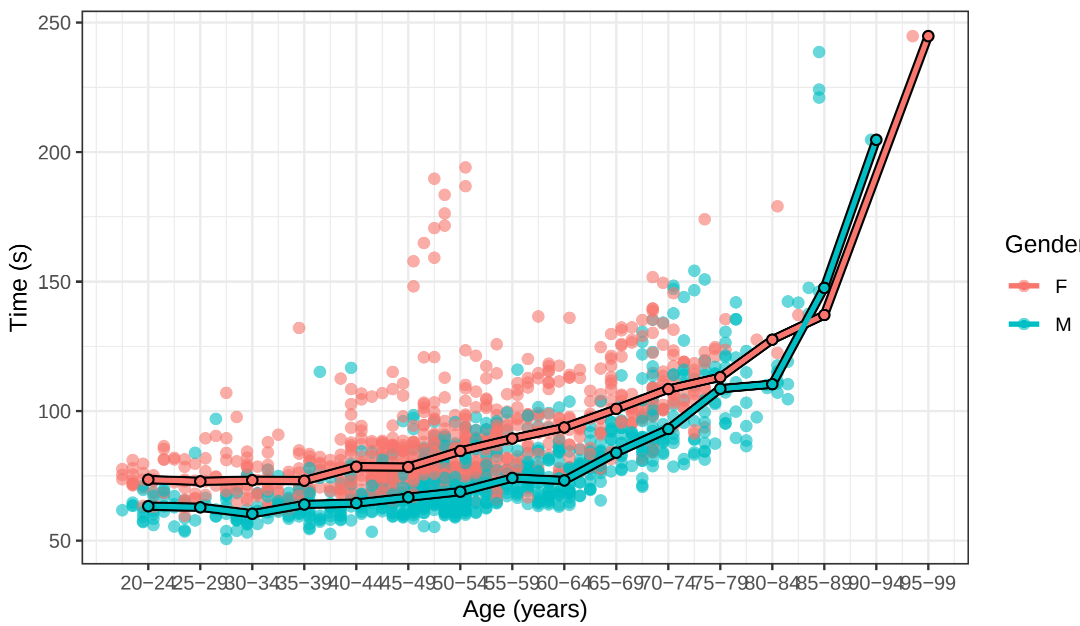

Fig. 1: Influence of age on 100 m Freestyle performance in New Zealand Masters swimmers by Gender. The points show recorded results while the solid lines show the age-group-specific medians.

While the definitions of the disability classes play an important role, they are not the only problematic aspect of the system. Hogarth et al. [16] showed that age is an important factor that predicts the performance of young para swimmers. It has been shown that the differences between age groups is smallest for younger master swimmers (35-49 year old) and highest within older swimmers (> 70) [12, 19, 29, 31, 33]. As a consequence, the difference in performance between two adjacent groups, such as 85-89 and 90-94 is so severe, that the 89-year-old swimmers have to wait until they are 90 to be able to have a chance to win competitions again (see Figure 1). Even the decline in performance within one year can be so severe that a swimmer with a birthday in January has a significant disadvantage to swimmers with a birthday in December. Long distance swimmers, such as 10km open water swimmers, seem to peak later in life than the shorter distance swimmers [38].

Many studies that investigate the influence of age on swim performance are based on data obtained from high profile international events, such as national competitions [13, 20], the Federation Internationale de Notation (FINA, renamed to World Aquatic in 2023) World Cup or the (Para) Olympics [22]. The observed population is therefore focused on young, high performance athletes. The popularity of these data sets is probably based on their open online availability.

1.1 Swimming Point Systems

The goal of any sports handicap system, such as the World Handicap System for golf or the World Para Swimming Point System, is to enable athletes to have a fair competition [15, 36]. A swimmer with no arms should be able to compete against a swimmer with one or two arms. While the game balancing system should compensate for such a specific factor, it should not confound other factors, such as the skill or fitness of the athlete. If two athletes with the same disability compete, the more skilled and fitter should win. It should also not change the ranking order within the different para classifications, sex or age. The ranking order within, for example, the 50-year-old female S10 swimmers should remain intact. The fastest swimmer in this specific group should always get the highest points within this group, although the absolute difference in points between the fastest and second fastest swimmer can be altered.

The International Paralympic Committee (IPC) developed a point system for World Para Athletics [27]. A few years later, the World Para Swimming Federation, which is associated to the IPC, developed their Para Swimming Point system based on the concepts in the World Para Athletics System, such as the use of the Gompertz function (see equation (2)). The World Para Swimming Federation provides a Micosoft Excel based calculator for swim official to convert times to points and the reverse [2].

Mathematical methods have already been developed to compensate age effects for able-bodied youth swimmers [1, 7]. The current World Para Swimming Point System, however, does not take age into consideration. 80-year-old para swimmers are expected to compete with 20-year-old swimmers. The results of such competitions are predictable. This formula is primarily used for the open competition on the international and Olympic level. It is applied on high performance swimmers. It is therefore not directly suitable for para masters swimmers.

The Amateur Swimming Association (Swim England) offers a different approach. Their British Para-Swimming Points Calculator[3] uses the FINA able-bodied formula:

$$P = 1000 \times \left(\frac{B}{T}\right)^3$$

where B is the baseline time and T is the time swum by the athlete. The score is then rounded to the nearest integer. The original FINA points use the current world record as the baseline time split across sex and length of the pool (short course 25m vs. long course 50m)[4]. The British Para-Swimming Points Calculator uses the world para records as their baseline times. This formula does not take the age of the athlete into consideration either nor does it allow athletes to compete across classifications. S9 swimmers can only compete against other S9 swimmers. Graham Spratt from Sportsystems, the company producing the British Para-Swimming Points, pointed out in a support forum post that they were dissatisfied with the World Para Swimming Point System giving too many slow swimmers zero points. This prevents a ranking between those swimmers. In addition, being given zero points does carry a certain psychological value which could discourage slow swimmers.

Able-bodied master swimmers compete within age blocks of five years (20-24, 25-29, etc.). A similar method could be applied to master para swimmers. They could use the para point system to compete across their disability levels within their age & sex group. Fortunately, the number of para athletes is low compared to able-bodied athletes. Unfortunately, it is predictable that even events that are held at a national levels would not offer much competition. The swimmers would be sparsely scattered across the age and sex groups. It is likely that all athletes would almost automatically win medals. The situation would become even worse if swimmers were further confined to compete within their class. This would likely guarantee a gold medal for everyone.

It would also be desirable for para swimmers to compete with able-bodies swimmers in a fair way. In particular for swimmers with only minor disabilities, such as those in S10, who are often already competing with able-bodied swimmers. They are still able to compensate their physiological disadvantage with an increased amount of training. They are therefore able to at least not be the slowest swimmers in their age group. Daly et al. [10] showed that the Paralympic swimmers use similar race speed patterns as able-bodied competitive swimmers. Still, their systematic disadvantage makes it hard for them to win gold.

1.2 Game Balancing

There are alternative methods to enable competitors with different skills and abilities to have an enjoyable competition. In a running competition, the slower athlete could be given a couple of meters of a head start.

2 https://www.paralympic.org/file/2021-world-para-swimming-senior-points-calculator-xlsx
3 https://www.swimmingresults.org/downloads/para-points/
4 https://www.fina.org/swimming/points

In table tennis, for example, the higher skilled player could be required to play with his/her weak hand or would be required to use a smaller paddle. The size of the permissible table area can also be dynamically adjusted to account for the state of the game [2]. Such computer adjusted games are often referred to as exertion games [24].

In computer games it is common practise to dynamically adapt the performance of non-player characters (NPC) to the skills of the human player. In a car racing game, for example, the NPCs can be dynamically adjusted by the game to always give the human player an interesting challenge. If the player falls behind, the NPCs drive slower and allow the human player to catch up. If the human player has taken the lead, the NPCs driver faster and better. This results in the commonly know "Rubber Banding" effect. While such a system is a guarantee for an enjoyable game, it is not generally being used for official competitions where humans compete against other humans.

Still, the World Handicap System for golf dynamically calculates each players' index by taking the most recent performances into account. The advantage of such a dynamic adjustment is that it also compensates for the current skill and training level of players. The strength of such a system is also its greatest weakness, since a novice player could beat a world champion.

1.3 Research Goal

The goal of this project is to analyse the current World Para Point System for swimming and to develop improvements that make the system easier to use and potentially fairer. In addition, we intend to expand the system to include age related effects into a World Para Master Points System. The mathematical model should have the following characteristics:

1. enable para swimmers to compete across impairment classes, sex, and age
2. be useful for high performance swimmers as well as casual swimmers
3. should be easy to use for swim event organisers
4. should be fair

Definitions of fairness are notoriously difficult and depend on their context [21]. For the purpose of this study, we consider a para-point system to be fair if:

1. It takes the age of the athlete into consideration
2. It is transparent
3. The results can be reproduced and verified
4. It does not require any ad-hoc adjustments
5. It takes the whole data set into consideration
6. It can resolve ties effectively

To formalise the definition of fairness a bit further, let X refer to a competitor's result recorded as speed (inverse time) in a certain swimming event, and consider two athletes with attributes (such as disability and age) A_1 and A_2 respectively. Assume they complete a certain competition with speed (inverse time) x_1 and x_2 respectively. Let

$$p_i = \Pr(X \leq x_i | A_i) \qquad \text{for } i = 1, 2 \tag{1}$$

refer to the theoretical proportion of the population with attributes A_i who would have had a worse (slower) result than x_i for $i = 1, 2$. When comparing the two athletes between themselves, the athlete who is better relative to the group with their attributes, i.e., the one with the greater p_i would get a better score. Thus, for example, a blind 90-year-old swimmer who is the best among blind 90-year-old swimmers, will get a better score than an able bodied 20-year-old, who is the worst among able bodied 20-year-olds, even though the latter's time result might be much better in absolute terms.

This study does not have the goal to question the number of classifications nor their criteria. We also do not investigate the consistency of athletes within each classification.

2 Current World Para Swimming Point System

Before being able to propose an improved World Para Swimming Point System it is important to fully understand the current system. We will shortly describe the system, provide explanatory commentary and assess its strength and weaknesses. Consider a competitor assigned to a disability class g, $g = 1, ..., G$, competing in an event k, (e.g., men's 50 m Freestyle,) $k = 1, ..., K$ with the time result t. There is a total of $G = 38$ events (19 each for men and for women), and $G = 14$ disability classes. The current World Para Swimming Point System formula is:

$$P(t, k, g|a, \mathbf{b}, \mathbf{c}) = ae^{-e^{b_k - c_{kg}/t}}, \tag{2}$$

where a is the maximum of available points (set to $1,200$), $\mathbf{b} \in \mathbb{R}^K$ is a vector of event-specific positive parameters, and $\mathbf{c} \in \mathbb{R}^G$ is a matrix of event- and disability class-specific positive parameters [18]. The result is rounded down to the nearest integer. Because, only c is class-specific, the median performance for the more impaired classes is worse than for those less impaired. Also, $\mathbf{b}$ and $\mathbf{c}$ are chosen so that for the world's best t_B, $P(t_B, k, g|a, \mathbf{b}, \mathbf{c})$ is as close to $1,000$ as possible.

To understand the current World Para Swimming Point System, let's start with a simple example. Consider two male para-swimmers from two different classes $S03$ and $S09$, where $S03$ corresponds to more disability than $S09$, competing in 50 m Freestyle. Assume the results are exactly $30s$ for the $S09$ swimmer and exactly $45s$ for the $S03$ swimmer. Finally, assume the parameters for the S03 are $a = 1200$, $b = 6.09$, $c_{S03} = 334.86$ and $c_{S09} = 198.02$. Without any point system, the faster ($S09$) swimmer wins. However, the points obtained from the equation (2) are only 659 for the faster $S09$ swimmer and 926 for the slower $S03$ swimmer, making the latter the winner.

To see the possible rationale behind this, we need to have a closer look at the mathematical properties of the formula (see equation (2)). Let X denote a random variable, which describes the inverse time result of an individual. Assume that this random variable X comes from a distribution with a cumulative density function (c.d.f.) $F_A(x)$, where A denotes the attributes of the individual such as sex, age, or disability class. By definition, $F_A(x) = Pr(X \leq x)$. In other words, for a particular swimming result $x = 1/\text{time}$, the value of the c.d.f. can be interpreted as the proportion of population which would have a worse result than the observed one. We show in Section 2.2 that the formula in the equation (2) is exactly the c.d.f of a Gumbel distribution multiplied by a constant a. Assuming that the swimming results expressed via inverse time have a Gumbel distribution, one can thus say that the $S09$ swimmer with the $30s$ result is better than 55% of their group, and the $S03$ swimmer with the $45s$ result is better than 77% of their group, justifying the win of the latter.

Such score based on group-specific cumulative density functions thus arguably allow for a fair comparison between different groups, since it in effect makes the top p% in each group compete with each other. It should be noted that this result applies to any c.d.f., and furthermore, functions from completely different families can be used for different groups although that is unlikely to be necessary.

2.1 Strength and weaknesses

The current system has several weaknesses. According to the IPC's Explanatory Report, the lower one-third of the dataset has been excluded from the estimation of the parameters a, b and c [26]. The technical report only explains that this exclusion improved the model fit. Still, excluding such a high proportion of data points is unusual for dealing with outliers. It also biases the system towards high performing athletes.

Second, the current system relies on several ad-hoc adjustments in the stepwise parameter estimation process that have to be implemented manually. The Explanatory Report does not document all the adjustments. Since there is no openly available documentation of these adjustments the system does become opaque. It is not possible to replicate the results and hence verify the system.

We also have to point out that the goal of giving 1,000 points to the current world record for each combination of event and class does not work as intended. We entered all the current world records into the current para point calculator and received more than 1,000 points for 67.2% of the long course events and 41.4% for the short course events. Despite all the manual adjustments they made, the current para model still cannot guarantee the symbolic meaning of 1,000 points.

The current World Para Point system therefore does not meet all the desired mathematical characteristics. Specifically, characteristic two is neglected. The current system does not meet any of the six fairness criteria defined above.

In the following sections, we explain further the technical aspects of the underlying rationale for the equation (2), and develop a transparent, systematic, replicable way of fitting the model parameters.

2.2 Gumbel and the Current Para System.

Since sports results usually represent the best of the best, it is common to model them using generalised extreme value distributions. One of them is the Gumbel distribution (also known as double exponential) with the c.d.f. defined as:

$$F(x|b,c) = e^{-e^{b-cx}} \tag{3}$$

for some positive b and c. If one assumes that this distribution has been truncated from above at a point x^*, the c.d.f. becomes:

$$F(x|a,b,c) = \begin{cases} \frac{a}{1000} e^{-e^{b-cx}}, & \text{for } x < x^*, \\ 1, & \text{for } x \geq x^*, \end{cases} \tag{4}$$

where

$$x^* = \frac{b - \log(-\log \frac{1000}{a})}{c}.$$

One can see that this is equation (2) with $x = 1/t$. Thus, in what follows x refers to the inverse of the recorded time result. Figure 2 shows how the scoring system works.

