Study

Phenomenology of Sprinting and Endurance: Toward a Uniform Performance Assessment Model

by Wim Westera

ABSTRACT

According to the IAAF scoring tables, Usain Bolt's 100m world record of 9.58 sec is worth 1374 points, whereas Kenenisa Bekele's 10,000m world record of 26:17.59 yields only 1295 points. This demonstrates the immanent weakness and unfairness of the current scoring methodology. This paper studies the relationship between running distance and running speed, and proposes an alternative scoring method. First it presents the personal predictor model (PPM). This model uses two personal bests of an athlete for calibration, and then it allows predicting an athlete's hypothetical personal bests for any other distance. The accuracy is well below 1% and it thereby greatly outperforms existing models. Second, it presents the normalised multi-event scoring model (NMSM). This model overcomes the manifest flaws of the current IAAF scoring tables; it demonstrates greater fairness, consistency and transparency. The impact of the new model is explained using empirical data. It substantiates the need for replacing the existing IAAF scoring tables. Finally, the two explained models (PPM and NMSM) are combined for composing personalised scoring tables. These tables convert an athlete's performances for any distance into a single score, which allows for a ranking of the athlete's performances across various distances.

AUTHOR

Wim Westera is a full professor at the Open University of the Netherlands. His background is in physics and digital media for learning. He is Director of the Learning Media Research Programme of the Centre for Learning Sciences and Technologies. He is a masters athlete and racing cyclist.

Introduction

It is well established that average running velocity decreases with racing distance. In a 10,000m the average speed of top athletes is 20% lower than what is achieved in an 800m. This is not without logic. If this were not the case, the 10,000m runner could do a 9,200m race-pace warm-up and still beat the 800m specialists during the last two laps. Apparently, for races that require prolonged efforts, the internal physiology of energy and power reinforces effort levels that are well below maximum capacity. Naturally the same holds for swimming, speed skating and any other sports that involve traversing a certain distance within the fastest possible time.

Comparison of performances across different running events is a delicate topic. How would one ever be capable of comparing Usain Bolt's outstanding sprinting performances with the likewise outstanding 10,000m world record of Kenenisa Bekele? This would be like comparing apples and pears. Yet, in athletics this is common practice, even within the official framework of the sport's governing body, the International Association of Athletics Federations (IAAF).

The decathlon would be a good case in point. The IAAF uses a detailed set of tables for converting performances in the various events to a single numerical value so that these can be added into a result score. For instance a high jump of 2.00m and 400m in 50.0 sec would yield a result score of (803 + 815)=1618 points.

WESTERA[1] has severely criticised the validity of the decathlon tables. He argues that the tables display unacceptable bias since they favour some events over others. Sprinting-based performances yield disproportionately higher scores than the throwing events or the 1500m. Historical bias and frequent changes seem to demonstrate the arbitrariness of performance alignments in the decathlon tables [2]. In fact, comparison to an alternative scoring model proposed by the author leads to the conclusion that the official tables have assigned the world record to the wrong athlete.

Similar objections hold for the other IAAF tables that are used for comparing different running events or for comparing athletes from different age groups. The IAAF Technical Committee, which is responsible for the validity of these tables, needs to regularly make modifications because of apparent irregularities in the relationship between results and assigned points.

But even in the latest version[3] inconsistencies and suspect data can easily be tracked. Figure 1 displays the result scores that the IAAF assigns to the diverse track running world records.

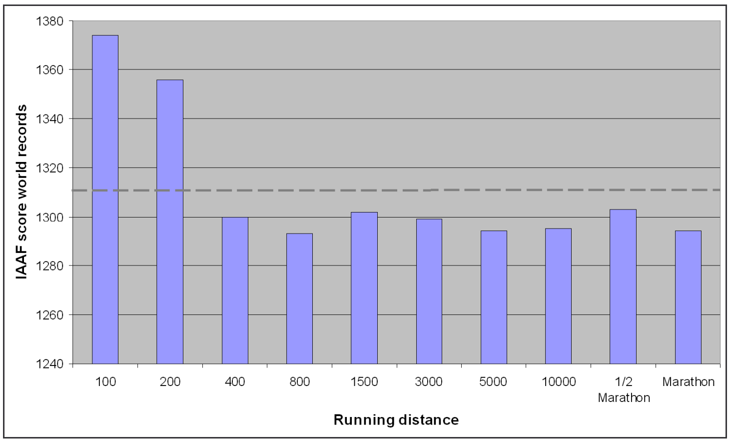

Figure 1: IAAF scores for men's track running world records

The large variability of the ratings demonstrates the weakness of the tables. The 100m world record is rated up to 80 points higher than the long distance records. The different ratings would indicate that Kenenisa Bekele is an amateur when compared with Usain Bolt. It is hard to understand why the 5,000m and the 10,000m world records are less valued. The world records, indicating the ultimate limits of human power, would be the perfect reference for aligning the scoring tables. Therefore each world record should, in principle, receive the same score. For unclear reasons, however, sprinting is highly overrated in the IAAF tables. Besides this, other unexplained differences are manifest in Figure 1.

