Original Article

Prediction of marathon performance time on the basis of training indices

GIOVANNI TANDA

DIMSET, University of Genova, Italy

ABSTRACT

Tanda G. Prediction of marathon performance time on the basis of training indices. *J. Hum. Sport Exerc.* Vol. 6, No. 3, pp. 511-520, 2011. The purpose of this study was to examine the relationship between marathon performance time (MPT) and some training factors recorded for a given number of weeks prior to a race. Twenty-two runners, age 28-54 years, participated as subjects in this investigation. They kept daily exercise records during their marathon training for an overall number of 46 races, whose marathon time ranged from 167 to 216 min. Among the several parameters investigated, MPT was found to be affected mainly by the mean distance run per week K, during the training period under observation, and by the mean training pace P. These two training parameters have been combined by a mathematical approach to give a correlation for the prediction of the mean marathon pace P_m (easily related to MPT), based on an 8-week training period, as follows: P_m (sec/km) = 17.1 + 140.0 *exp*[-0.0053 K(km/week)] + 0.55 P (sec/km). The above correlation is able to estimate the MPT with a SEE of about 4 min. **Key words**: EXERCISE, ENDURANCE, TRAINING.

Corresponding author. DIMSET, University of Genova, via Montallegro 1, 16145 Genova, Italy
E-mail: giovanni.tanda@unige.it

Submitted for publication February 2011
Accepted for publication September 2011
JOURNAL OF HUMAN SPORT & EXERCISE ISSN 1988-5202

doi:10.4100/jhse.2011.63.05

INTRODUCTION

Efforts to correlate human endurance performance to physiological, physical and training factors have been ongoing for about a century. For endurance sports like marathon running, the maximal oxygen uptake, the lactate threshold and the energy cost of running appear to play key roles in endurance performance (see, for instance, Sjödin & Svedenhag, 1985; Joyner & Coyle, 2008; Faude et al., 2009). Slovic (1977), Hagan et al. (1981, 1987) and Bale et al. (1985) were among the first to point out that some training indices (such as the training pace or the mean distance run per day) are also highly predictive of MPT. Florence & Weir (1997) established a correlation between critical velocity CV (a velocity determined during tests involving a series of fatiguing runs on a treadmill) and MPT; their data showed that marathon time correlated more highly with CV than either peak oxygen consumption or ventilatory threshold. A study conducted by Roecker et al. (1998), based on an incremental treadmill test, revealed that the individual anaerobic threshold had the highest predictive value for long distance competitions among 16 other parameters. Billat et al. (2001) showed that the discriminating factors for international top-class marathoners, when compared with runners at a slightly lower level, are the maximal oxygen consumption $\dot{V}O_{2max}$ for males, and the velocity on a 1000m run (after 10 km run at marathon velocity) for females. In a further investigation, Billat et al. (2003) stated that the velocity at the $\dot{V}O_{2max}$ is the main factor predicting the 10-km performance in elite male and female Kenyan runners; they also investigated the type of training and concluded that a high-intensity training contributes to a higher $\dot{V}O_{2max}$ in men. More recently, Legaz et al. (2006) correlated MPT for a top-level homogeneous group of males and females to some physiological measurements. Their model for the male group used as independent variables the lactate value at 10km/h, left ventricular telediastolic diameter (LVD) and lactate value at 22 km/h, whereas for the female group the independent variables were the subscapular skinfold, serum ferritin and sum of six skinfolds.

Most of the above studies have been based on measurements of such variables as maximal oxygen consumption, heart rates, lactic acid during submaximal and maximal treadmill running, performed on high-level distance runners and involving expensive laboratory facilities. However, a method of predicting performance based on training indices may be, for the increasing mass of non-competitive runners, an attractive and inexpensive alternative to metabolic testing (typically reserved to elite athletes) in events like the marathon.

The aim of this study is to correlate marathon performance time (MPT) only with training characteristics. Correlations for MPT mainly or only based on training variables are already documented in the literature (Slovic 1977; Hagan et al. 1981, 1987; Bale et al. 1985); moreover, a direct relationship between training and physiological factors is likely to exist (Billat et al., 2002, 2003) and this could justify the assumption of a correlation completely based on the training data. To comply with this purpose, a regression analysis has been performed in order to identify those training factors, recorded for a given number of weeks prior to a race, that contribute significantly to the prediction of final race time.