Note, that the corresponding probability density function is:

$$f(x|a,b,c) = \frac{dF(x|,a,b,c)}{dx} = \begin{cases} \frac{a}{1000} c e^{b-cx} e^{-e^{b-cx}}, & \text{for } x < x^*, \\ 0, & x \geq x^*. \end{cases} \tag{5}$$

2.3 A note on ties.

It can be shown that $F(X)$, which is itself a random variable, has a uniform distribution in the range from 0 to 1 [6]. Assuming the chosen distribution is a good fit, this implies a uniform distribution of the resulting scores. (If the scores are rounded to the nearest integer, as is the custom, it implies a *discrete* uniform distribution). In practical terms, for any scoring system transforming a result $X = x$ into a score $S(x)$, it is desirable to avoid ties where $S(x_1) = S(x_2)$, unless, of course, $x_1 = x_2$. However, the current scoring systems are meant to bin the uniformly distributed $F(X)$ into $b = 1,000$ bins. By definition, ties must occur if there are 1,001 or more results.

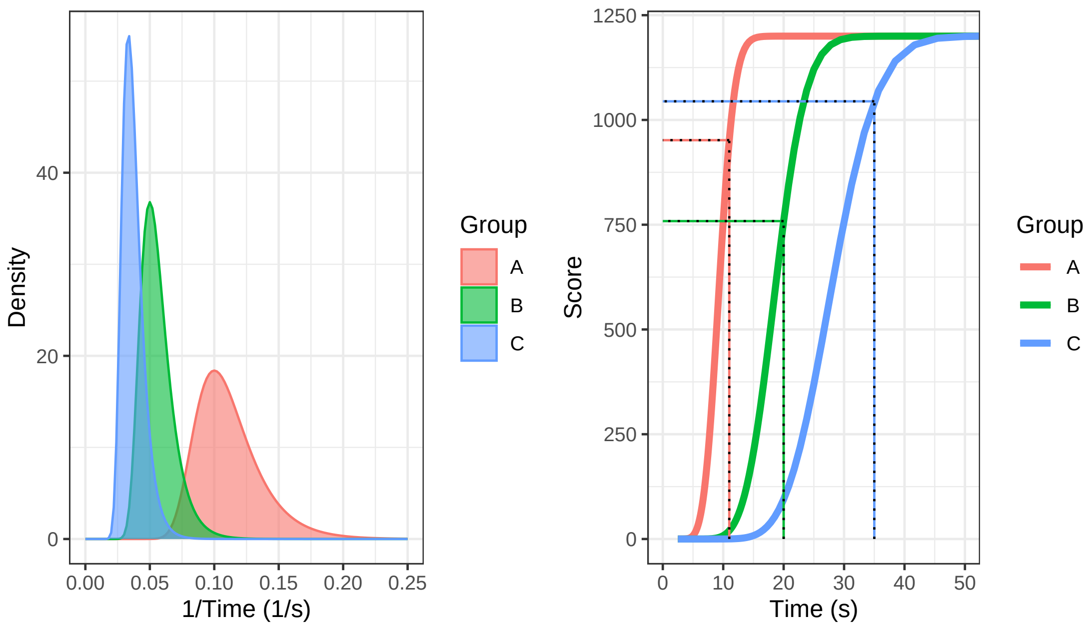

Fig. 2: Transforming competition times into scores comparable across classes of competitors. For each of the classes, A, B, and C, the distribution function is estimated (left panel) and transformed into a score on the scale from 0 to $a = 1,200$ (right panel). Notice, that although the competitor in class C took more time than the competitors A and B, their scores will be higher.

Basic probability theory shows that the probability of having at least one tie when placing n participants into b equally probable bins is:

$$\Pr(\text{at least one tie}) = \begin{cases} 1 - \frac{b!}{(b-n)!b^n}, & \text{for } n < b, \\ 1, & \text{for } n \geq b. \end{cases}$$

Thus, for example, if $n = 50$ participants are placed into $b = 1,001$ bins, the probability of at least one tie is already quite high at 0.712 [5].

2.4 Fitting a Single Gumbel Distribution.

In classical statistics, the traditional way to estimate distribution parameters given data is maximum likelihood estimation. While sometimes analytically tractable, it may often require the use of approximate numerical optimisation algorithms. These, in turn, become harder to manage when one deals with simultaneous estimation of multiple parameters in the presence of various restrictions on the above parameters, because of the issues with non-differentiability and convergence as well as sensitivity to initial values. An alternative approach is quantile matching, where one can use least squares to fit a parametric c.d.f. to the empirical one.

Transforming equation (4) gives:

$$\log\left(-\log\left(\frac{1000}{a}F(x|a,b,c)\right)\right) = b - cx \quad \text{for } x < x^* \text{ and } a \geq 1000. \tag{6}$$

5 To be exact, because the continuous score lies between 0 and $1,000$, rounding to the nearest integer will result in probabilities of $\frac{1}{1000}$ for scores 1 to 999 and to lower probabilities of $\frac{1}{1000\times 2}$ at the extreme scores of 0 and $1,000$. The probability of at least one tie for $n = 50$ participants in this set-up is slightly higher at 0.713.

The cloglog (complimentary log-log) transformation of $\frac{1000}{a}F(x|a,b,c)$ is thus a linear function of x with intercept b and slope c.

Based on that, for a sample of results $x_1,\dots,x_n$ for a specific class and result, given a the estimation of b and c can proceed as follows:

i evaluate $r_i = \frac{\text{order}(x_i)}{n+1}$,
ii scale them to $r'_i = r_i \times 1000/a$, and evaluate the cloglog transformation as $z_i = \log(-\log(r'_i))$.
iii find the estimators for b and c by minimising the sum of squares $\sum_i (z_i - (b - cx_i))^2$ s.t. $c > 0$ and $\max_i(x_i) \le x^*$.

The problem of minimising the sum of squares subject to linear constraints is solved by quadratic programming (see, for example, [25]), which we have implemented in R via the quadprog package [34]. See Appendix B for details and code.

Note, that the solution is conditional on a. We therefore implemented the following grid-search algorithm:

1. define a grid of values $a^{(1)},...,a^{(J)}$.
2. conditional on each value $a^{(j)}$, $j = 1,...,J$, use the quadratic programming algorithm to estimate the parameters b and c and evaluate the respective log-likelihood $L = \sum_i \log(f(x_i|a^{(j)},b,c)$.
3. choose $a = a^{(j)}$ (and the respective estimates for b and c) which maximize the log-likelihood L.

2.5 Addressing Ordering by Disability

Consider now our original scoring problem. In the context of para swimming, we have $G = 14$ disability classes, of which the sport classes S1-S10 are ordered by the severity from highest to lowest. The S11-S13 classes for visually impaired swimmers are also ordered in the same way. The S14 class for intellectually impaired swimmers stands by itself. Consider an event k with two disability classes $g_{(1)}$ and $g_{(2)}$ represented by the parameter sets $\{a_k, b_{k,g_{(1)}}, c_{k,g_{(1)}}\}$ and $\{a_k, b_{k,g_{(2)}}, c_{k,g_{(2)}}\}$ respectively, where $g^{(1)}$ is the more disabled class. Note, that a is assumed common to all classes, while the b and c parameters are, in principle, allowed to be class-specific. In order for our scoring to be consistent, for the same result x we need the score for class $g^{(1)}$ always to be higher than the score for class $g^{(2)}$:

$$S(x|a_k, b_{k,g_{(1)}}, c_{k,g_{(1)}}) \ge S(x|a_k, b_{k,g_{(2)}}, c_{k,g_{(2)}}) \quad \forall x. \tag{7}$$

Ensuring the above ordering within each event is equivalent to making sure that the graphs of the class-specific cumulative density functions do not intersect. As shown in equation (6), a Gumbel c.d.f. can be linearised with b as an intercept and c as a slope. We thus need to find intercepts b and slopes c such that the straight lines defined by them do not intersect, see Figure 3. There are two ways to ensure it within the event: (i) keep the parameters c constant and order class-specific bs or (ii) keeps the parameter b constant and order class-specific cs. Here we choose to follow the current Paralympic system and use the second option. Mathematically, we have the following model for the result x in the event k with the associated disability class g:

$$F(x|\mathbf{a},\mathbf{b},\mathbf{c},g) = \begin{cases} \frac{a_k}{1000} e^{-e^{b_k - c_{k,g}x}}, & \text{for } x < x^*, \\ 1, & \text{for } x \ge x^*, \end{cases} \tag{8}$$

where $a_k \ge 1000$, $c_{k,1} > \cdots > c_{k,10} > 0$, $c_{k,11} > \cdots > c_{k,13} > 0$, and

$$x^* = \frac{b_k - \log(-\log(\frac{1000}{a_k}))}{c_{k,g}}$$

Note, that unlike the current Para Point System, which sets $a = 1200$ for all events, we allow a to be event-specific.

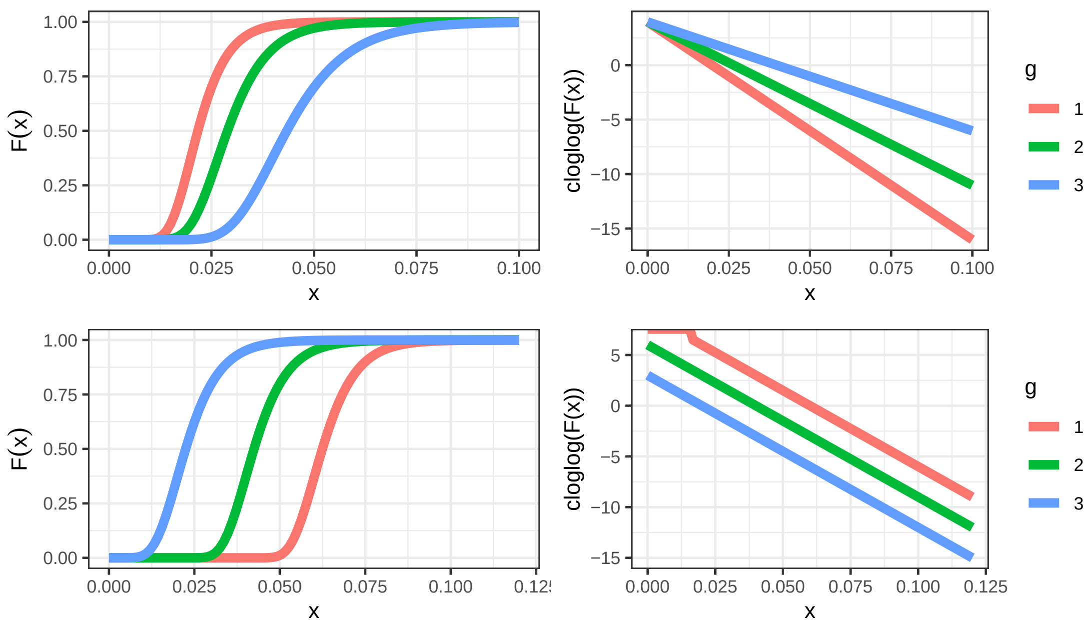

Fig. 3: Using linear transformation of the Gumbel c.d.f. to illustrate the two ways to ensure that among two competitors with the same result x, the one from the more disabled class will always get the higher score: (i) common intercept b with class-specific slopes c (upper panel), and (ii) common slope c with class-specific intercepts b (lower panel).

Given **a**, the parameters **b** and **c** can be estimated, one event at a time, within the quadratic programming set-up as described in Section 2.3 with additional linear constraints on **c**. The parameter **a** can be estimated via grid search as also described in Section 2.3.

For an "empty" class, i.e., a class with no observations available, no estimate can be produced by the method described above. However, the constraint structure can be used to impute an estimate *post hoc*.

3 Fair World Para Point System

Although we now have 16 parameters to estimate for each event (14 disability class-specific b-coefficients, a common c and a common a), and multiple constraints, this still fits easily within the quadratic programming set-up. We still can do a grid search over possible values of a, and fit the model conditional on each. Note, that the empirical percentiles have to be calculated separately within each group.

It should be noted that in a practical application some classes may have no observations in them. It is thus impossible to estimate the slope, and thus the parameter c for such a class. However, such an estimate may be necessary for scoring future participants. Keeping in mind the ordering restrictions on the parameters we suggest local linear interpolation as shown in Figure 4. Thus, for example, if a missing class is between two non-missing classes, then the average of their slopes should be used.

3.1 Addressing Age

A somewhat different problem that hasn't been addressed as far as we know is adjusting by age. Intuitively, on a population level, there is an age at which top fitness and thus best results can be reached. Moving away from that optimal age is associated with a decrease in fitness [4]. Note that for the truncated Gumbel distribution specified in equation (4) the median m is such that

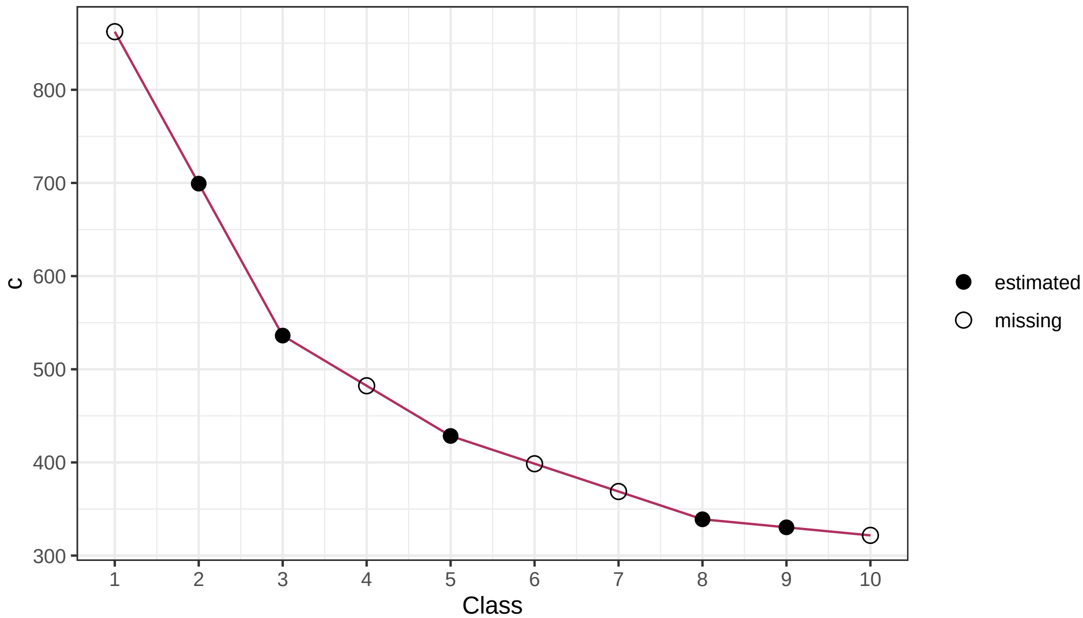

Fig. 4: In case of missing data, the coefficients for the 'empty' groups can be recovered by linear inter- and extrapolation as shown here for an imaginary situation where the coefficients for the classes 2,3,5,8, and 9 could be estimated from the data, but those for the classes 1,4,6,7 and 10 had to be recovered after the analysis.

$$F(m|a,b,c) = \frac{a}{1000} e^{-e^{b-cm}} = 0.5.$$

That is,

$$m = \frac{b - \log\left(-\log\left(.5\frac{1000}{a}\right)\right)}{c}. \tag{9}$$

It is reasonable to assume that this median inverse time has a quadratic dependence on age with the global maximum corresponding to the age of peak fitness at which top speed would be reached (see e.g. [8] and [11]). We thus posit

$$b = \beta_0 + \beta_1 \mathrm{age} + \beta_2 \mathrm{age}^2, \tag{10}$$

where $\beta_2 > 0$ to make sure that the quadratic function has a maximum (rather than a minimum).