This paper will examine this rating problem. First it will elaborate the relationship between running distance and running speed. It will explain the personal prediction model (PPM), an analytical model that can be used for determining an athlete's personal best times at various distances, provided that two best times are available for calibration. Next, the paper will present the normalised multi-event scoring model (NMSM). This model overcomes the manifest flaws of the current IAAF scoring tables. The impact of the new model will explained. Finally, the two separate models will be combined for composing personalised scoring tables. These tables convert an athlete's performances for any distance into a single score, which allows composing an absolute ranking of the athlete's performances across various distances.

Research Approaches

To able to properly compare the quality of performances in different running events, one needs insight into the relationship between human performance and the duration of the efforts involved. Although quite some research has been devoted to the topic, only very little is known at an analytical level. If a formula were available, one might - starting from some individual characteristics – be able to predict the athlete's personal performance limits for a range of distances. In the ideal (but unrealistic) case, one would just collect an athlete's shoe size, weight, height, leg length, lung capacity or any other relevant characteristic, and then use the formula to calculate the athlete's anticipated performances on various distances. With these predictions one might be able to decide what distances offer the best chances and devise personalised training schedules for these. Unfortunately, things aren't that straightforward.

In the studies that have been done three different approaches can be distinguished: 1) physiological models, 2) statistical models, 3) phenomenological models.

Physiological models

These approaches search for explanations based on the underlying physiological processes. From various studies it is known that endurance performance is directly related with maximal aerobic power[4]. Measurement of maximum aerobic running speed or speed at maximal oxygen uptake can be used for predicting performances in the range from middle distances to long distances[5], covering the range from say 1500m up to the marathon. At shorter distances predictions tend to be greatly unreliable because of disturbing interferences of anaerobic metabolisms. Bundle et al.[6] proposed a model that combined the physiological limits of anaerobic and aerobic power for predicting performances in both the sprinting and mid-range distances (ranging from a few seconds to a few minutes). They suggested a simple negative exponential relationship between velocity and running duration, and incorporated the different time scales that anaerobic and aerobic metabolisms are active. The approach isn't very accurate though: the predictions deviate on average well above 3% from realised performances. This may seam a negligible percentage, but for a 15 minute effort duration it would mean an uncertainty of plus or minus half a minute. For athletes such inaccuracy doesn't make sense. An additional disadvantage is that tests for assessing the athlete's running velocity at maximal aerobic and anaerobic power are required. In practice, such tests are inaccurate as such, and because of the required maximum effort the administering of the tests may easily interfere with the pursued training approach. Various researchers carried out biomechanical studies of the running start-up process[7], but these are mostly concerned with the dynamics of body angle, the required metabolic power and the techniques for start-up optimisation, neglecting the speed-distance relationship.

Statistical models

Likewise, progressive scoring tables based on statistical processing of large numbers of performance data are known to be inaccurate and unreliable. The IAAF provides and uses such scoring tables for several purposes: to determine the result score of a performance for the World Rankings, to evaluate the competitions, to make the best athlete award in a specific competition, and to produce national, club, school, etc. rankings[3]. HARDER[8] tried to produce better tables by considering population fractions achieving a certain performance level. Calibration of the fractions between different events enables statistical mapping for inter-event comparison. Although Harder's tables deviate from the IAAF-tables, they correlate very well with these, and thus unfortunately display the same inaccuracies. The statistical approaches have two things in common. First, they reflect a population-based average representing a mixture of many different human features and conditions. Such general approach may severely affect their applicability for individuals. Second, they are phenomenological in kind and do not rely on an underlying theory that would improve our understanding of the mechanisms involved.

Phenomenological models

Phenomenological models use mathematical expressions for describing observed phenomena. The models should be consistent with underlying theories, but don't necessarily include these. This reflects both a strength and a weakness. Blankly starting from manifest phenomena while neglecting underlying theories may help greatly reduce complexity, which makes it more feasible to find solutions. Inevitably, this neglect goes at the expense of explanatory power. KATZ & KATZ[9] demonstrated that athletics world records can be covered by power relationships between distance and exertion time. It should be noted that the IAAF scoring tables are all based on power laws that are calibrated via population statistics. STANKIEWICZ[10] used a power law dependence between velocity an distance for the comparison of decays in energy over time in road and track events. In all cases severe anomalies were observed. WESTERA[11] introduced a phenomenological model for describing the dependence of running velocity and duration. Basically, it uses exponential decay rather than a power law. The accuracies are claimed to be typically around 1%, which is much better than existing models (typically 3% or higher). The model uses a limited set of "mechanical" presuppositions and then uses an interpolation technique for predicting an athlete's personal best. It doesn't use any physiological or biomechanical test data. Instead, the model uses two personal bests of an athlete for calibration, and then it allows predicting an athlete's hypothetical personal bests for any other distance. The model is based on a first order estimate of the way lap time (which is equivalent with reciprocal speed) increments with total distance. Also, the model accounts for delays that occur during start-up. This way the model covers the entire range including endurance and sprinting distances. The model was validated with empirical data of four different groups of athletes: world-class male athletes, world class female athletes, committed male sub-elite athletes and committed female sub-elite athletes. WESTERA[3] also proposed a phenomenological model for the accurate calculation of scores in the decathlon. The model converts performances for any of the events into a numerical result score, so that a total score over the events can be calculated by addition.

This paper builds on both the prediction model and the scoring model to improve the overall quality of scoring tables (cf. the anomalies in Figure 1). Below we will first briefly explain the two models.