MATERIAL AND METHODS

Research participants

This study collected training data of twenty-two runners (twenty-one males, one female), age 28-54 years for a five-year period, over which they run 46 marathons, with MPT ranging from 167 min to 216 min (2h 47min to 3h 36min). All runners gave informed consent to participate in this investigation. Table 1 gives the anthropometric characteristics of the subjects as mean ± SD (standard deviation), minimum and maximum.

The subjects provided their daily exercise records during their pre-marathon training (typically a three-months period), including the distance and time run for each workout. To ensure the required level of reliability of the study, the marathon races considered had the same level of difficulty (predominantly on flat terrain), similar weather conditions and were run by the athletes at a regular pace throughout the race (with a difference between first half and second half times less than 4 min) and at the highest intensity in line with their training level.

Table 1. *Anthropometric characteristics of the subjects.*

	mean	SD	min	max
Height (cm)	173	6.4	168*	185
Weight (kg)	65.2	6.5	57.5*	79
BMI (kg/m^2)	21.7	1.3	19.2	24.7
Age (years)	42.8	6.7	28	54

**male sample group (for the sole female runner: height:155 cm, weight 53.5 kg)*

Research design and training documentation

The purpose of the study was to develop a regression equation to predict the marathon performance time. Unfortunately, it was not possible to record physiological measurements for the athletes' sample; thus the study was limited to training variables. The training data were accumulated over a twelve-week period before the race. Since the workouts performed during the week immediately preceding the race are typically of lesser importance, analysis was developed neglecting data collected in the final week.

The variables considered for the MPT prediction were: the number of previous marathons run N_m, the number of training days per week N_d, the mean distance run for each workout K_{day}, the maximum workout distance per week K_{max}, the mean workout distance per week K, the mean training pace P. In order to provide reliable data, subjects were used to train on track or in external flat environment with a GPS device in order to correctly estimate their training distances. Moreover, each athlete included, in his/her training programme, one or more long-distance workouts (typically a 30-35 km long steady run 3-4 weeks prior the race).

Parameters based on the distance run in a given time lap (day, week) typically include the distance run also in the warm-up and recovery; the same consideration applies to the calculation of the training pace, given by the ratio between the time, employed to run the workout distance (including the warm-up and recovery), and the distance itself.

No tests for a given distance (i.e. faster time for a 5km or 10km run) were considered in this study for the following reasons: the difficulty in scheduling a running test under standard conditions (i.e., similar weather and terrain, same time allocation prior to the marathon, etc.) for athletes living and training in different cities, and the consideration that numerous subjects of the sample did not include a fast 5km or 10km workout in their programme of marathon training.

Training and performance characteristics of the research participants are summarised in Table 2. Even though the analysis has been conducted by averaging the collected training data over 11, 10 and 8 consecutive weeks prior to the race, the time-averaged parameters reported in Table 2 refer to the 8-week period since no significant alterations of data were recorded by changing (from 8 to 11 weeks) the period of observation. The reference, 8-week, period started 9 weeks before the race and finished seven days prior to the race.

Table 2. *Performance and training of the subjects.*

	mean	SD	min	max
Finishing time MPT (min)	191	12	167	216
Race pace P_m (sec/km)	271.8	17.7	237.4	306.8
Previous marathons N_m	7.6	6.3	0	22
#No. of training days per week N_d	4.0	1.0	2.7	6.8
#Mean distance for each workout K_{day} (km/day)	15.6	1.4	11.4	19.1
Maximum workout distance per week K_{max} (km/week)	87	15.2	59	132
#Mean workout distance per week K (km/week)	65.9	15.9	40.4	110.7
#Mean training pace P (sec/km)	284.6	18.1	253.3	330.6

 # averaged over an 8-week training period prior to the race

Statistical analysis

The identification of the parameters significantly affecting the MPT was performed by using a standard software for the regression analysis of data according to different shapes (linear, polynomial, exponential, power law, etc.) of the mathematical correlating function. The standard error of estimate SEE and the correlation coefficient r were considered to evaluate the accuracy of a given regression curve. Too large values of SEE or values of r relatively far from unity were considered as indices of poor quality of the correlation.

RESULTS

Firstly, the effects of the selected training characteristics on the recorded marathon pace P_m (sec/km) were investigated. It is useful to keep in mind that marathon performance time MPT, expressed in min, is related to P_m, expressed in sec/km, by the equation: $MPT = 42.195 P_m/60$.