Consider now a sample of results $x_1, \ldots, x_n$ for competitors with respective ages $\mathrm{age}_1, \ldots, \mathrm{age}_n$. In order to benefit from the convenience of the quadratic programming approach, we aggregate our data into 1-year age groups, order the observations within each age-group and scale them up to obtain empirical percentiles z_i as before, and minimize the sum of squares

$$\sum_i (z_i - (\beta_0 + \beta_1 \mathrm{age}_i + \beta_2 \mathrm{age}_i^2 - cx_i))^2$$

subject to constraints $\beta_2 > 0$, $c > 0$ and

$$\max_i \left(\frac{a}{1000} \exp(-\exp(\beta_0 + \beta_1 \mathrm{age}_i + \beta_2 \mathrm{age}_i^2 - cx_i)) \right) \leq 1.$$

3.2 Putting Disability and Age Together

In order to keep the ordering between disability classes as well as incorporate adjustment for age, we put the two proposed methods together to produce the final global model and the associated algorithm.

We assume that the inverse time result x in an event k of a person from disability class g of given age has a truncated Gumbel distribution with the c.d.f:

$$F(x|\mathbf{a},\beta,\mathbf{c},k,g,\text{age}) = \begin{cases} \frac{a_k}{1000}\exp(-\exp(\beta_{k,0}+\beta_{k,1}\text{age}+\beta_{k,2}\text{age}^2-c_{k,g}x)), & \text{for } x < x^*_{k,g,\text{age}}, \\ 1, & \text{for } x \geq x^*_{k,g,\text{age}}. \end{cases} \quad (11)$$

such that

$$a_k \geq 1000 \quad \text{for } k=1,\dots,K,$$

$$\beta_{k,2} < 0 \quad \text{for } k=1,\dots,K,$$
$$c_{k,1} > \cdots > c_{k,10} > 0, c_{k,11} > \cdots > c_{k,13} > 0 \quad \text{for } k=1,\dots,K.$$

and

$$x^*_{k,g,\text{age}} = \frac{\beta_{k,0}+\beta_{k,1}\text{age}+\beta_{k,2}\text{age}^2-\log(-\log\frac{1000}{a_k})}{c_{k,g}}. \quad (12)$$

We apply the quadratic programming to quantile matching for results x_i, $i = 1...,n$ within each event k as follows.

1. define a grid of value for a_k, $a^{(1)}$,...,$a^{(J)}$.
2. evaluate empirical percentiles within each disability class and 1-year age-group as $r_i = \frac{order(x_i)}{n+1}$, where n is the number of observations within that disability class of that age group.
3. for each grid value $a^{(j)}$, scale and transform the empirical percentiles to obtain $z_i = \log(-\log(r_i 1000/a^{(j)}))$.
4. use quadratic programming to minimize the sum of squares $\sum_i(z_i-(\beta_{k,0}+\beta_{k,1}\text{age}+\beta_{k,2}\text{age}^2-c_{k,g_i}x_i))^2$ subject to the constraints listed above. Evaluate the corresponding log-likelihood.
5. choose $a_k = a^{(j)}$ corresponding to the highest log-likelihood and use the associated model estimates.

4 Application to synthetic data

4.1 Data Source

In order to demonstrate how the method works in ideal conditions, we have created an artificial dataset with 100 competitors in each age group from 15 to 35 years old, in each of the 14 disability classes. For each competitor, we have sampled inverse times from Gumbel distribution with parameters $\beta_0 = -4.75$, $\beta_1 = 0.75$, $\beta_2 = -0.015$ and $c_g = 750 - 50 \times g$ for $g = 1, ..., 14$. Note that the parameters imply that the class-specific median depends on age as

$$m_g = \frac{5+\log(-\log(.5))-0.015\times(\text{age}-25)^2}{c_g},$$

i.e. reaching the peak fitness at age 25. After generating these observations, we've truncated our sample at $a = 1,200$.

4.2 Results

The estimated value of a was quite close to the true $a = 1200$ at $\hat{a} = 1,251$. The estimated coefficients for the age dependency were $\hat{\beta}_0 = -4.2103$, $\hat{\beta}_1 = 0.6742$, $\hat{\beta}_2 = -0.0135$, again close to the true ones. Figure 5 shows the profile log-likelihood associated with different values of a and the quantile-matching fit without adjusting for age. Figures 6 and 9 show the observed and fitted dependencies between the inverse time and

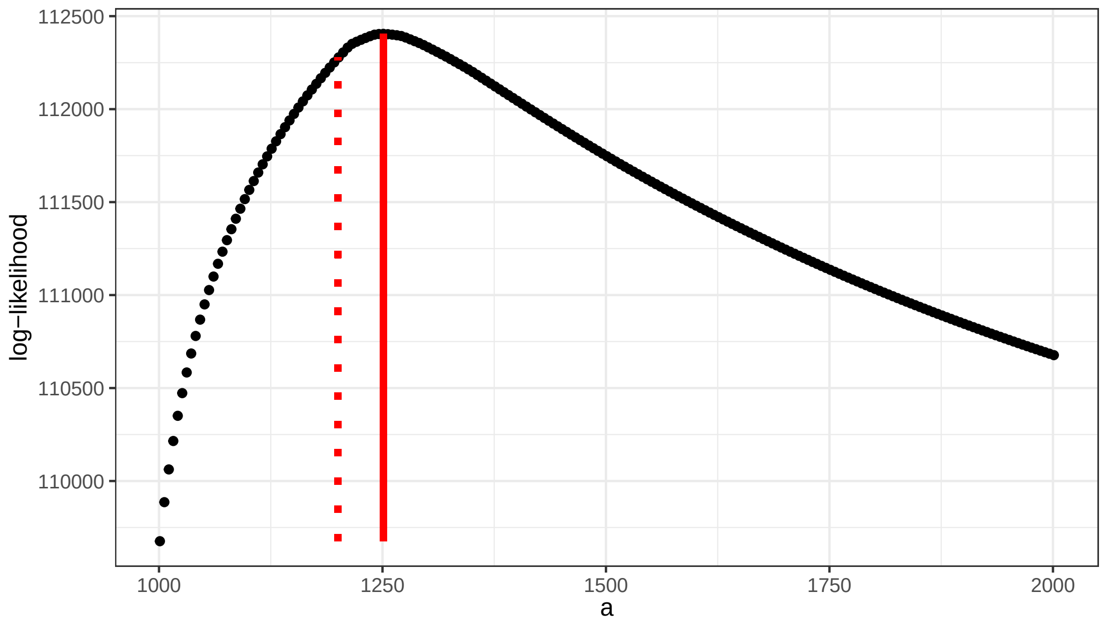

Fig. 5: Log-Likelihood corresponding to each possible value of a for the model applied to the simulated data. The dotted red line indicates the "true" value $a = 1200$ used for simulating the data set. The value corresponding to the highest likelihood ($\hat{a} = 1,251$) is indicated by the solid red line.

the cloglog-transformed empirical percentiles by class and by age-group within class 9 respectively. Figure 7 shows the good fit of median inverse time on age by class, and Figure 8 shows the estimated and observed class-specific coefficients c for our simulated dataset.

Having understood how the methods works for the simulated data, we are now moving on to the real datasets.

5 Application to the Paralympics Data

The International Paralympic Committee (IPC) provided us with a dataset from their swimming events results database. While their results are openly available, including the name of athlete[6], they only provided us with anonymous data that did not allow us to identify any swimmer directly.

The original dataset contained 356,609 individual valid swimming results. For each result the following information was available: year, category (long course vs. short course), event distance, event stroke, time, gender, age and classification. At the data cleaning stage, we have dropped observations with recorded age above 100 as well as observations where the recorded time was either negative or faster than the current world record, assuming these to be clerical errors. Finally, we have removed records of any competitors in 150 m Individual Medley with recorded classes of SM5 and above. The resulting dataset contained 356,528 observations.

The dataset ranges from the year 2006 to 2021 with 155,156 (43.5%) observations for female and 201,372 (56.5%) observations for male participants. The age of the participants ranged from 7 to 68 with a median of 21.75.

The number of participants by event and disability class are shown in Figure 10. While some combinations of events, such as 100 m Freestyle, are very popular across all classifications and gender, there have been a

6 https://www.paralympic.org/swimming/results

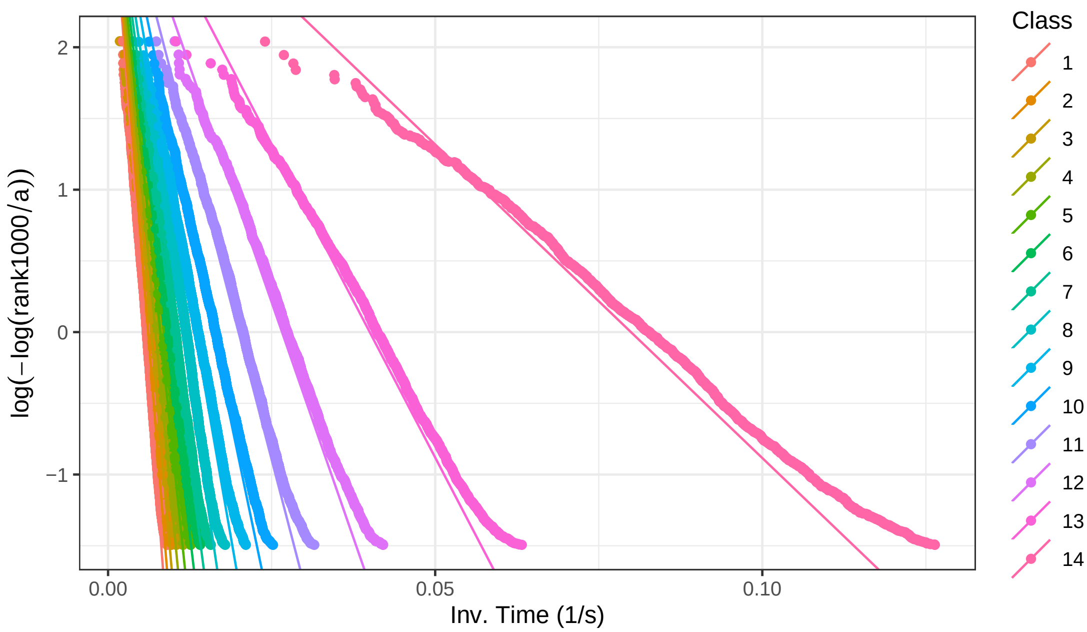

Fig. 6: Class-specific observed (points) and estimated (lines) relationships between inverse time x and the cloglog transformed empirical percentiles for the optimal $\hat{a} = 1,251$ for the simulated data.

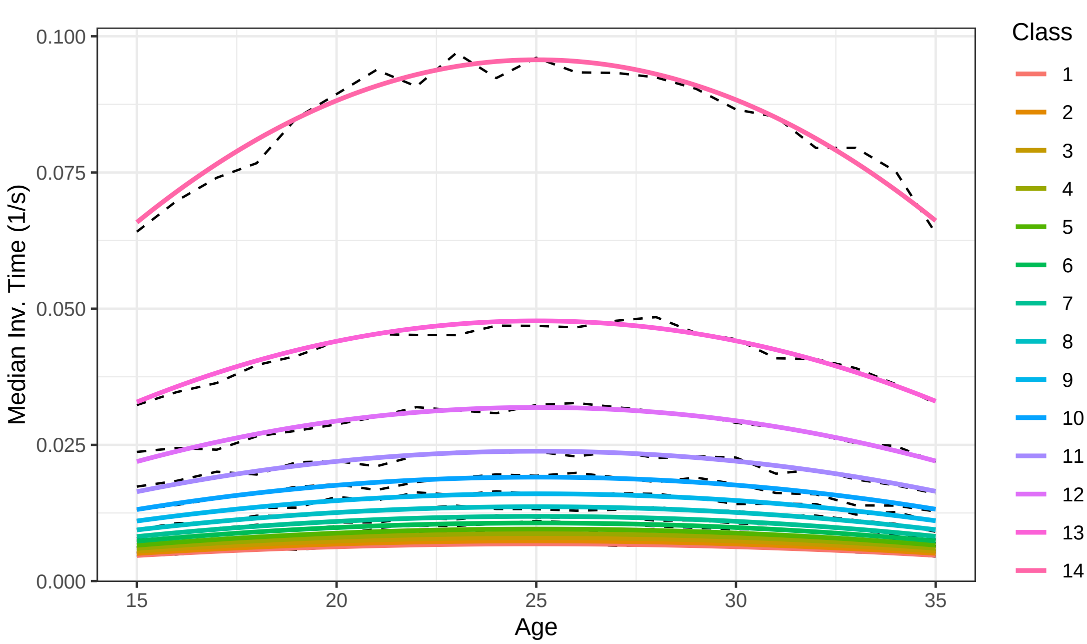

Fig. 7: Left panel: Class-specific observed (black dotted lines) and estimated (colorful lines) class-specific medians as functions of age for the simulated dataset.

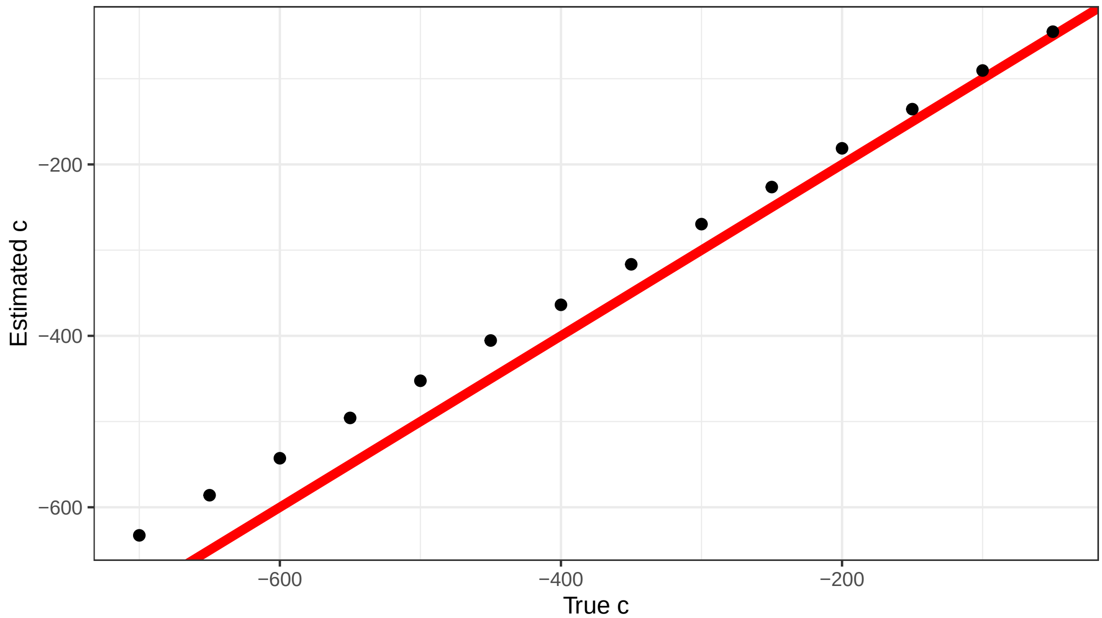

Fig. 8: Observed and estimated class-specific coefficients c for the simulated dataset.