The Personal Prediction Model

The personal prediction model (PPM)[11] covers two different issues: 1) the general relationship between running velocity and running distance, 2) corrective formulas for delays that occur during start-up. Such a split has been suggested before in order to explain anomalies as a result of the different metabolic processes for sprinting (anaerobic) and endurance (aerobic)[12].

Connecting average running speed and running distance

The PPM considers the total time (t) and total distance (s) of a race. But it does not reflect the dynamics during the race, for instance intervening accelerations or weakening. We introduce the average lap time *L* (this corresponds with reciprocal running velocity), which is given by:

$$L = \frac{t}{s} \quad . \qquad (1)$$

As a first order estimate the lap time increment *dL* at fixed t is assumed to be reciprocally proportional to the distance s, yielding

$$dL = \alpha \cdot \frac{ds}{s} \quad , \qquad (2)$$

where α is a constant.

Integration over *s* gives a simple logarithmic expression

$$L = \alpha \cdot \ln(\frac{s}{\beta}) \quad , \tag{3}$$

where *β* is a constant. The first order approximation reflected in equations (1) - (3) is not only theoretically grounded. Empirical evidence of its appropriateness can be found by using some existing data. This is done in Figure 2, which displays a single logarithmic plot of lap time *L* against the log of distance *s* for men's track running world records. Although the fit of the data is not superior, the (log-)linear relationship comes encouragingly close, even though data of different athletes were used.

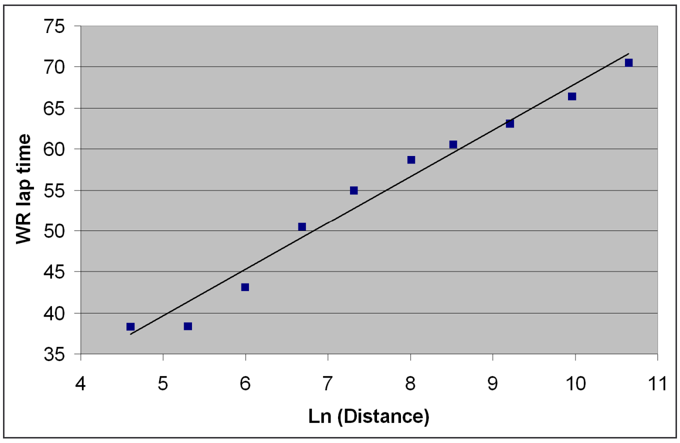

Figure 2: World records lap times against the logarithm of running distance

The straight line in figure 2 representing the world records is a special case defining the lower limits of lap time for each distance: humans cannot go any faster, at least for the time being. In the upper left part of the graph each individual athlete having its unique talents and powers may be represented with its own personal line. For determining this personal line the athlete just chooses two separate personal bests and then calculates the two associated positions in the graph. Because of the log-linear relationship the two points can then be connected with each other and thus reveal predicted personal best lap times (or velocity, or total time) for any other distance. Since mathematical formulas are available, this can be done analytically and accurately. The greatest accuracy will be obtained when the two distances that are used for the calibration span a sufficiently wide interval (interpolation). For instance, using someone's personal records at 800m and 5000m would probably make a reliable interpolation basis for forecasting the person's achievable performance on 1500m, 3000m or even an arbitrary 3968.7m. However, extrapolation, for instance when forecasting marathon performance from these points, is likely to be less accurate.

Compensating for start-up losses

The model described so far is based on cruising velocity and therefore works alright for the middle- and long-distances. For sprint distances, say up to 400m, time delays incurred during the early phase of the race when the athlete has to accelerate from standstill to cruising velocity will greatly confound the outcomes. Errors up 5% are reported by WESTERA[11]. In the model so far described, it was silently assumed that cruising velocity is average velocity, but for sprinting this doesn't quite hold: for sprinting distances cruising velocities are considerably higher than average velocities. Obviously, the difference increases at shorter distances because of the relatively long time spent on accelerating and the higher cruising velocities that the athlete tries to achieve. WESTERA[11] used available split times of various world-class sprinters[13,14]. In the 100m, the data indicate that cruising velocities are reached after about 30m. So by using the 30m split time, the cruising velocity over the last 70m can easily be calculated. The same could be done for 200m, although only few split time data could be found in the literature[15]. The disturbing effect of start-up losses will gradually disappear at longer distances. By assuming exponential decay of this correction with distance and using the 100m and 200m cruise speed data for calibration, the start-up losses can be accounted for in an analytical way.

In sum, the general prediction procedure now reads as follows:

1. Establish two sound personal bests: s_1, t_1, and s_2, t_2.
2. To account for start-up losses replace all distances s_1 and s_2 with s_1* and s_2*, respectively by using

$$s*(t) = s(t) + \gamma \cdot t \cdot e^{-\delta \cdot s} \quad , \qquad (4)$$

   with γ=21.3 and δ=0.00365 (these data come from the 100m and 200m split times)

3. Calculate the personal coefficients α and β through that determine the personal line offset and slope,

$$\alpha = \frac{\frac{t_1}{s_1*} - \frac{t_2}{s_2*}}{\ln(\frac{s_1*}{s_2*})} \qquad (5)$$

and

$$\ln(\beta) = \ln(s_1*) - \frac{\ln(\frac{s_1*}{s_2*})}{(1 - \frac{t_2 \cdot s_1*}{t_1 \cdot s_2*})} \quad . \qquad (6)$$

4. Choose a distance *s* for predicting time *t*.
5. Replace distance *s* with *s**, using equation (4).
6. Calculate predicted time *t* by using:

$$t(s*) = \alpha \cdot s* \cdot \ln(\frac{s*}{\beta}) \quad . \qquad (7)$$

Application of the PPM

A sample of cases is given in the tables below. Table 1 presents the calculations of a sample of male, middle- and long-distance world-class athletes [16].