The relationship of marathon pace P_m to considered training indices is reported in Table 3, where the shape of the best regression line, the correlation coefficient (r) and the standard error of estimate (SEE) are indicated for each independent variable. Results showed that the mean distance for each workout K_{day} (km/day) and the maximum distance per week K_{max} (km/week) are not effective predictors for MPT. Similarly, neither the number of workout days per week N_d, nor the number of previous marathons N_m were found to be predictive of the marathon pace. Conversely, the mean distance per week K and the mean training pace were found to be strongly correlated with MPT. The regression lines giving the marathon pace P_m versus K and P are shown in Figures 1 and 2, respectively.

Table 3. *Relationship of marathon pace to selected training indices.*

	shape of best regression line	coefficient of determination **r**	**SEE**(sec/km)
Previous marathons N_m	linear	0.11	17.9
#No. of training days per week N_d	linear	0.72	12.8
#Mean distance for each workout K_{day} (km/day)	linear	0.36	17.2
Maximum workout distance per week K_{max} (km/week)	linear	0.71	12.6
#Mean workout distance per week K (km/week)	exponential decay	0.81	10.7
#Mean training pace P (sec/km)	linear	0.85	9.3

 # averaged over an 8-week training period prior to the race.

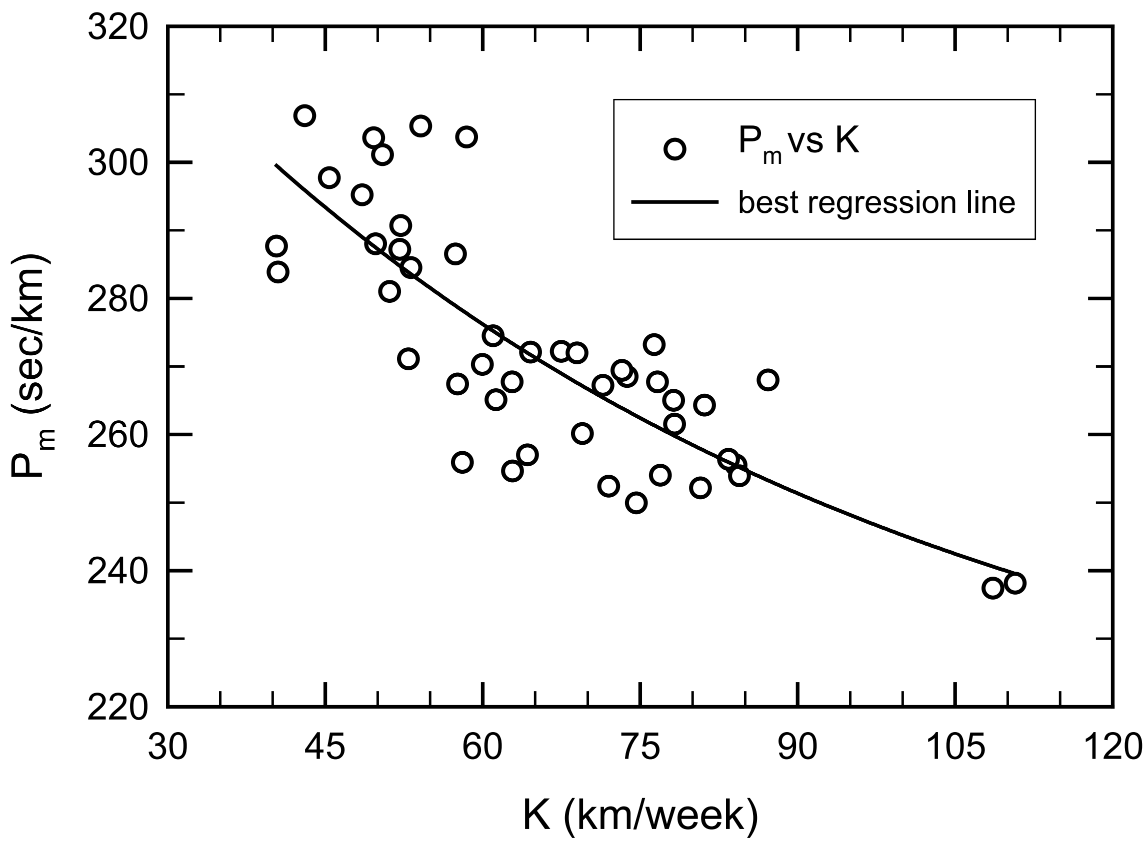

Figure 1. *Relationship between P_m (marathon pace) and K (mean distance run per week).*