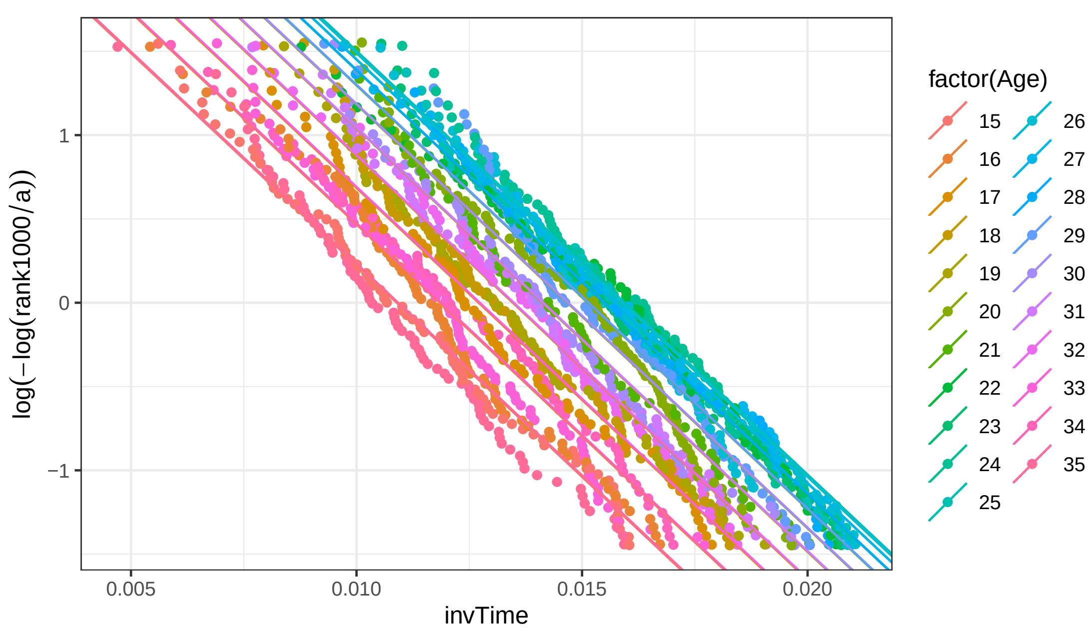

Fig. 9: Age-group-specific observed (points) and estimated (lines) relationships between inverse time x and the cloglog transformed empirical percentiles for the optimal $\hat{a} = 1,251$ for class S09 of the simulated dataset. Note that some lines lie too close together to be distinguishable.

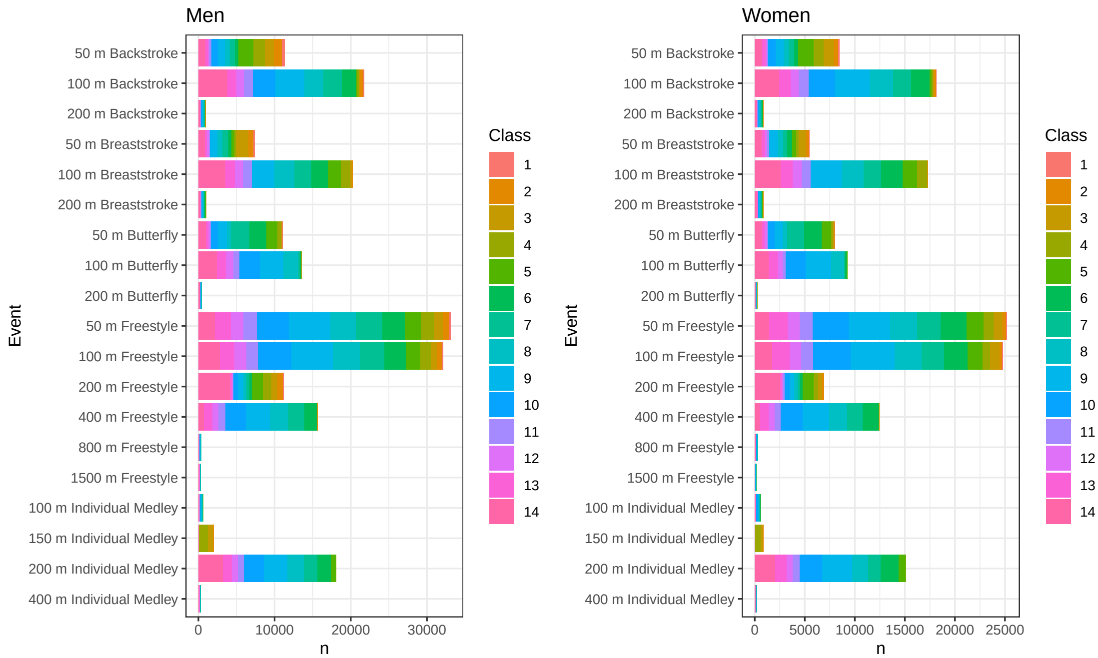

Fig. 10: Participant numbers by event and class in the Paralympics data set.

few combinations where the number of data points is small. However, it does not cause a problem for the model fitting.

It also has to be mentioned that the dataset is based on events that may use qualifying times. This means that all athletes must have reached a high performance level before even being allowed to participate in these competitions. This does have an impact on the distribution of the data since almost no times are being recorded above the qualifying time.

5.1 Results

5.1.1 Fitting

We started by identifying the cut-off point a, common to all 19 events for both genders, which would maximize the joint likelihood for the entire dataset. We found it to be $\hat{a} = 1282$, which is not too far from the currently used one. The log-likelihood associated with that value was $1,609,889$ whereas the log-likelihood associated with the currently used value of $a = 1,200$ was $1,608,657$, i.e. substantially lower.

The summary of the estimated coefficients and performance statistics when adjusting for disability are shown in Tables 3 and 4 respectively (see Appendix A). For all but 3 events, the maximal score was 1000, and the maximum number of ties at the top end was 8 for 100 m Backstroke for women. Looking at the raw scores before rounding to integers, the top three are 1000.389, 1000.389, 1000.387 from the classes S01, S06 and S03 respectively, and their respective cloglog transformations (i.e., the $-(b - cx)$ values) are 1.394197, 1.394195, and 1.394189 (i.e., no ties). We will refer to this score as the Tie Breaking Score, based on its application.

The situation is worse with the ties at the slower end, with the maximum number of ties being 147 out of 30571 male competitors in the 100 m Freestyle event. The five smallest scores are 10^{-25}, 10^{-18}, 10^{-17}, 10^{-13}, and 10^{-10}, i.e. all zero effectively, but their cloglogged version, the Tie Breaking Scores, are 4.164247, 3.837502, 3.790417, 3.596481, and 3.340833 respectively. We can resolve all these results using the Tie Breaking Score leaving us with no ties.

We have also considered the practice of estimating percentiles as $rank/n$ rather than $rank/(n+1)$, where $rank$ is the observed rank, and n is the group size. The system implies that the top performer in the group gets approximately $1,000$ points. When the group is large, there is obviously very little difference, but when n is small the fitted distribution can be quite different depending on the choice made. In the extreme case of one person in a group, for example, it is a difference of assuming that person represents the median result rather than the top result. Statistically, when we use the $rank/n$ assumption, the best estimated a was unchanged at 1282 with the corresponding log-likelihood of $1,696,958$, i.e., higher than that for $rank/(n+1)$. There were understandably substantially more ties at the top (on average 4.26 per event vs. 3.55) and approximately the same average number of ties at the bottom (33.76 vs 34.02).

The parameters and the summary of the fitted score when adjusting for both, disability class and age are shown in Table 5 and Table 6 respectively (see Appendix A). For the age adjusted model, the best common $\hat{a}$ was found to be 1187 with the log-likelihood of $1,595,735$ (i.e., lower than the age-free model) compared to $1,595,690$ for $a = 1,200$. The age adjustment did not change the number of times at the top (still a maximum of 8) and slightly more ties at the bottom (maximum 214 for 100 m Breaststroke, men). Predictably, in general, the age-adjusted scores highly correlated with the age-free scores. ($r = 0.98$). However, there were some substantial differences, the most extreme of which was a difference of 362 points for a 61-year-old participant from S02, one of the oldest recorded. Without the age adjustment, the participant's score was evaluated at 258 points, while after the adjustment the score became 620 points.

5.1.2 Para Point Calculator

The estimation for the various parameters of the statistical model are available in Appendix Appendix A for both, an age inclusive and exclusive model. The same data is also available as CSV files at the Open Science Foundation[7]. These parameters should be updated on an annual basis as more data points become available. For an additional convenience for event organisers, we created an Excel Spreadsheet that calculates the points based on the classification, time and event[8]. A second calculator is available that extends the model with an age component on the same web page. This will be particularly useful for masters competitions. Our Fair World Para Point Calculator for swimming also include the Tie Breaker Score which enables event organisers to resolve potentially all observed ties.

6 Application to New Zealand Masters Data

We used the openly available data from the New Zealand Masters Swimming[9], which contained data from all Masters competitions in the years 2013-2020. The total number of observations in the World Masters data was 13,247 with 6,490 (49.0%) men and 6,757 (51%) women. The age ranged from 20 to 98 years of age with a median of 53. The numbers of participants by event is shown in Figure 11.

6.1 Results

Figures 12 and 13 shows the results of fitting our age-adjusted method to the results of the Freestyle 100m event for men. The event had 806 results recorded with the age of the participants ranging from 20 to 92 with a median of 54. For the fit, we have attempted grouping the data into 1-year, 5-year and 10-year age groups with the resulting estimates for a of 1284, 1294, and 1277 respectively. The corresponding

7 https://doi.org/10.17605/OSF.IO/SJ958
8 https://doi.org/10.17605/OSF.IO/SJ958
9 https://www.nzmastersswimming.org.nz/competition/results/

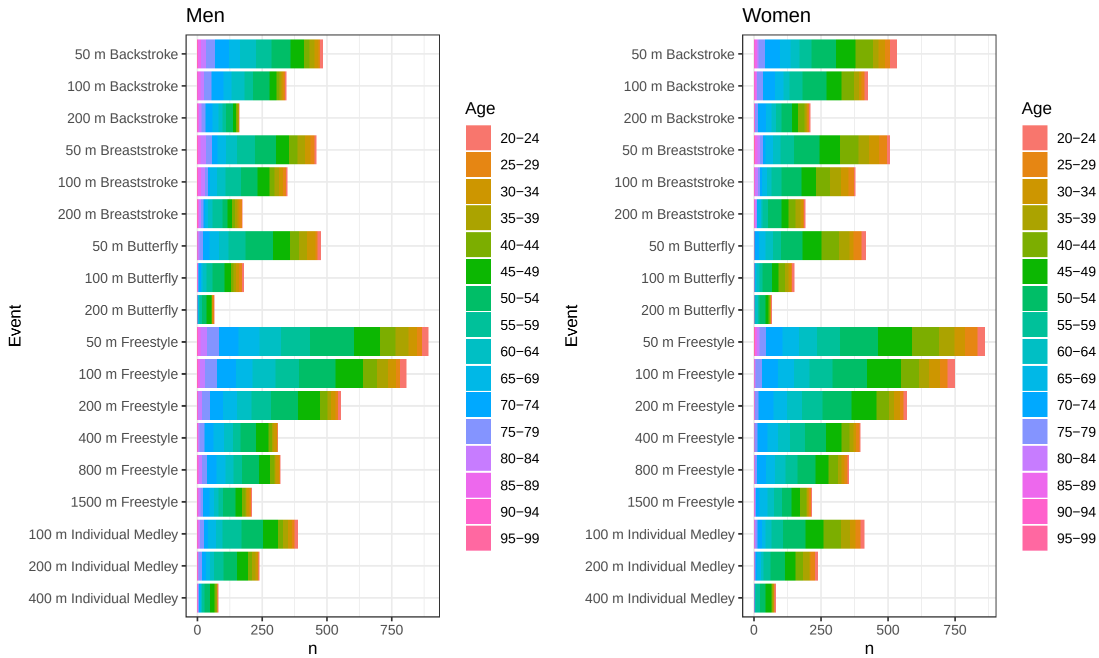

Fig. 11: Participant numbers by event and age group in the New Zealand Masters data set.

log-likelihood values were 3966.777, 3993.364, and 3988.353, indicating the best fit (higher log-likelihood) for the 5-year group aggregation. Figure 12 shows the original data as well as the three fitted median curves. Figure 13 shows the distribution (density) of the scores evaluated based on the 5-year group aggregation fit as well as the cumulative density. Notice that this 5-year group aggregation is not equivalent to the five-year age brackets currently in use with Masters swimming. Our system takes the age of each swimmer into account as an integer.

Recall that our modelling assumptions imply that the distribution of scores should be uniform, i.e., the observed (empirical) cumulative density function (blue line) should coincide with the theoretical c.d.f. of the uniform distribution (the red line) in Figure 13. This doesn't quite happen in this case. The highest score attributed in this analysis was 1000 with the next one being 995, i.e. no ties at the top. The lowest score was 0 assigned to three participants simultaneously.

The common best $\hat{a}$ for all the New Zealand Masters events was 1182. The coefficients for the scoring function for all the New Zealand Masters events as well as the highest and lowest scores and the number of ties are shown in Table 1 for the common a and in Table 2 for the event-specific a (see Appendix A). We found only one tie at the top for the 100 m Breaststroke event for Women (the more precise values were 993.9495 and 993.6281, which were both rounded to 994).The ties at the bottom end were more frequent, with the highest number (and proportion) reached for 50 m Freestyle for Women of 13 ties (1.5%). The Tie Breaking Score was able to resolve these ties.

7 Conclusions and Discussion

We propose a Fair World Para Point System for swimming, based on a Gumbel distribution of inverse time results which simultaneously adjusts for disability and age. The implementation of a quadratic programming algorithm requires specialist software, such as the freely available R software[10], but is

10 http://www.r-project.org

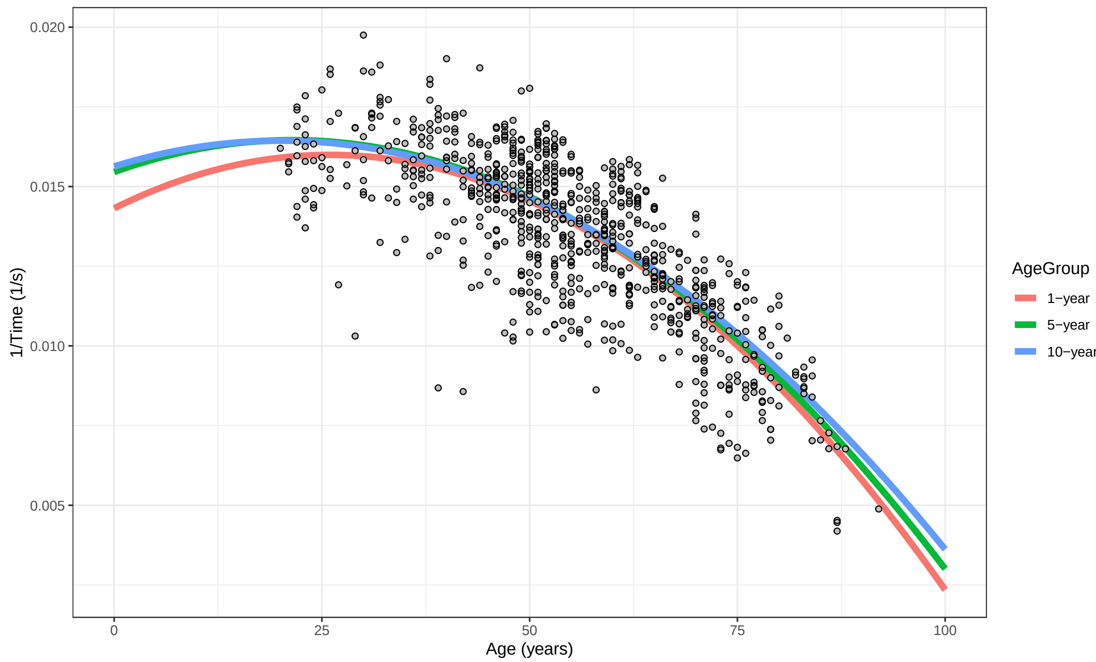

Fig. 12: Observed and estimated relationship between median inverse time and the age of the New Zealand Masters participants.