Table 1: Performance predictions for male world-class athletes

Athlete	Distance (m)	Personal best (IAAF)	Predicted personal best	Absolute deviation
H.E.G.	1500	3:26.00	3:26.00	
	3000	*7:23.09*	*7:20.87*	*0.005*
	5000	12:50.24	12:50.24	
B.L.	1500	3:26.34	3:26.34	
	3000	*7:33.15*	*7:24.26*	*0.020*
	5000	12:59.22	12:59.22	
K.B.	3000	7:25.79	7:25.79	
	5000	*12:37.35*	*12:42.41*	*0.007*
	10000	26:17.53	26:17.53	
H.G.	1500	3:33.73	3:33.73	
	3000	*7:25.09*	*7:24.77*	*0.001*
	5000	*12:39.36*	*12:42.53*	*0.004*
	10000	26:22.75	26:22.75	
R.R.	800	1:44.05	1:44.05	
	1500	*3:29.14*	*3:32.61*	*0.017*
	3000	7:43.85	7:43.85	
D.K.	1500	3:29.46	3:29.46	
	3000	*7:20.67*	*7:20.18*	*0.001*
	5000	12:39.74	12:39.74	
K.M.	3000	7:45.44	7:45.44	
	5000	*13:13.06*	*13:15.85*	*0.004*
	10000	27:26.29	27:26.29	
J.H.	3000	7:44.40	7:44.40	
	5000	*13:21.90*	*13:18.03*	*0.005*
	10000	27:41.25	27:41.25	
K.L.	1500	3:38.83	3:38.83	
	3000	*7:52.50*	*7:44.61*	*0.017*
	5000	*13:36.10*	*13:27.44*	*0.011*
	10000	28:24.70	28:24.70	
Overall deviation				0.003

World-class athletes provide an important sample because they are usually well-prepared and perform near the limits of human capability. For each of the athletes three or occasionally four official personal bests are listed in the third column. Outer distances (shortest and longest) have been used for calibration, viz. the calculation of α and β according to equations (5) and (6). Substituting these parameters in equation (7) produces the predictions for the intermediate distance events. The predictions are quite close to the official personal bests. The average deviation (minus signs neglected) of the sample is only 0.3% (0.003, cf. bottom row of Table 1). More supportive evidence is given in WESTERA[11].

Few cases for sprint distances could be elaborated. Because of specialisation only a very few world-class sprinters excel in three disciplines (100-200-400). Yet, a few exceptions [6] could be found and these are presented in Table 2.

Table 2: Performance predictions for world-class sprinters

Athlete	Distance (m)	Personal best (IAAF)	Predicted personal best	Absolute deviation
I.S. (female)	100	0:11.10	0:11.10	
	200	*0:22.21*	*0:22.86*	*0.029*
	400	0:49.29	0:49.29	
M.K. (female)	100	0:10.83	0:10.83	
	200	*0:21.71*	*0:22.18*	*0.022*
	400	0:47.60	0:47.60	
M.J.P (female)	100	0:10.96	0:10.96	
	200	*0:21.99*	*0:22.47*	*0.022*
	400	0:48.25	0:48.25	
H.M. (male)	100	0:10.30	0:10.30	
	200	*0:20.40*	*0:21.25*	*0.042*
	400	0:45.90	0:45.90	
T.S. (male)	100	0:10.10	0:10.10	
	200	*0:19.83*	*0:20.71*	*0.045*
	400	0:44.50	0:44.50	
Overall deviation				0.03

In the sprinting range the results are a bit less accurate, but despite lacking statistics in 200m split times they are still better than those of other models.

To make these calculations, a computer programme can be devised, a preliminary version of which is available on the web[17]. Users enter their two personal bests required for calibration and enter one or more distances for which they receive their prophesised times.

The logarithmic linearity of equation (7) also offers the opportunity of a simple graphical representation of the model. This is displayed below in Figure 3. Since the vertical axis denotes lap time and the horizontal axis covers the (logarithmic) scale of distance, the performances of an individual athlete are given by a unique straight line. For reasons of convenience, performance times (derived from the product of lap time and distance) for each event are projected at the appropriate coordinates.