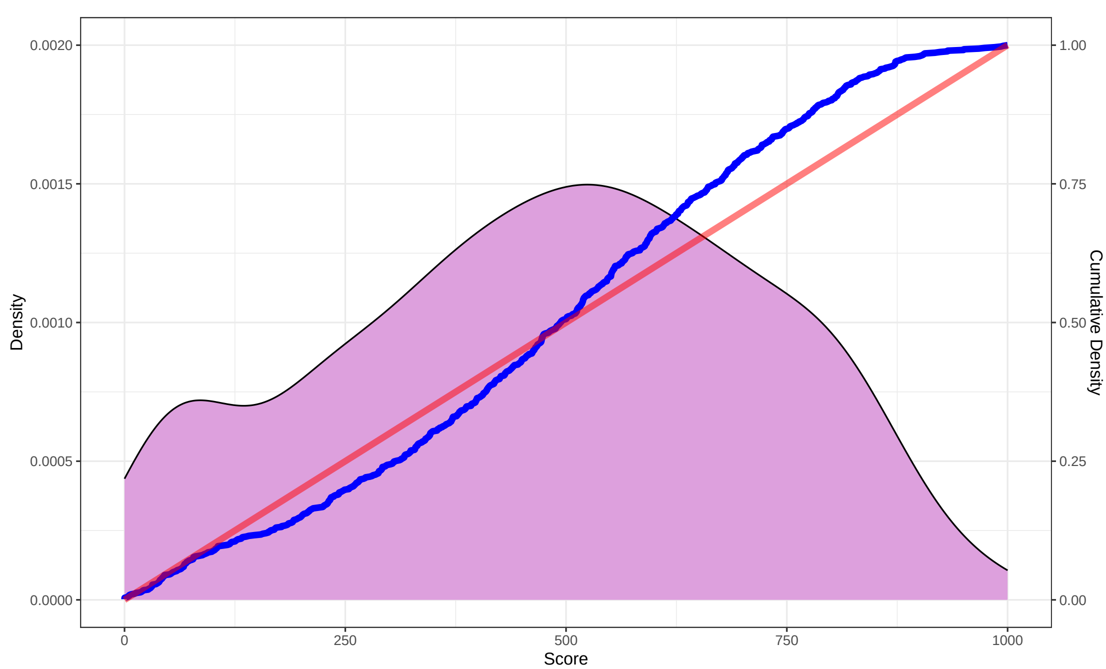

Fig. 13: The fitted distribution of scores. The purple area is the resulting density, and the blue line is the empirical cumulative density function, and the red line is the cumulative density function for a uniform distribution. If the modelling assumptions are correct, the scores should be uniformly distributed and the blue line should coincide with the red one.

otherwise straightforward in a sense of avoiding problems of convergence or sensitivity to initial values and thus not requiring expert supervision. It also allows for the simultaneous estimation of all the parameters involved while taking into account all the required constraints (such as ordering by disability). This avoids awkward step-wise procedures with non-transparent post-hoc adjustments and tweaks that are currently used to ensure the constraints work.

This new system also solves the problem of resolving point ties. No two swimmers will be forced to share a ranking. This is particularly useful for dealing with results at the tail ends of the distribution. The lower tail suffers from more point ties than the upper tail. While the current Para Point System also has the mathematical property of being able to break ties using the Tie Breaking Score, it simply ignored this option so far. Our Fair World Para Point System includes the calculation of the Tie Breaking Score and hence enables this highly useful feature.

The Fair World Para Point System is more transparent and offers a better parameter estimation. It no longer relies on manual adjustments and ad-hoc modifications. This will enable governing bodies to update their system more easily and more frequently without the need for an expert statistician. This holds true not only for the open competitions at international events at the highest performance level, but also for other events.

Including parameters for age will, for the first time, enable disabled masters swimmers to compete. No longer will the performance of a 60-year-old swimmer be measured on the same scale as a 20-year-old swimmer. For both, abled and disabled swimmers, age groups are currently only being considered for young athletes, such as 13 years, 14 years, 15 years, 16 years, and 17-18 years. For older swimmers, the masters competitions offer athletes aged twenty and older to compete. This is currently reserved for abled bodies. There are currently no para masters competitions. Partly because there is no method for athletes to compete fairly against each other. The Fair World Para Point System fills this gap and will enable masters para swimming competitions to take place. We hope that this new system will help overcome this inequality. The free and easy to use point calculators we published will enable para masters swimming events.

7.1 Fit considerations

While the system works very well with the simulated data, the reality is never that simple. Below we discuss several issues which may arise in the process of applying the system to real data and outline possible solutions.

Needless to say, any mathematical model is only as good as the data provided. The coefficient estimates, underlying the resulting scoring system may not work as well for a completely different population of athletes. For example, Olympic competitors are top performers, who have dedicated their lives to excelling in their sport of choice. The system based on their results may not apply well to local school competitions or local masters competitions. To create a more inclusive system, refraining from cutting out a third of the data, as it is currently done by the IPC for some combinations of event/class, does seem necessary.

Any cumulative density function $F(X)$ is able to transform an input $X = x$ into a score on the scale from 0 to 1, which will monotonically increase with x. Within the framework we have proposed,

1. a person with a better time will always have a better score than a person with a worse time from the same group, and
2. of the people with the same time from two different groups, the more disabled group will always get the higher score whatever that same time is

However, using an incorrect and badly fitting distribution means that the interpretation of the score as a percentile - $F(x) \times 100\%$ of the competitors would have a worse score making the fairness of comparison between different classes questionable. Furthermore, it can be shown that $F(X)$, which is itself a random variable, has a uniform distribution in the range from 0 to 1 implying a uniform distribution of the resulting scores [6]. This property is lost when an inappropriate distribution is fitted. In the case of Gumbel distribution, the cloglog-transformed empirical percentiles should depend on x linearly within each class. We have found substantial deviations from that when analysing Para Olympics data. The current Para point system omits the lower third of the results (i.e. slowest times) from their estimation process, which is a practical although not a particularly elegant or transparent decision.

7.2 Resolving Ties

Ties are generally perceived as a problem since they fail to establish order between athletes. Hence, we made a point to monitor for them, especially at the top and the bottom. We haven't found ties to be a problem at the top: there were only a handful of them in the entire multi-year data set, they typically came from different disability classes, and further distinction could be made between them when looking at more exact numbers, such as the Tie Breaking Score, which is defined as the cloglog score $b - c/x$ before transformation into 0-1,000 integer score system if desired. The ties at the zero-end of the distribution were much more prevalent, and potentially raise two issues: the already mentioned problematic fit and the generally demoralising effect of many competitors being given zero points.

The latter issue is, perhaps, easier to address. If one wants to avoid giving zero points, the system can be easily scaled to start at 50 points. The solution is admittedly whimsically similar to anecdotal evidence of aeroplanes not having unlucky row numbers and elevators not having unlucky floor numbers. The issue with a large number of participants receiving the lowest score, whatever it may be, should be addressed from the probabilistic point of view. Some ties are to be expected. If m people are randomly allocated points from 0 to $1,000$ with equal probabilities (uniform distribution), the number of people expected to get exactly zero points or exactly 1,000 points is $m/1001$. However, for the 50 m Freestyle event for men, with 33,108 participants, the expected number of zero scores is thus about 33, much lower than the reported 222. On the other hand, the number of participants who get 1,000 points is much lower than 33. This points to the problems with model fit.

7.3 Distribution

Our analysis has shown that the fit was often problematic also in very far from the perfectly straight lines expected under the assumption of the Gumbel distribution. In practice, however, there is no right answer at the end of the book and when fitting a statistical model one should ask whether the best model is good enough. There are few alternatives to Gumbel in the literature. For example, [14] suggest a generalised extreme value distribution (of which Gumbel is a special case), but admit that some aspects of parameter fitting are still an open question. Moreover, the advantage of the Gumbel distribution, which allows linearisation and imposition of disability class- and age-related constraints within the quadratic programming routine is lost. Thus fitting these other distributions would involve non-linear optimisation with constraints in a multi-dimensional parameter space, which makes it non-trivial to execute. Such an optimisation algorithm will be sensitive to initial values and might run into convergence problems. Therefore our proposed framework does not move away completely from the system currently used by the IPC, but rather formalises it in a more elegant and holistic way and extends it to take account of age.

A potential reason for the weakness of the model fit is that each disability class might include a diverse group of disabilities. While each athlete is unique, it is necessary to group them into a limited number of classes. This can force two (or more) distinct populations into a single class, resulting in poor within-class homogeneity.

7.4 Parameter Estimation

The framework assumes truncation for practical as well as statistical reasons. Statistically, there is evidence that the maximum likelihood is substantially higher for $a > 1,000$, i.e. for a truncated distribution, which assumes that the best results are yet to come. In practice, there is a desire to have a buffer for future results. We have chosen to have common a for all events for the sake of consistency, but it is possible to allow for individual values. To illustrate the difference, we have estimated the age-adjusted model with both, fixed and dynamic a in tables 1 and 2 respectively (see Appendix A). Furthermore, since our algorithm is straightforward to implement, it can be re-applied, for example, annually to new data and the parameter values can be updated as necessary. In comparison to the current World Para Point System that uses a fixed value of a of 1,200, we based the fixed value of a on the calculation of an optimal value.

7.5 Fairness

The goal of any handicap system is to enable fair competition between athletes. The current World Para Point system manually aspires to give 1,000 points for a world record within each event without actually achieving it. As a result, many of the world records do not receive 1,000 points. Moreover, the many ad-hoc adjustments made to approximate 1,000 points might serve the top athletes but the overall performance of the World Para System suffers.

The Fair World Para Point system guarantees that the best time within the event will not exceed 1,000 points, but it may be less. Let's take the example of 100 Freestyle Men. Here, the leading S9 athlete receives 897 points, while the leading S2 athlete receives 1,000 points. The S9 athlete might feel disadvantaged since he would lose out despite being the best in class. Being in a certain classification might make it harder to win. A case could be made that athletes in a certain classification might be systematically disadvantaged.

While this situation is not ideal, the current World Para Swimming Point System is facing the exact same constraint. Still, the maths behind the para point systems is agnostic. It assumes that the athletes' performance can be adequately described by the Gumbel distribution. It also implies that the best athlete in the entire population of potential competitors will get the top score. In practice, however, we only observe a sample of the potential competitors, whose performance is, moreover, subject to random variation. The only way to guarantee 1,000 points to the top performer in each class is to explicitly force the model to do so. This could induce serious biases for the rest of the distribution. Unfortunately, in addition to the vagaries of random sampling, the model fit remains, not entirely unexpectedly, far from perfect.

The fundamental question is to what degree the model should be manipulated to suit the top performers. On the one hand, one might argue that winning is everything and that the overall fairness of the rankings is less important. On the other hand, one could bring the point forward that fairness is better served by benefiting the largest number of athletes, no matter if they are top performers or not.

Austin [3] pointed out that the Olympic Creed is "The important thing in life is not the triumph, but the fight; the essential thing is not to have won, but to have fought well." It does not follow the Vince Lombardi principle of "Winning isn't everything. It's the only thing." The winning mentality can have serious negative consequences, not only for the athletes but the sport overall. It follows that a system that is going to be used for the Para Olympics should use the Fair World Para Point system that optimises fairness for all.

This adoption of the Fair World Para Point system will require the IPC to update its software system and processes. While this does require some resources, it is worthwhile for the reasons mentioned above. For regular uses of the IPC they only provide an Excel spreadsheet similar to the one we prepared and hence the is no difference in the workload for most event organisers.

After reviewing the performance of the Fair World Para Point system against the criteria set out in the introduction, we can conclude that it earned the attribute "fair". It does meet all the requirements, which cannot be said for the current World Para Point system. Still, there are some underlying mathematical issues that do leave room for improvement.

7.6 Future Work

Our system allows for an extension that would enable abled and disabled swimmers to compete. Abled swimmers would simply become their own class, increasing the number to a total of 15. This would allow disabled swimmers to compete against a much larger number of competitors. We intend to include this option in the next version of our system.

It would also be beneficial to analyse the homogeneity within each class. There is a chance that some classes, far from representing a single cluster, consist of several homogeneous groups. It would also be interesting to see if the groups identified by an algorithm will correspond to the existing disability classes. One example of such an analysis would involve the implementation of a Reversible jump MCMC algorithm for Gumbel mixtures (see, for example, [30]). The RjMCMC algorithm allows not only the estimate of the parameters of each mixture but also the estimate of the most likely number of groups involved as well as the likeliest associated membership.

Another avenue of research is to reconsider the distribution being fitted and the nature of the sample available. The former may include using the *peaks* over threshold (POT) methodology outlined by [14]. The latter would mean fitting a distribution truncated both, above and below. Moreover, one may consider stochastic truncation where the probability of inclusion into the sample is not simply zero or one depending on whether the performance is below or above a certain threshold but rather depends in a monotonic parametric way on that performance.

The outlined approaches should help better understand and improve the fit, thus promoting a more fair and equitable competition framework.

8 Acknowledgements

We would like to thank Alexander Picolin from the International Paralympic Committee for providing data and for providing feedback about their para point system. We would like to thank Liz Davidson from the New Zealand Masters Swimming for providing data.

9 Appendix A

								max. score			min. score		
Sex	M	Stroke	n	β_0	β_1	β_2	c	points	ties	ties %	points	ties	ties %
F	50	Butterfly	417	5.06	0.0563	-0.00102	219.24	997	0	0.0	0	0	0.0
F	50	Backstroke	532	5.52	-0.0109	-0.00025	209.10	1000	0	0.0	0	6	1.1
F	50	Breaststroke	507	5.35	0.0353	-0.00076	252.31	983	0	0.0	0	0	0.0
F	50	Freestyle	861	6.43	0.0102	-0.00056	210.91	964	0	0.0	0	13	1.5
F	100	Butterfly	151	4.74	0.0332	-0.00072	435.23	929	0	0.0	56	0	0.0
F	100	Backstroke	425	5.09	0.0001	-0.00035	433.95	979	0	0.0	0	2	0.5
F	100	Breaststroke	379	6.43	0.0172	-0.00062	600.70	994	1	0.3	0	0	0.0
F	100	Freestyle	749	5.62	0.0127	-0.00053	420.19	1000	0	0.0	0	11	1.5
F	100	IM	412	6.54	0.0228	-0.00071	561.61	1000	0	0.0	4	0	0.0
F	200	Butterfly	67	3.09	0.0861	-0.00122	914.44	981	0	0.0	106	0	0.0
F	200	Backstroke	210	6.14	-0.0094	-0.00032	1040.19	998	0	0.0	26	0	0.0
F	200	Breaststroke	191	7.04	-0.0242	-0.00021	1215.64	941	0	0.0	10	0	0.0
F	200	Freestyle	570	5.03	0.0285	-0.00065	890.21	1000	0	0.0	12	1	0.2
F	200	IM	239	5.66	0.0298	-0.00067	1120.31	977	0	0.0	21	0	0.0
F	400	Freestyle	396	4.79	0.0389	-0.00075	1873.25	1000	0	0.0	20	0	0.0
F	400	IM	81	4.24	0.0377	-0.00052	2047.06	916	0	0.0	75	0	0.0
F	800	Freestyle	355	4.29	0.0499	-0.00077	3909.13	1000	0	0.0	21	0	0.0
F	1500	Freestyle	215	4.44	0.0358	-0.00063	7216.89	1000	0	0.0	29	0	0.0
M	50	Butterfly	476	5.90	0.0592	-0.00110	209.08	1000	0	0.0	0	1	0.2
M	50	Backstroke	485	4.30	0.0337	-0.00067	167.76	1000	0	0.0	1	0	0.0
M	50	Breaststroke	460	5.31	0.0616	-0.00102	238.89	1000	0	0.0	0	1	0.2
M	50	Freestyle	891	6.64	0.0404	-0.00089	202.67	1000	0	0.0	0	6	0.7
M	100	Butterfly	179	3.63	0.0905	-0.00133	368.49	967	0	0.0	10	0	0.0
M	100	Backstroke	344	5.78	0.0265	-0.00077	429.57	1000	0	0.0	0	1	0.3
M	100	Breaststroke	347	5.41	0.0676	-0.00110	544.26	1000	0	0.0	13	0	0.0
M	100	Freestyle	806	6.35	0.0501	-0.00103	441.44	1000	0	0.0	0	2	0.2
M	100	IM	387	5.94	0.0350	-0.00083	457.30	1000	0	0.0	0	0	0.0
M	200	Butterfly	66	3.96	0.0207	-0.00055	663.87	896	0	0.0	49	0	0.0
M	200	Backstroke	162	5.85	-0.0078	-0.00039	831.66	942	0	0.0	0	0	0.0
M	200	Breaststroke	173	4.79	0.0807	-0.00117	1171.03	996	0	0.0	66	0	0.0
M	200	Freestyle	554	5.94	0.0619	-0.00111	985.76	985	0	0.0	0	0	0.0
M	200	IM	238	4.58	0.0752	-0.00112	1000.87	971	0	0.0	13	0	0.0
M	400	Freestyle	311	6.07	0.0467	-0.00096	1988.56	977	0	0.0	3	0	0.0
M	400	IM	81	3.19	0.0487	-0.00076	1462.98	879	0	0.0	120	0	0.0
M	800	Freestyle	319	5.55	0.0334	-0.00079	3685.36	1000	0	0.0	24	0	0.0
M	1500	Freestyle	211	6.09	0.0394	-0.00087	7823.52	1000	0	0.0	6	0	0.0

Tab. 1: Estimated coefficients and performance statistics for the Fair World Para Point calculation for the New Zealand Masters, age adjustment only, $a = 1294$.

										max. score			min. score	
Sex	M	Stroke	n	a	β_0	β_1	β_2	c	max. score	Ties (top)	Ties (top) %	min. score	Ties (bottom)	ties (bottom) %
F	50	Butterfly	417	1193	5.00204	0.05569	-0.00101	216.3874	991	0	0.0	0	0	0.0
F	50	Backstroke	532	1185	5.48484	-0.01062	-0.00025	207.8302	997	0	0.0	0	6	1.1
F	50	Breaststroke	507	1252	5.01543	0.03313	-0.00071	233.5131	944	0	0.0	0	0	0.0
F	50	Freestyle	861	1577	4.87423	0.00662	-0.00040	148.1262	750	0	0.0	0	9	1.0
F	100	Butterfly	151	1821	3.42190	0.02500	-0.00051	282.8484	649	0	0.0	52	0	0.0
F	100	Backstroke	425	1274	4.67328	0.00077	-0.00032	392.3384	928	0	0.0	1	4	0.9
F	100	Breaststroke	379	1204	6.28448	0.01701	-0.00060	585.2388	982	0	0.0	0	0	0.0
F	100	Freestyle	749	1231	5.10683	0.01464	-0.00051	384.0212	960	0	0.0	0	9	1.2
F	100	IM	412	1104	8.07581	0.01867	-0.00077	686.2911	1071	0	0.0	1	0	0.0
F	200	Butterfly	67	1001	4.18116	0.10748	-0.00156	1245.0316	1093	0	0.0	93	0	0.0
F	200	Backstroke	210	1190	6.08473	-0.00920	-0.00032	1029.9283	993	0	0.0	26	0	0.0
F	200	Breaststroke	191	1606	5.29512	-0.01452	-0.00018	872.2103	736	0	0.0	17	0	0.0
F	200	Freestyle	570	1058	7.88469	0.01800	-0.00075	1314.7246	1117	0	0.0	0	1	0.2
F	200	IM	239	1281	5.16546	0.02789	-0.00061	1008.1346	923	0	0.0	25	0	0.0
F	400	Freestyle	396	1087	6.55758	0.03171	-0.00080	2427.8245	1087	0	0.0	6	0	0.0
F	400	IM	81	1001	5.91226	0.03876	-0.00058	2828.8176	1053	0	0.0	53	0	0.0
F	800	Freestyle	355	1068	6.41108	0.04894	-0.00089	5500.5356	1107	0	0.0	3	0	0.0
F	1500	Freestyle	215	1132	5.09458	0.03479	-0.00066	8133.6793	1044	0	0.0	20	0	0.0
M	50	Butterfly	476	1211	5.77495	0.05334	-0.00103	200.2414	976	1	0.2	0	1	0.2
M	50	Backstroke	485	1062	6.33187	0.04195	-0.00093	242.7973	1113	0	0.0	0	2	0.4
M	50	Breaststroke	460	1192	5.19894	0.06113	-0.00100	234.1280	992	0	0.0	0	1	0.2
M	50	Freestyle	891	1329	5.87413	0.02448	-0.00067	166.5887	889	0	0.0	0	6	0.7
M	100	Butterfly	179	1323	3.25626	0.08044	-0.00117	320.9848	893	0	0.0	13	0	0.0
M	100	Backstroke	344	1131	6.07939	0.03593	-0.00089	472.9108	1045	0	0.0	0	1	0.3
M	100	Breaststroke	347	1099	6.82942	0.07586	-0.00129	677.7270	1075	0	0.0	4	0	0.0
M	100	Freestyle	806	1294	5.78932	0.03526	-0.00082	377.8111	913	0	0.0	0	2	0.2
M	100	IM	387	1120	6.35540	0.04841	-0.00100	515.8670	1055	0	0.0	0	0	0.0
M	200	Butterfly	66	2379	2.63932	0.01657	-0.00035	364.9747	497	0	0.0	40	0	0.0
M	200	Backstroke	162	1582	4.46421	-0.00255	-0.00031	605.1387	747	0	0.0	0	0	0.0
M	200	Breaststroke	173	1001	6.67579	0.10595	-0.00158	1649.4686	1110	0	0.0	38	0	0.0
M	200	Freestyle	554	1244	5.59359	0.05838	-0.00104	920.1093	950	0	0.0	0	0	0.0
M	200	IM	238	1316	4.07989	0.06769	-0.00099	874.5848	898	0	0.0	17	0	0.0
M	400	Freestyle	311	1283	5.51999	0.04345	-0.00087	1787.8470	921	0	0.0	5	0	0.0
M	400	IM	81	1001	4.41781	0.05627	-0.00094	2021.5399	1025	0	0.0	106	0	0.0
M	800	Freestyle	319	1090	7.25714	0.03137	-0.00090	4728.9923	1084	0	0.0	9	0	0.0
M	1500	Freestyle	211	1143	6.68950	0.03964	-0.00092	8552.9758	1034	0	0.0	3	0	0.0

Tab. 2: Estimated coefficients and performance statistics for the Fair World Para Point calculation for the New Zealand Masters, age adjustment only, event-specific a.

Sex	M	Stroke	n	b	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}
F	50	Butterfly	7193	3.51	248.71	248.71	248.71	196.99	189.89	169.27	159.03	145.79	131.58	123.22	139.56	121.97	121.97	127.45
F	100	Butterfly	8136	5.00	518.26	518.26	518.26	518.26	518.26	518.26	502.54	435.77	399.22	396.43	477.30	389.74	378.78	397.08
F	200	Butterfly	224	5.82	2617.13	2335.10	2053.07	1822.66	1592.25	1361.84	1288.83	1063.80	999.13	979.95	1016.98	1016.98	926.57	1061.09
F	50	Backstroke	7545	2.81	279.77	249.70	201.80	184.98	156.01	156.01	139.53	128.73	115.58	108.20	128.24	109.67	109.35	109.84
F	100	Backstroke	15777	4.27	795.14	719.86	594.40	588.56	504.56	450.70	424.64	385.39	363.77	347.31	404.09	357.27	342.80	351.36
F	200	Backstroke	731	4.46	1292.44	1887.37	1292.44	1292.44	1139.67	998.84	975.86	884.06	812.74	753.26	871.68	795.88	775.46	780.29
F	50	Breaststroke	4825	3.55	400.58	316.76	259.56	230.71	210.67	209.19	193.36	167.75	153.87		166.33	145.89	145.89	153.19
F	100	Breaststroke	15316	4.16	1013.63	900.93	718.01	576.79	521.15	499.05	480.86	408.20	392.35		431.08	386.93	377.64	386.68
F	200	Breaststroke	743	5.26	1955.61	2497.84	1955.61	1413.37	1351.87	1351.87	1214.95	1068.58	1040.74		1155.06	993.80	993.80	1033.36
F	50	Freestyle	22440	3.52	251.92	251.92	197.89	183.70	175.85	156.93	143.39	132.60	119.82	113.99	128.55	114.54	110.08	115.07
F	100	Freestyle	21983	3.73	566.29	566.29	461.67	426.97	381.30	359.00	328.68	306.10	275.48	263.44	306.19	267.67	259.48	267.01
F	200	Freestyle	6328	3.54	1179.64	1179.64	984.16	836.29	802.97	730.07	670.00	610.84	573.63	537.29	600.24	540.96	540.85	546.26
F	400	Freestyle	10762	5.43	2837.60	4256.93	2837.60	2515.28	2373.55	2078.77	2040.89	1905.82	1776.55	1719.92	1991.70	1774.05	1686.38	1751.47
F	800	Freestyle	310	6.55	7216.84	6806.00	6395.17	5984.34	5573.51	5162.68	4898.43	4559.79	4188.66	4172.80	4790.35	4329.86	4051.92	4390.96
F	1500	Freestyle	131	4.84	13823.15	12689.11	11555.08	10421.04	9287.00	8152.97	6946.11	6549.42	6179.16	6179.16	6724.41	5993.26	5583.59	6387.3
F	100	IM	513	5.20	1398.07	1075.79	753.52	685.42	617.32	558.29	533.16	476.76	448.18	408.28	457.64	408.51	404.77	415.03
F	150	IM	852	3.11	1016.66	1008.85	783.92	684.89										
F	200	IM	13223	5.35	1615.57	1583.93	1552.3	1483.78	1304.35	1158.43	1135.77	1047.23	966.44	926.61	1056.07	937.49	905.28	926.46
F	400	IM	176	5.86	4762.38	4392.3	4022.22	3652.14	3282.06	2911.98	2799.17	2290.57	2198.78	2101.59	2480.69	2153.34	1996.97	2182.55
M	50	Butterfly	10561	3.72	339.53	301.56	263.60	199.62	156.67	152.93	142.62	131.81	120.46	115.22	122.33	113.57	111.95	116.65
M	100	Butterfly	12922	5.82	1369.21	1123.08	876.95	572.84	561.29	496.14	492.39	427.09	409.17	388.04	436.70	388.21	383.56	391.24
M	200	Butterfly	416	6.69	1545.61	1523.37	1501.14	1478.90	1456.67	1434.43	1308.33	1061.69	1038.06	997.07	1082.34	1008.16	995.10	1008.08
M	50	Backstroke	10720	2.85	254.55	200.25	179.20	166.63	133.39	133.39	126.81	115.34	102.43	97.54	114.50	96.23	93.00	99.84
M	100	Backstroke	20417	4.62	769.47	632.66	558.00	558.00	445.42	425.82	406.21	375.99	345.35	331.00	386.98	333.40	323.99	335.23
M	200	Backstroke	863	5.03	1882.51	1600.24	1564.40	1519.63	1061.22	996.31	963.74	884.90	805.30	744.42	931.18	810.93	728.61	811.55
M	50	Breaststroke	7049	3.79	406.25	262.39	234.61	221.41	199.73	181.37	173.03	152.38	142.39		153.83	138.92	134.26	139.69
M	100	Breaststroke	19327	4.34	1037.33	652.97	561.87	524.68	488.21	451.39	408.95	384.29	358.06		388.90	350.91	342.19	346.87
M	200	Breaststroke	974	5.41	2472.14	1838.77	1497.43	1369.72	1284.29	1161.98	1136.47	964.24	954.41		1034.16	959.94	933.33	938.39
M	50	Freestyle	31583	3.97	341.88	271.41	207.97	185.36	161.62	152.54	138.03	129.27	118.64	111.88	123.93	111.22	109.43	112.53
M	100	Freestyle	30571	4.24	765.08	600.60	517.92	435.29	367.98	356.80	326.32	302.95	278.00	261.00	299.26	264.30	258.96	263.62
M	200	Freestyle	10756	3.92	1438.11	1177.42	1004.69	876.07	763.45	718.39	673.66	624.79	566.71	531.47	634.74	582.38	537.63	531.81
M	400	Freestyle	14661	5.75	5005.80	4021.57	3145.21	2486.98	2262.33	2018.84	1936.12	1840.25	1711.15	1640.73	1880.95	1717.91	1649.78	1681.21
M	800	Freestyle	352	5.66	8987.21	7584.79	6182.37	4779.95	4118.33	4118.33	3859.38	3834.28	3446.11	3291.78	4064.86	3567.59	3354.77	3544.09
M	1500	Freestyle	289	5.76	20812	17124.85	13437.69	13437.69	9152.76	8030.02	7573.95	7468.49	7071.24	6105.26	7325.74	7325.74	6500.49	6839.40
M	100	IM	552	6.25	1375.68	1149.35	923.02	696.69	603.88	570.85	543.78	492.09	451.59	437.16	494.09	418.42	418.42	432.43
M	150	IM	2001	3.91	1227.09	1188.99	849.88	748.97										
M	200	IM	17157	5.64	1999.41	1821.17	1642.94	1352.80	1142.38	1111.87	1047.81	964.48	909.34	868.33	968.66	876.53	854.40	862.3 0
M	400	IM	274	7.44	4100.50	3936.26	3772.02	3607.78	3443.55	3302.05	2778.62	2578.48	2480.47	2407.90	2688.46	2340.68	2298.02	2495.28

Tab. 3: Estimated coefficients for disability adjusted Fair World Para Point Score ($a = 1282$).