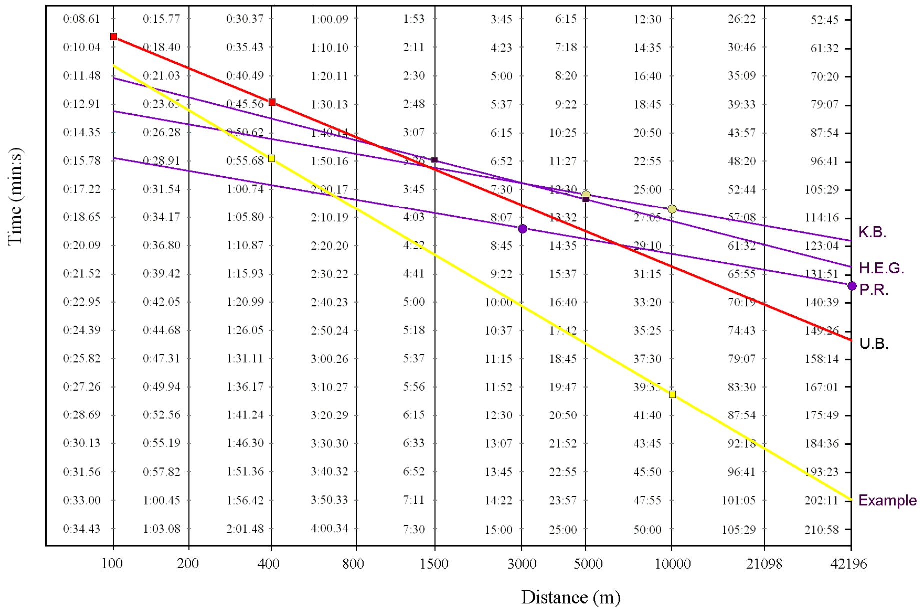

Figure 3: Graphical representation of the logarithmic model

Any individual could basically mark two different personal bests in the figure and then connect the two points with a straight line. This is illustrated by the yellow line (labelled "Example") where two calibration points are indicated with yellow squares (0:55 at 400m and 40:00 at 10,000m, respectively). The intersections with the ordinates provide predicted outcomes at the various distances (10.89 at 100m, 24.08 at 200m, 2:06.57 at 800m, 4:27.57 at 1500m, 10:02.65 at 3000m, 18:07,40 at 5,000m, 1:32:55 at the half marathon and 3:21:41 at the marathon, respectively).

The dark lines in Figure 3 are the performance curves of selected world-class athletes who cover multiple distances: Kenenisa Bekele (K.B.), Hicham El Guerrouj (H.E.G.), and Paula Radcliffe (P.R.), respectively. According to the model, Bekele is supposed to break the marathon world record (2:00:07). Also, he will be able to run 400m in 52.20 sec, which is only slightly below times recorded for his final laps in 10,000m races. Radcliffe seems to have a "weak" 10,000m best: the prognosis is 29:49 (against 30:01.09 personal best). Note that the curves of these two athletes display about the same slope, indicating the same type of decay that is probably distinctive for long distance runners. The curve of El Guerrouj shows a steeper slope, which can be attributed to the higher cruise speeds that he is capable of in the mid-range distances. One of the things that turns out is that El Guerrouj would be capable of running 10,000m in 27:17. His marathon of 2:09:10 would be respectable, but not world-class level.

The performance curve of Usain Bolt (U.B.) is also displayed in Figure 3. Here, his personal bests at 100m and 400m are used, although the latter (45.28) dates back to 2007, which is well before his breakthrough as a world class sprinter; indeed the predicted 200m time of 20.41 sec is far behind his actual best. Nevertheless, using the 100m and 200m events for calibration does not make much sense exactly because of the inaccuracies due to reinforced extrapolation. For the same reason the predictions for the longer distances (e.g. 30:39 at 10,000m) do not seem to make much sense here.

The Normalised Multi-event Scoring Model

For comparing the performances in different competitions, the IAAF uses pre-fixed scoring tables[3]. The tables are based on a power law formula of the type:

$$S = A \cdot (B - P)^C \qquad (8)$$

The formula converts a performance P into a result score S. Here A is a scaling constant, performance P is exertion time, B is a lower limit threshold performance (i.e. time above which no score is assigned), and C is a power slightly larger than 1 for obtaining progressive scoring curves. For each event different coefficients A, B and C are prescribed by IAAF[18]. For instance, for the 100m it is compulsory to use A= 25.43470, B= 18.0, C=1.810.

For decathlon scoring tables WESTERA[1] proposed three alternative models that produce better alignment. The models are all based on the idea of normalisation, that is, all performances are reduced to performance rates and than equally treated in a progressive scoring formula, e.g. a power law. The power law solution for the result score S reads[1] :

$$S = A \cdot \left[\frac{(P - P_L)}{(P_H - P_L)} \right]^C \qquad (9)$$

Here A is a scaling constant, P is performance (reciprocal time rather than time), P_L is a lower limit threshold performance (i.e. reciprocal time below which no score is assigned), P_H is a high level performance reference for calibration, and C is a power coefficient that is larger than 1 for obtaining slightly progressive scoring curves. Using empirical data of the world's top 100 decathletes the parameters A, C, P_H and P_L could be determined. Rather than 30 parameters in the existing IAAF power law model of equation (8) (viz. A, P_L and C 10 times each), the alternative model according to equation (9) requires only 22 (A, C, and 10 times P_L and P_H). It was established that the alternative model offers a balanced, transparent, uniform approach, uses less coefficients, and eliminates bias, arbitrariness and unfairness. The overrating of sprinting-based events in decathlon was eliminated. With the proposed tables decathletes would be able to collect their scores to the same extent from any of the 10 disciplines.