Sex	M	Stroke	n	max. score	Ties (top)	Ties (top) %	min. score	Ties (bottom)	Ties (bottom)%	Best time
F	50	Butterfly	7193	1000	3	0.0	0	32	0.4	27.98
F	100	Butterfly	8136	1000	4	0.0	0	20	0.2	61.91
F	200	Butterfly	224	981	0	0.0	0	0	0.0	144.07
F	50	Backstroke	7545	1000	4	0.1	0	1	0.0	30.33
F	100	Backstroke	15777	1000	7	0.0	0	51	0.3	64.05
F	200	Backstroke	731	1000	3	0.4	0	2	0.3	140.54
F	50	Breaststroke	4825	1000	2	0.0	0	12	0.2	32.14
F	100	Breaststroke	15316	1000	4	0.0	0	44	0.3	69.57
F	200	Breaststroke	743	1000	0	0.0	0	2	0.3	158.11
F	50	Freestyle	22440	1000	3	0.0	0	77	0.3	26.25
F	100	Freestyle	21983	1000	3	0.0	0	80	0.4	56.58
F	200	Freestyle	6328	1000	2	0.0	0	8	0.1	122.09
F	400	Freestyle	10762	1000	4	0.0	0	32	0.3	259.59
F	800	Freestyle	310	1000	0	0.0	0	0	0.0	527.59
F	1500	Freestyle	131	977	0	0.0	1	0	0.0	1005.51
F	100	IM	513	1000	0	0.0	3	2	0.4	65.01
F	150	IM	852	1000	3	0.4	9	1	0.1	152.14
F	200	IM	13223	1000	3	0.0	0	44	0.3	138.37
F	400	IM	176	963	0	0.0	6	0	0.0	308.86
M	50	Butterfly	10561	1000	4	0.0	0	13	0.1	24.48
M	100	Butterfly	12922	1000	4	0.0	0	116	0.9	53.72
M	200	Butterfly	416	1000	0	0.0	0	2	0.5	124.06
M	50	Backstroke	10720	1000	3	0.0	0	7	0.1	26.21
M	100	Backstroke	20417	1000	5	0.0	0	117	0.6	55.42
M	200	Backstroke	863	1000	5	0.6	1	2	0.2	123.59
M	50	Breaststroke	7049	1000	3	0.0	0	36	0.5	28.89
M	100	Breaststroke	19327	1000	4	0.0	0	95	0.5	62.97
M	200	Breaststroke	974	1000	2	0.2	0	1	0.1	141.91
M	50	Freestyle	31583	1000	4	0.0	0	144	0.5	22.44
M	100	Freestyle	30571	1000	4	0.0	0	146	0.5	48.70
M	200	Freestyle	10756	1000	4	0.0	0	12	0.1	110.34
M	400	Freestyle	14661	1000	2	0.0	0	74	0.5	234.42
M	800	Freestyle	352	1000	0	0.0	0	0	0.0	503.16
M	1500	Freestyle	289	1000	1	0.3	2	0	0.0	943.15
M	100	IM	552	1000	3	0.5	1	1	0.2	57.68
M	150	IM	2001	1000	1	0.0	0	0	0.0	141.17
M	200	IM	17157	1000	3	0.0	0	81	0.5	122.70
M	400	IM	274	1000	0	0.0	0	0	0.0	280.95

Tab. 4: Performance statistics for disability adjusted Fair World Para Point Score.

Sex	M	Stroke	n	β_0	β_1	β_2	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}
F	50	Butterfly	7193	3.3074	0.033	-0.0008	274.80	274.80	266.50	216.49	206.91	175.56	163.80	147.86	133.55	124.95	142.39	123.95	123.95	130.42
F	100	Butterfly	8136	4.9806	0.0396	-0.0010	577.18	577.18	577.18	577.18	577.18	565.06	536.59	459.93	426.05	416.27	518.25	411.77	399.26	420.25
F	200	Butterfly	224	2.836	0.0082	-0.0009	1394.93	1191.20	987.47	881.86	776.25	670.64	642.92	545.29	514.28	503.20	515.48	515.48	475.37	540.13
F	50	Backstroke	7545	2.8028	0.0277	-0.0007	300.97	286.20	233.20	210.40	178.80	172.30	150.70	137.82	123.89	115.90	138.16	117.66	116.77	118.10
F	100	Backstroke	15777	4.5023	0.0243	-0.0006	913.24	827.98	680.28	676.54	574.71	514.14	474.62	443.61	401.10	382.15	448.72	395.79	377.79	388.13
F	200	Backstroke	731	2.6908	0.0217	-0.0007	859.88	1257.71	859.88	859.88	794.25	664.75	638.04	576.99	528.56	491.69	566.31	512.20	500.31	512.01
F	50	Breaststroke	4825	3.1042	0.0413	-0.0009	425.17	339.62	269.14	226.98	207.38	203.61	186.56	160.88	146.29		160.74	140.60	140.60	148.73
F	100	Breaststroke	15316	4.3502	0.0285	-0.0005	1090.70	993.82	807.55	657.94	600.16	567.49	531.55	468.46	430.64		480.72	431.22	416.59	428.66
F	200	Breaststroke	743	3.1814	0.021	-0.0006	1279.94	1646.00	1279.94	913.88	909.98	877.94	787.82	690.53	671.58		742.58	641.75	641.75	677.01
F	50	Freestyle	22440	3.4462	0.0172	-0.0002	283.48	283.48	224.40	205.24	187.12	164.31	149.99	137.73	124.38	118.07	135.04	119.59	114.21	119.56
F	100	Freestyle	21983	3.7647	0.0141	-0.0002	639.85	639.85	525.23	475.86	429.52	382.07	349.17	323.67	291.41	278.09	326.21	284.32	274.18	282.49
F	200	Freestyle	6328	3.3765	0.0098	-0.0002	1264.86	1264.86	1056.34	890.92	808.97	726.26	669.46	608.12	570.57	534.70	597.17	536.62	536.62	546.40
F	400	Freestyle	10762	5.7184	0.0184	-0.0005	3144.15	4591.49	3144.15	2705.40	2560.04	2327.60	2239.59	2061.56	1952.58	1864.99	2168.24	1928.41	1829.77	1903.40
F	800	Freestyle	310	2.8403	0.0262	-0.0010	3536.69	3312.01	3087.33	2862.64	2637.96	2413.28	2310.89	2140.41	1980.21	1953.48	2245.33	2022.22	1893.49	2088.83
F	1500	Freestyle	131	1.066	0.0289	-0.0016	3281.73	3059.29	2836.84	2614.40	2391.95	2169.51	1868.56	1763.64	1654.60	1654.22	1783.45	1560.63	1492.82	1733.58
F	100	IM	513	3.1488	0.0591	-0.0021	977.85	736.67	495.49	460.67	425.84	378.02	363.75	325.34	299.94	272.47	312.56	271.36	270.96	282.80
F	150	IM	852	2.8215	0.0318	-0.0008	1023.80	993.94	783.57	725.73										
F	200	IM	13223	5.5932	0.022	-0.0005	1764.27	1738.38	1712.49	1582.74	1461.95	1288.47	1248.34	1152.98	1046.49	1001.46	1150.13	1018.00	980.77	1005.67
F	400	IM	176	2.9909	0.0083	-0.0006	2491.31	2303.33	2115.35	1927.37	1739.39	1551.41	1486.44	1224.99	1181.54	1126.58	1305.12	1144.52	1068.78	1173.01
M	50	Butterfly	10561	3.4944	0.0373	-0.0010	361.61	321.47	281.33	204.61	168.66	157.58	146.91	134.06	122.03	117.34	124.59	116.34	114.43	119.88
M	100	Butterfly	12922	5.5748	0.0587	-0.0016	1366.88	1140.25	913.63	612.42	596.94	510.62	510.62	444.18	418.88	399.00	452.35	402.49	395.05	-403.50
M	200	Butterfly	416	3.6696	0.0494	-0.0021	928.57	913.44	898.30	883.17	868.04	852.00	780.40	636.75	611.61	593.52	654.08	600.70	589.71	609.31
M	50	Backstroke	10720	2.8427	0.0271	-0.0006	278.08	231.05	205.99	184.01	152.52	151.04	136.68	123.64	109.52	105.11	123.46	104.56	100.48	108.65
M	100	Backstroke	20417	4.5914	0.0358	-0.0009	796.46	704.91	620.80	620.80	496.29	472.50	430.47	397.66	363.54	349.72	405.26	354.60	342.44	355.06
M	200	Backstroke	863	3.3807	0.024	-0.0009	1336.85	1170.29	1092.07	1077.52	799.80	710.45	690.11	637.66	570.01	531.11	661.92	576.28	520.46	589.98
M	50	Breaststroke	7049	3.2799	0.0546	-0.0012	408.97	286.72	240.25	224.20	199.76	177.91	169.34	147.23	137.49		150.74	136.99	130.96	138.23
M	100	Breaststroke	19327	4.4155	0.0478	-0.0010	1112.79	737.65	636.83	599.50	562.97	503.03	477.15	423.35	392.65		433.93	391.56	377.12	383.56
M	200	Breaststroke	974	3.4185	0.0416	-0.0013	1486.11	1267.45	1043.34	981.38	880.30	816.23	785.66	667.09	656.98		718.51	663.83	645.49	656.59
M	50	Freestyle	31583	4.0078	0.023	-0.0005	379.51	305.34	234.00	208.21	181.39	163.94	148.36	138.17	126.68	119.62	133.26	119.63	117.27	120.61
M	100	Freestyle	30571	4.2669	0.0233	-0.0006	839.23	668.41	575.78	480.42	409.52	379.96	347.28	320.84	294.30	276.83	318.26	281.60	274.91	280.11
M	200	Freestyle	10756	3.9701	0.008	-0.0003	1481.12	1277.08	1076.44	920.06	823.05	746.30	702.89	650.10	591.22	555.53	657.30	605.75	561.55	558.04
M	400	Freestyle	14661	5.8788	0.0138	-0.0003	5312.59	4197.86	3335.15	2613.84	2409.10	2190.22	2083.58	1937.69	1801.94	1728.86	1985.80	1813.58	1738.53	1773.88
M	800	Freestyle	352	2.6516	0.0175	-0.0009	4566.15	3824.41	3082.67	2340.93	2011.18	2001.94	1938.01	1914.96	1727.68	1658.79	1986.23	1756.55	1683.53	1790.99
M	1500	Freestyle	289	2.8732	0.0179	-0.0008	9543.73	8239.42	6935.11	6935.11	4892.95	4389.96	4016.00	3954.45	3724.47	3244.85	3877.04	3877.04	3429.56	3657.68
M	100	IM	552	3.5427	0.0468	-0.0015	870.48	725.11	579.74	434.37	371.51	359.60	338.58	303.25	279.69	269.72	309.77	260.84	258.36	271.27
M	150	IM	2001	3.7092	0.0272	-0.0005	1284.62	1182.91	904.55	788.92										
M	200	IM	17157	5.6066	0.0337	-0.0009	2106.04	1916.56	1727.09	1420.06	1249.38	1168.28	1104.14	1017.36	946.71	905.09	1017.11	918.96	891.28	901.82
M	400	IM	274	4.3396	0.0282	-0.0011	2449.75	2371.87	2293.99	2216.11	2138.23	2013.57	1707.26	1586.74	1521.46	1479.93	1668.47	1438.54	1406.95	1556.75

Tab. 5: Estimated coefficients for the disability and age adjusted Fair World Para Point Score ($a = 1187$).

Sex	M	Stroke	n	max. score	Ties (top)	Ties (top)%	min. score	Ties (bottom)	Ties (bottom)%
F	50	Butterfly	7193	1000	2	0.0	0	29	0.4
F	100	Butterfly	8136	1000	1	0.0	0	26	0.3
F	200	Butterfly	224	1000	0	0.0	119	0	0.0
F	50	Backstroke	7545	1000	2	0.0	0	4	0.1
F	100	Backstroke	15777	1000	7	0.0	0	73	0.5
F	200	Backstroke	731	972	0	0.0	27	0	0.0
F	50	Breaststroke	4825	995	0	0.0	0	14	0.3
F	100	Breaststroke	15316	1000	0	0.0	0	83	0.5
F	200	Breaststroke	743	911	0	0.0	0	0	0.0
F	50	Freestyle	22440	1000	2	0.0	0	115	0.5
F	100	Freestyle	21983	1000	4	0.0	0	111	0.5
F	200	Freestyle	6328	1000	2	0.0	0	5	0.1
F	400	Freestyle	10762	1000	2	0.0	0	36	0.3
F	800	Freestyle	310	1000	0	0.0	105	0	0.0
F	1500	Freestyle	131	1000	0	0.0	378	0	0.0
F	100	IM	513	1000	0	0.0	36	0	0.0
F	150	IM	852	989	0	0.0	11	2	0.2
F	200	IM	13223	1000	1	0.0	0	58	0.4
F	400	IM	176	832	0	0.0	168	0	0.0
M	50	Butterfly	10561	1000	1	0.0	0	14	0.1
M	100	Butterfly	12922	1000	1	0.0	0	154	1.2
M	200	Butterfly	416	1000	0	0.0	50	0	0.0
M	50	Backstroke	10720	1000	3	0.0	0	8	0.1
M	100	Backstroke	20417	1000	5	0.0	0	132	0.6
M	200	Backstroke	863	1000	0	0.0	25	0	0.0
M	50	Breaststroke	7049	1000	0	0.0	0	33	0.5
M	100	Breaststroke	19327	1000	0	0.0	0	213	1.1
M	200	Breaststroke	974	1000	0	0.0	16	0	0.0
M	50	Freestyle	31583	1000	5	0.0	0	165	0.5
M	100	Freestyle	30571	1000	4	0.0	0	152	0.5
M	200	Freestyle	10756	1000	2	0.0	0	10	0.1
M	400	Freestyle	14661	1000	0	0.0	0	86	0.6
M	800	Freestyle	352	1000	0	0.0	124	0	0.0
M	1500	Freestyle	289	1000	1	0.3	139	0	0.0
M	100	IM	552	1000	0	0.0	32	0	0.0
M	150	IM	2001	1000	0	0.0	0	1	0.0
M	200	IM	17157	1000	1	0.0	0	87	0.5
M	400	IM	274	853	0	0.0	68	0	0.0

Tab. 6: Performance statistics for the disability and age adjusted Fair World Para Point Score.

10 Appendix B

For the examples of implementation please see our repository at https://doi.org/10.17605/OSF.IO/SJ958

Note, that in the simple single distribution model, the constraint $x_i \leq x^*$ for $i = 1, \ldots, n$ is equivalent to the more concise constraint $\max_i(x_i) \leq x^*$. However, for the linear programming algorithm, it makes more sense to define this constraint in terms of a linear combination of b and c. Hence

$$b - cx_i \geq \log(-\log \frac{1000}{a})) \quad \text{for } i = 1, \ldots, n.$$

Similarly, the constraint in equation (12) can be rewritten as

$$\beta_{k,0} + \beta_{k,1}\text{age}_i + \beta_{k,2}\text{age}_i^2 - c_{k,g}x_i \geq \log(-\log \frac{1000}{a_k})) \quad \text{for } i = 1, \ldots, n.$$

```
pgumbel <- function(x,b,c){exp(-exp(b-c * x))}

# pdf
dgumbel <- function(x,b,c){c * exp(b-c * x) * exp(-exp(b-c * x))}

# pdf on a log-scale
dgumbel.log <- function(x,b,c){log(c)+(b-c * x)-exp(b-c * x)}

# random number generator (via inverse sampling)
rgumbel <- function(n,b,c){
  u <- runif(n)
  x <- (b-log(-log(u)))/c
  return(x)
}

library(quadprog)

### DISABILITY CLASSES ONLY

my.Q.classes <- function(p.raw,a,x,g){
  # p are the raw percentiles
  # a is the common truncation parameter
  # x are the observations
  # g are the groups (assumed to go from 1 to 14)

  Y <- log(-log(p.raw * 1000/a))

  X <- array(dim=c(length(x),15))

  # 1st is intercept
  X[,1] <- 1

  for(j in 1:14){
    X[,j+1] <- (g == j) * x
  }


  # if some classes are missing
```

```
is.ma <-  which(table(factor(g,levels=1:14))==0)
is.not.ma <- which(table(factor(g,levels=1:14))>0)
if(length(is.ma)!=0){ X <- X[,c(1,1+is.not.ma)]}

Rinv <- solve(chol(t(X) %*% X))
d <- t(Y) %*% X

### constraints
xmax <- tapply(x,g,max)

# it is OK to have tautological constraints, presumably

# all c coefs are negative

A.neg <- diag(c(0,rep(-1,length(is.not.ma))))[-1,]; b.neg <- rep(0,length(is.not.ma))

# ordering for S1-10, S11-13, S14
A.ord <- t(array(c(
  0,-1,1,0,0,0,0,0,0,0,0,0,0,0,0,
  0,0,-1,1,0,0,0,0,0,0,0,0,0,0,0,
  0,0,0,-1,1,0,0,0,0,0,0,0,0,0,0,
  0,0,0,0,-1,1,0,0,0,0,0,0,0,0,0,
  0,0,0,0,0,-1,1,0,0,0,0,0,0,0,0,
  0,0,0,0,0,0,-1,1,0,0,0,0,0,0,0,
  0,0,0,0,0,0,0,-1,1,0,0,0,0,0,0,
  0,0,0,0,0,0,0,0,-1,1,0,0,0,0,0,
  0,0,0,0,0,0,0,0,0,-1,1,0,0,0,0,
  0,0,0,0,0,0,0,0,0,0,0,-1,1,0,0,
  0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0),
  dim=c(15,11)))

b.ord <- rep(0,11)

A.ord <- A.ord[,c(1,1+is.not.ma)]
A.ord <- A.ord[apply(A.ord!=0,1,sum)==2,]

b.ord <- rep(0,dim(A.ord)[1])

# truncation restrictions

Amat.bot <- cbind(1,diag(xmax))
b.bot <- rep(log(-log(1000/a)),dim(Amat.bot)[1])


m.quad <- solve.QP(Dmat = Rinv, factorized = TRUE, dvec = d,
                   Amat = cbind(t(A.neg),t(A.ord),t(Amat.bot)),
                   bvec = c(b.neg,b.ord,b.bot), meq = 0)

# repairing for the missing coefficients

b.est <- m.quad$solution[1]
```

```
  c.est <- rep(NA,14)
  c.est[is.not.ma] <- -m.quad$solution[-1]

  return(list(b.est=b.est,c.est=c.est))
}


### AGE ONLY


my.Q.age <- function(p.raw,a,x,age){
  # p are the raw percentiles
  # a is the common truncation parameter
  # x are the observations
  # age are the age-groups

  Y <- log(-log(p.raw * 1000/a))

  # we now have 4 coefficients

  X <- array(dim=c(length(x),4))

  # 1st is intercept
  X[,1] <- 1

  # 2nd is linear in age
  X[,2] <- age
  # 3rd is quadratic in age
  X[,3] <- age^2

  # 4th is slope of x
  X[,4] <- x


  Rinv <- solve(chol(t(X) %*% X))
  d <- t(Y) %*% X

  # constraints
  # b2 is negative (optimal age for speed)
  # c is negative

  Amat.top <- t(array(c(
    0,0,-1,0,
    0,0,0,-1),dim=c(4,2)))


  b.top <- c(0,0)

  # maximum cdf is smaller than the truncated one

  Amat.bot <- X
```

```
  b.bot <- rep(log(-log(1000/a)),length(x))

  m.quad <- solve.QP(Dmat = Rinv, factorized = TRUE, dvec = d,
                     Amat = t(rbind(Amat.top,Amat.bot)),
                     #Amat=t(Amat.top), bvec=b.top,
                     bvec = c(b.top,b.bot),
                     meq = 0)
  return(m.quad)
}


### DISABILITY CLASSES AND AGE

my.Q.age.class <- function(p.raw,a,x,age,age.orig,class){
  # p are the raw percentiles
  # a is the common truncation parameter
  # x are the observations
  # age are the age-groups
  # class are classes

  Y <- log(-log(p.raw * 1000/a))

  # we now have 4 coefficients

  X <- X.orig <- array(dim=c(length(x),17))

  # 1st is intercept
  X[,1] <- X.orig[,1] <- 1

  # 2nd is linear in age
  X[,2] <- age
  X.orig[,2] <- age.orig

  # 3rd is quadratic in age
  X[,3] <- age^2
  X.orig[,3] <- age.orig^2

  # 4th is slope of x times class dummy
  for(j in 4:17){
    X[,j] <- X.orig[,j] <- (class==j-3) * x
  }

  # if some classes are missing
  is.not.ma <- which(table(factor(class,levels=1:14))>0)
  is.ma <- which(table(factor(class,levels=1:14))==0)
  if(length(is.ma)!=0){ X <- X[,c(1:3,3+is.not.ma)]}
  if(length(is.ma)!=0){ X.orig <- X.orig[,c(1:3,3+is.not.ma)]}

  Rinv <- solve(chol(t(X) %*% X))
  d <- t(Y) %*% X
```

```
  ### constraints
  xmax <- tapply(x,class,max)

  # it is OK to have tautological constraints, presumably

  # all c coefs are negative

  A.neg <- diag(c(0,0,0,rep(-1,length(is.not.ma))))[-(1:3),]; b.neg <- rep(0,length(is.not.ma))

  # ordering for S1-10, S11-13, S14
  A.ord <- t(array(c(
    0,0,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0, # b2 < 0
    0,0,0,-1,1,0,0,0,0,0,0,0,0,0,0,0,0,
    0,0,0,0,-1,1,0,0,0,0,0,0,0,0,0,0,0,
    0,0,0,0,0,-1,1,0,0,0,0,0,0,0,0,0,0,
    0,0,0,0,0,0,-1,1,0,0,0,0,0,0,0,0,0,
    0,0,0,0,0,0,0,-1,1,0,0,0,0,0,0,0,0,
    0,0,0,0,0,0,0,0,-1,1,0,0,0,0,0,0,0,
    0,0,0,0,0,0,0,0,0,-1,1,0,0,0,0,0,0,
    0,0,0,0,0,0,0,0,0,0,-1,1,0,0,0,0,0,
    0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0,0,0,
    0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0,0,
    0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0),
    dim=c(17,12)))

  b.ord <- rep(0,12)

  A.ord <- A.ord[,c(1:3,3+is.not.ma)]
  A.ord <- A.ord[apply(A.ord!=0,1,sum)==2,] # potentially cuts out a middle restriction, so not ent
  # ... but seems to be working.

  b.ord <- rep(0,dim(A.ord)[1])

  # truncation restrictions

  Amat.bot <- X.orig
  b.bot <- rep(log(-log(1000/a)),length(x))

  m.quad <- solve.QP(Dmat = Rinv, factorized = TRUE, dvec = d,
                     Amat = cbind(t(A.neg),t(A.ord),t(Amat.bot)),
                     bvec = c(b.neg,b.ord,b.bot), meq = 0)
  b.est <- m.quad$solution[1:3]
  c.est <- rep(NA,14)
  c.est[is.not.ma] <- m.quad$solution[-(1:3)]

  return(list(b.est=b.est,c.est=c.est))
}
```

References

[1] Abbott, S., C. Hogan, M. T. Castiglioni, G. Yamauchi, L. J. G. Mitchell, J. Salter, M. Romann, and S. Cobley (2021). Maturity-related developmental inequalities in age-group swimming: The testing of 'mat-caps' for their removal. *Journal of Science and Medicine in Sport 24*(4), 397–404.

[2] Altimira, D., J. Clark, F. Mueller, M. Billinghurst, and C. Bartneck (2017). Enhancing player engagement through game balancing in digitally augmented physical games. *International Journal of Human Computer Studies 103*(July 2017), 35–47.

[3] Austin, M. W. (2010). Why winning matters. *Think 9*(26), 99–102.

[4] Berry, S. M., C. S. Reese, and P. D. Larkey (1999). Bridging different eras in sports. *Journal of the American Statistical Association 94*(447), 661–676.

[5] Burkett, B., C. Payton, P. Van de Vliet, H. Jarvis, D. Daly, C. Mehrkuehler, M. Kilian, and L. Hogarth (2018). Performance characteristics of para swimmers: How effective is the swimming classification system? *Physical Medicine and Rehabilitation Clinics 29*(2), 333–346.

[6] Casella, G. and R. L. Berger (2021). *Statistical inference*. Cengage Learning.

[7] Cobley, S., S. Abbott, J. Eisenhuth, J. Salter, D. McGregor, and M. Romann (2019a). Removing relative age effects from youth swimming: The development and testing of corrective adjustment procedures. *Journal of Science and Medicine in Sport 22*(6), 735–740.

[8] Cobley, S., S. Abbott, J. Eisenhuth, J. Salter, D. McGregor, and M. Romann (2019b). Removing relative age effects from youth swimming: The development and testing of corrective adjustment procedures. *Journal of Science and Medicine in Sport 22*(6), 735–740.

[9] Daly, D. J., D. J. Daly, and Y. Vanlandewijck (1999). Some criteria for evaluating the "fairness" of swimming classification. *Adapted Physical Activity Quarterly 16*(3), 271–289.

[10] Daly, D. J., S. K. Djobova, L. A. Malone, Y. Vanlandewijck, and R. D. Steadward (2003). Swimming speed patterns and stroking variables in the paralympic 100-m freestyle. *Adapted Physical Activity Quarterly 20*(3), 260.

[11] Doane, D. P., L. E. Seward, and K. Murphy (2022). Gender and age patterns in nsga swim competitions. *Journal of Aging Research 2022*, 8459520.

[12] Donato, A. J., K. Tench, D. H. Glueck, D. R. Seals, I. Eskurza, and H. Tanaka (2003). Declines in physiological functional capacity with age: a longitudinal study in peak swimming performance. *Journal of Applied Physiology 94*(2), 764–769.

[13] Fairbrother, J. T. (2007). Age-related changes in top-ten men's u.s. masters 50-m freestyle swim times as a function of finishing place. *Perceptual and Motor Skills 105*(3), 1289–1293.

[14] Gomes, D. T. and L. Henriques-Rodrigues (2019). Swimming performance index based on extreme value theory. *International Journal of Sports Science & Coaching 14*(1), 51–62.

[15] Hansen, J. (2015). The origin of the term handicap in games and sports – history of a concept. *Physical Culture and Sport. Studies and Research 65*(1), 7–13.

[16] Hogarth, L., V. Nicholson, C. Payton, and B. Burkett (2021). Modelling the age-related trajectory of performance in para swimmers with physical, vision and intellectual impairment. *Scandinavian Journal of Medicine & Science in Sports 31*(4), 925–935.

[17] Hogarth, L., C. Payton, P. Van de Vliet, M. Connick, and B. Burkett (2018). A novel method to guide classification of para swimmers with limb deficiency. *Scandinavian Journal of Medicine & Science in Sports 28*(11), 2397–2406.

[18] International Paralympic Committee (2017). World para swimming points system.

[19] Knechtle, B., T. M. Barbosa, and P. T. Nikolaidis (2019). The age-related changes and sex difference in master swimming performance. *Mov Sport Sci/Sci Mot* (104), 29–36.

[20] Kollarz, C., B. Knechtle, C. A. Rüst, and T. Rosemann (2013). Comparison of age of peak swimming speed in elite backstroke swimmers at national and international level. *OA Sports Medicine 1*(2), 19.

[21] Krakovsky, M. (2022, jul). Formalizing fairness. *Commun. ACM 65*(8), 11–13.

[22] König, S., F. Valeri, S. Wild, T. Rosemann, C. A. Rüst, and B. Knechtle (2014). Change of the age and performance of swimmers across world championships and olympic games finals from 1992 to 2013 – a cross-sectional data analysis. *SpringerPlus 3*(1), 652.

[23] Mann, D. L., S. M. Tweedy, R. C. Jackson, and Y. C. Vanlandewijck (2021). Classifying the evidence for evidence-based classification in paralympic sport. *Journal of Sports Sciences 39*(sup1), 1–6.

[24] Mueller, F. F., D. Edge, F. Vetere, M. R. Gibbs, S. Agamanolis, B. Bongers, and J. G. Sheridan (2011). Designing sports: A framework for exertion games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, New York, NY, USA, pp. 2651–2660. Association for Computing Machinery.

[25] Nocedal, J. and S. J. Wright (2006). *Quadratic Programming*, pp. 448–492. New York, NY: Springer New York.

[26] Picolin, A. (2014a). IPC athletics raza point score system – parameter determination procedure 2014. Report, International Paralympic Committee.

[27] Picolin, A. (2014b). IPC athletics raza point score system – review 2014. *International Paralympic Committee*.

[28] Puce, L., L. Marinelli, I. Pallecchi, L. Mori, and C. Trompetto (2020). Impact of the 2018 world para swimming classification revision on the race results in international paralympic swimming events. *German Journal of Exercise and Sport Research 50*(2), 251–263.

[29] Ransdell, L. B., J. Vener, and J. Huberty (2009). Masters athletes: An analysis of running, swimming and cycling performance by age and gender. *Journal of Exercise Science & Fitness 7*(2, Supplement), S61–S73.

[30] Rantini, D. and N. Iriawan (2021). On the reversible jump markov chain monte carlo (rjmcmc) algorithm for extreme value mixture distribution as a location-scale transformation of the weibull distribution. *Applied Sciences 11*(16), 7343.

[31] Rubin, R. T., S. Lin, A. Curtis, D. Auerbach, and C. Win (2013). Declines in swimming performance with age: a longitudinal study of masters swimming champions. *Open access journal of sports medicine 4*, 63–70.

[32] Serrano-Dura, J., P. Molina, and A. Martinez-Baena (2021). Systematic review of research on fair play and sporting competition. *Sport, Education and Society 26*(6), 648–662.

[33] Tanaka, H. and D. R. Seals (1997). Age and gender interactions in physiological functional capacity: insight from swimming performance. *Journal of Applied Physiology 82*(3), 846–851.

[34] Turlach, B. A. and A. Weingessel (2011). *quadprog: Functions to solve Quadratic Programming Problems.* R package version 1.5-4.

[35] Tweedy, S. M. and Y. C. Vanlandewijck (2011a). International paralympic committee position stand-background and scientific principles of classification in paralympic sport. *British Journal of Sports Medicine 45*(4), 259.

[36] Tweedy, S. M. and Y. C. Vanlandewijck (2011b). International paralympic committee position stand—background and scientific principles of classification in paralympic sport. *British Journal of Sports Medicine 45*(4), 259–269.

[37] Uscinowicz, N., W. Seidel, P. Zostawa, and S. Klich (2013). Objective and parametric methods used in functional classification disabled swimmers. *Fizjoterapia 21*(3), 50–59.

[38] Zingg, M. A., C. A. Rüst, T. Rosemann, R. Lepers, and B. Knechtle (2014). Analysis of swimming performance in fina world cup long-distance open water races. *Extreme Physiology & Medicine 3*(1), 2.