Redefining IAAF scoring tables

With the normalised multi-event scoring model (NMSM) we are able to devise a set of scoring tables that are free of the existing anomalies and iniquities. The results are basically given by equation (9), provided that we have available the two calibration levels P_L and P_H and the parameters A and C. These parameters can than be used in equation (9), which yields the result score S for any performance P at any distance s.

-The high-level performances calibration P_H

The choice of these parameters is not critical, if only they correspond with high performances. Since it is a basic requirement of the scoring model that world records in different distances receive the same (high) result score, we will use the world record data for each distance as the high-level reference points.

-The low-level performances calibration P_L

These parameters refer to the threshold performances below which no score is assigned. The procedure we follow here is to conform to the IAAF tables and use the official threshold values B. For practical reasons we used the performances of result score S=1, because performances of result score S=0 are not listed in the IAAF tables. Table 3 lists the threshold performances on various distances. For comparison also $1/P_H$ is listed.

Table 3: Performance thresholds and calibration values for various distances

Distance (m)	Threshold IAAF B (score=1) (s)	$1/P_H$ (s)	$1/(B*P_H)$	$1/P_L$ ($=0.523/P_H$) (s)
100	16.41	9.58	0.584	18.31
200	35.54	19.19	0.540	36.68
400	80.92	43.18	0.534	82.53
800	181.69	101.01	0.556	193.07
1500	380.04	206.00	0.542	393.75
3000	828.92	440.07	0.531	841.15
5000	1421.02	757.35	0.533	1447.60
10000	3106.31	1577.53	0.508	3015.29
21098	7449.00	3503.00	0.470	6695.63
42196	17123.00	7439.00	0.434	14218.90
			Average= 0.523	

The fourth column $1/(B \cdot P_H)$ lists the ratio of threshold performance and top performance. The ratios in this column reveal substantial irregularities indicating arbitrariness in the official IAAF threshold values B. It is hard to understand why this ratio is not a constant so that each distance uses the same relative scale. For better coherence, it is proposed here to use the average ratio (which is 0.523) as the fixed ratio for P_L/P_H. The resulting thresholds $1/P_L$ replacing the IAAF thresholds B are listed in the column at the right. One might still wonder why such threshold should be exactly 0.523 or any other figure.

-Power C

The power C is responsible for progression of scores with distance. It should be larger than 1 to make sure that equal improvements of performance receive higher scores in the high performance range[1]. The underlying idea is that, for instance, improving your 5000m performance from 28 minutes to 26 minutes is less impressive than improving it from 14 minutes to 13 minutes. The IAAF tables use different values for C at different distances. We will use a uniform value of C=1.832, being the average of six values used for different distances by the IAAF[18].

-Scaling factor A

The scaling factor A represents the score we want to assign to the high level performance references P_H, which we have chosen to be the same as the current world records. So A (more or less) corresponds with the maximum score an athlete can get. As a scaling factor its value is a bit arbitrary, it could be 1000 to make a neat figure. But we may also want to link the new score system with the existing one. Therefore we have chosen that A equals the average of the scores that world records receive in the existing regime. The data used are the same as represented in figure 1. The average result score of world records is 1311 points (we only used the distances displayed in the figure). So A is set to 1311.

Having all parameters defined, we are now ready to evaluate the impact of the new scoring model. In Table 4 we have listed an excerpt of the tables.

Table 4. Performance times (m:s) required for different scores and distances according to the new model (NMSM) and the current model (IAAF)

Distance	100 m		800 m		5,000 m		42,196 m	
Score	New	Current	New	Current	New	Current	New	Current
1400	9.42	9.52	1:39.28	1:37.65	12:24.39	12:10.09	2:01:52	1:57:18
1300	9.60	9.77	1:41.23	1:40.79	12:39.01	12:35.92	2:04:15	2:03:35
1200	9.80	10.04	1:43.33	1:44.05	12:54.77	13:02.75	2:06:50	2:09:23
1100	10.02	10.32	1:45.61	1:47.46	13:11.84	13:30.73	2:09:38	2:16:55
1000	10.25	10.61	1:48.09	1:51.02	13:30.45	14:00.02	2:12:41	2:24:02
900	10.51	10.92	1:50.82	1:54.76	13:50.89	14:30.81	2:16:01	2:31:32
800	10.80	11.24	1:53.84	1:58.72	14:13.53	15:03.36	2:19:44	2:39:27
700	11.12	11.59	1:57.22	2:02.94	14:38.88	15:38.02	2:23:53	2:47:53
600	11.48	11.96	2:01.06	2:07.47	15:07.66	16:15.26	2:28:35	2:56:56
500	11.90	12.36	2:05.49	2:12.39	15:40.91	16:55:75	2:34:02	3:06:47
400	12.40	12.81	2:10.74	2:17.84	16:20.25	17:40.54	2:40:28	3:17:41
300	13.01	13.32	2:17.18	2:24.02	17:08.51	18:31.36	2:48:22	3:30:03
200	13.80	13.92	2:25.54	2:31.36	18:11.23	19:31.68	2:58:39	3:44:43
100	14.96	14.70	2:37.78	2:40.92	19:42.97	20:50.27	3:13:40	4:03:50
1	17.99	16.41	3:09.64	3:01:69	23:41.85	23:41.02	3:52:46	4:45:23

From the calculations it follows that over almost the whole range the new scoring system differs significantly from the existing one. In most cases, the athlete has to run faster to obtain the same score the existing model would give. Differences may be up to 25% in the mid-range of performances, for instance, receiving 700 points at 5000m requires being one full minute faster in the new model, which is a difference of 7%. Receiving 700 points in the marathon requires 18%, 1000 points still requires doing more than 3% better. Because world records are used for calibration the differences go down in the high-end range. For instance, at 5000m world record level the difference between the two scoring tables is only three seconds. Although this is only 0.4%, it still makes a big absolute difference: up to 20m. It demonstrates that the impact of (unjust) scoring tables is substantial over the whole range.

Toward Personalised Scoring Tables

The NMSM approach offers a normalised, uniform scoring model for a variety of distances. A disadvantage of the approach is that low-level and high level calibration values must be determined for every single distance. One might consider linking the NMSM with the personal prediction model (PPM) in order to arrive at a single analytical expression for scoring with distance as the independent variable. The basic idea would be to assume a fictitious super athlete that holds all world records. Then equation (7) of PPM could be used to calculate the high-level calibration performances for any distance s* (note that performance =1/t). We would need only two world records for calibration so that α and β can be determined. Since P_L/P_H=constant and while maintaining C and A, equation (9) would then provide a result score S for any distance s*.

Unfortunately, world records are not owned by one single individual but by diverse individuals, each of them specialised in a discipline and displaying the characteristics for excelling in that discipline. Although the overall fit in Figure 2 was encouraging, it isn't good enough for accurately describing the world records. Errors add up to over 7%, which is unacceptable. It is concluded that the PPM as a personal predictor does not hold for combining performances of different individuals.

As a replacement for equation (7), one might use a power law relationship between distance s and exertion time t. KATZ & KATZ[9] explained that a power relationship holds for world records. However, also in this case best-fit solutions fall short, displaying errors up to 9%, which is even worse than the PPM. So with these equations it is not possible to derive an analytical scoring formula using distance as the independent variable.

Nevertheless, at the individual level, the PPM and the NMSM still can be used to produce personalised scoring tables. Such scoring tables assign scores to individual performances for any distance: the personal score (not to be confused with the official IAAF result score) expresses the quality of the performance of the athlete, measured against other performances of the same athlete. Table 5 presents an example: it lists a ranking of Kenenisa Bekele's best performances[16].

Table 5: Kenenisa Bekele's personal performance rankings using PPM and NMSM

Rank	Date	Distance	Time	New score S	1000S/ 1375	IAAF score	IAAF rank
1	28-9-2007	1500	3:32,35	1375	1000	1212	19
2	25-7-2006	1500	3:33,08	1357	987	1202	20
3	31-5-2004	5000	12:37,35	1345	978	1295	2
4	1-7-2005	5000	12:40,18	1326	964	1284	4
5	7-8-2007	3000	7:25,79	1311	954	1266	6
5	26-8-2005	10000	26:17,53	1311	954	1295	1
7	8-6-2008	10000	26:20,31	1302	947	1291	3
8	8-6-2004	10000	26:25,97	1284	934	1282	5
9	17-7-2009	3000	7:28,64	1279	930	1248	11
10	25-8-2006	5000	12:48,09	1274	927	1254	7
11	28-7-2007	5000	12:49,53	1265	920	1249	8
12	29-8-2008	5000	12:50,18	1261	917	1246	12
13	24-8-2001	3000	7:30,67	1257	914	1235	16
14	28-8-2009	5000	12:52,32	1247	907	1238	15
14	27-6-2003	5000	12:52,26	1247	907	1239	14
16	31-8-2008	3000	7:31,94	1243	904	1234	17
17	19-8-2005	3000	7:32,59	1236	899	1223	18
18	14-9-2007	10000	26:46,19	1223	889	1249	9
19	17-8-2009	10000	26:46,31	1222	889	1249	10
20	24-8-2003	10000	26:49,57	1212	882	1243	13
21	17-9-2006	3000	7:36,25	1198	871	1201	21
22	13-9-2003	3000	7:36,98	1190	866	1195	22
23	26-5-2007	3218	8:13,51	1175	855	1189	23
24	27-6-2004	3000	7:41,31	1147	834	1169	24

The table shows the ranking of 24 of Bekele's race performances. Note that such ranking can be made for any athlete. Personal calibration - determining α and β in equation (7) - was done with Bekele's 3000m and 10,000m personal best times (cf. Table 1). Table 5 shows the calculated score S, the same score expressed as a ratio normalised to 1000 points, the IAAF result score[3], and the associated ranking based on IAAF score. Differences are striking. From the perspective of Bekele, neither the 5,000m nor the 10,000m world records are at the top, but instead are two 1500m races. Although the 1500m times are not anywhere near the word record, the performances are of exceptional level taking into account the fact that Bekele is a long-distance runner specialised at 5000 and 10,000m. In fact, this is exactly what the PPM approach takes into account.

Conclusion

This paper explained the personal predictor model (PPM) and the normalised multi-event scoring model (NMSM) and presented a variety of empirical evidence for their validity.

The PPM produces valid and reliable predictions of personal performance limits. The accuracy is well below 1% and thereby greatly outperforms existing models that typically achieve accuracies of 3% or higher. Besides its unchallenged accuracy, the model has some additional advantages. Importantly, the model is transparent, since it is based on theoretical principles rather than arbitrariness and negotiation. Furthermore it is self-contained, easy to use and affordable, because it does not require any physiological or biomechanical tests to be carried out: it just uses two personal bests for individual self–calibration. Since the model compensates for start-up delays it is valid across a wide range of events, including sprinting, middle- and long-distance running. Finally, it can be demonstrated that the model displays a universal validity covering any velocity and distance related sports event, including running, speed skating and swimming[11].

The NMSM was developed to address anomalies of the IAAF scoring tables. The NMSM demonstrates greater fairness, consistency and transparency. Over the whole midrange of performances the new scoring system differs significantly from the existing one, requiring higher performances for obtaining the same score as the current scoring method. Differences are up to 25%, but even small relative differences at world top level translate to appreciable absolute differences. It was demonstrated that the impact of (unjust) scoring tables is substantial over the whole range. This substantiates the need for a new model. Combining the PPM and NMSM for deriving a generic scoring formula that would cover any distance and any performance was not feasible because poor fits of the world record performances to either the PPM or existing power law models. Yet, a personalised scoring table was presented, converting individual performances at various distances into a unified personal score.

Please send al correspondence to:
Prof. dr. Wim Westera
CELSTEC - Centre for Learning Sciences and Technologies
Open University of the Netherlands
Wim.westera@ou.nl

REFERENCES

1. WESTERA, W. (2006). Decathlon, towards a balanced and sustainable performance assessment method. New Studies in Athletics 21(1): 39-51.

2. TRKAL, V. (2003). The development of combined events scoring tables and implications for the training of decathletes. New Studies in Athletics 18(4): 7-17.

3. SPIRIEV, B. (2008). IAAF Scoring Tables of Athletics, 2008 revised edition, retrieved December 18, 2010 from http://www.iaaf.org/mm/Document/Competitions/TechnicalArea/ScoringOutdoor2008_742.pdf.

4. COSTILL, D.L.; THOMASON, H. & ROBERTS, E. (1973). Fractional utilization of the aerobic capacity during distance running. Medicine & Science in Sports & Exercise 5: 248–252.

5. DANIELS, J. & DANIELS, N. (1993). Running economy in elite male and elite female runners. Medicine & Science in Sports & Exercise 24: 483–489.

6. BUNDLE, M.W.; HOYT, R.W. & WEYAND, P.G. (2003). High-speed running performance: a new approach to assessment and prediction. Journal of Applied Physiology 95: 1955-1962.

7. DI PRAMPERO, P.E.; FUSI, S.; SEPULCRI, L.; MORIN, J.B.; BELLI, A. & ANTONUTTO, G. (2005). Sprint running: a new energetic approach. Journal of Experimental Biology

208: 2809-2816.

8. HARDER, D. (2001). Sports Comparisons: You can compare apples to oranges. Castro Valley, CA, USA: Education Plus 1: 7-13.

9. KATZ L. & KATZ J.S. (1999). Power laws and athletic performance. Journal of Sports Sciences; 17: 467-476.

10. STANCKIEWICZ, J. (2005). On track for a 2-hour marathon, South African Journal of Science 101: 397-398.

11. WESTERA, W. (2011). Where Bolt and Bekele meet: the analytical basis of running performance estimates. International Journal of Sports Science and Engineering, 4, 3: 139-152, retrieved September 12, 2011 from http://www.worldacademicunion.com/journal/SSCI/SSCIvol04no03paper02.pdf.

12. WARD-SMITH, A.J. (1985). A mathematical theory of running, based on the first law of thermodynamics, and its application to the performance of world class athletes. Journal of Biomechanics 18: 337-349.

13. LEE, J. (2008). Usain Bolt 100m 10 meter Splits and Speed Endurance, Speedendurance.com 2008; August 22, retrieved December 28, 2010 from http://speedendurance.com/2008/08/22/usain-bolt-100m-10-meter-splits-and-speed-endurance/.

14. LEE, J. (2009). Usain Bolt 10 meter splits, Fastest Top Speed 2008 vs 2009, Speedendurance.com 2009; August 19, retrieved December 28, 2010 from http://speedendurance.com/2009/08/19/usain-bolt-10-meter-splits-fastest-top-speed-2008-vs-2009/.

15. LEE, J. (2009). Usain Bolt 200 meter splits, Speed Reserve and Speed Endurance, Speedendurance.com 2009; August 21, retrieved December 28, 2010 from http://speedendurance.com/2009/08/21/usain-bolt-200-meter-splits-speed-reserve-and-speed-endurance/.

16. IAAF. Top lists, International Association of Athletics Federations, retrieved December 28, 2010 from http://www.iaaf.org/statistics/recbycat/index.html.

17. WESTERA, W. Personal record predictor, preliminary version, JAVA applet retrieved December 28, 2010 from http://www.open.ou.nl/wim/calculate.

18. ATLETIEKUNIE (2004). Formules en Constanten. Commissie wedstrijdreglement KNAU, 9th edition 2004, retrieved December 28, 2010 from http://www.atletiekunie.nl/upload/File/Dutch_Athletes/Formules%20en%20constanten%20(20-11-03).doc