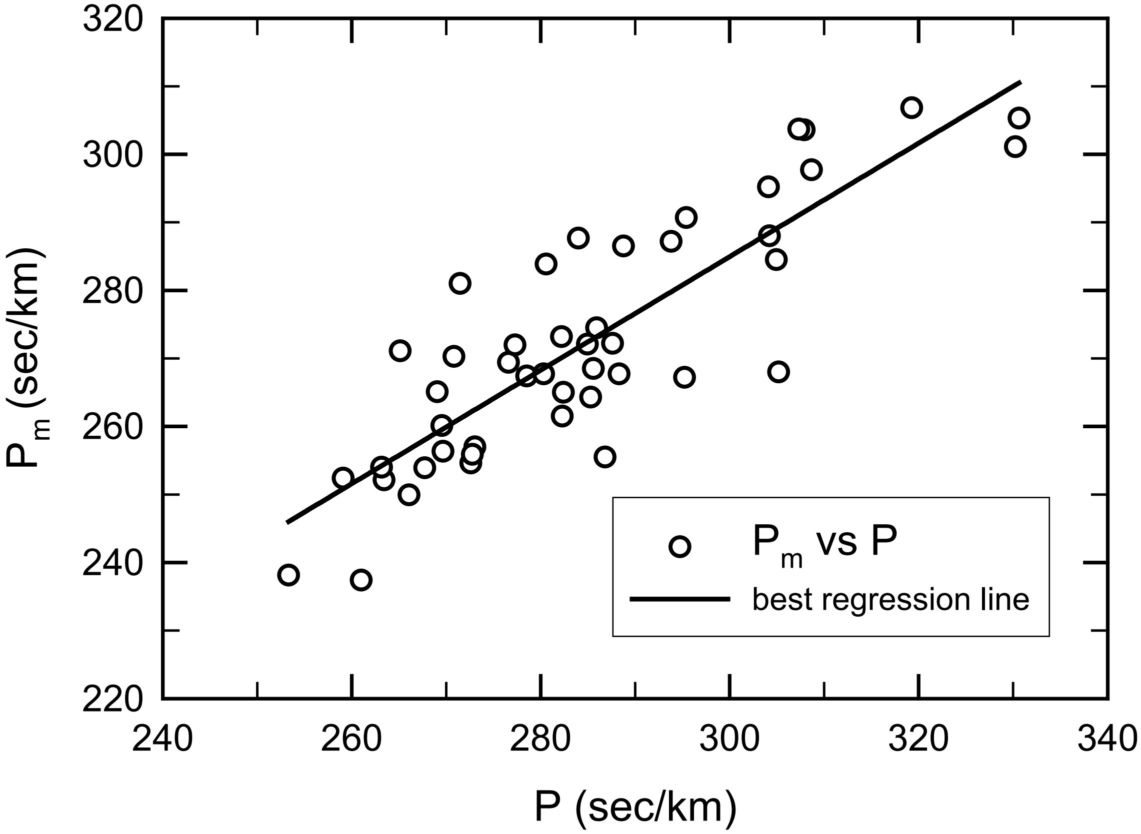

Figure 2. *Relationship between P_m (marathon pace) and P (mean training pace).*

Attention was then turned to the development, through a multiple non-linear regression analysis, of the relationship giving the marathon pace P_m as a function of the best predictive factors, K and P.

According to the complete set of training data, the marathon pace can be predicted by the following equation:

$$P_m \text{ (sec/km)} = 17.1 + 140.0\ exp[-0.0053\ K\text{(km/week)}] + 0.55\ P \text{ (sec/km)}$$

where *exp* denotes the exponential function. The standard error of estimate SEE of the regression equation is 5.77 sec/km, which corresponds to a SEE of only 4 min in the MPT prediction. Despite the reduced range of variation of BMI for the sample group, it was found that, when the sample of runners with BMI<23 kg/m^2 is considered (17 athletes out of 22, with 37 marathons run out of 46) the SEE of the regression equation drops to only 5.10 sec/km (i.e. 3 min 35 sec for MPT).

The predicted values of P_m (sec/km) are plotted in Figure 3 as functions of the recorded values. The solid line represents the line of perfect agreement. The correlation predicts 69% of MPT values within 4 min and 50% within 2 min. These figures increase to 73% and 54%, respectively, when only subjects with BMI<23 kg/m^2 are considered.

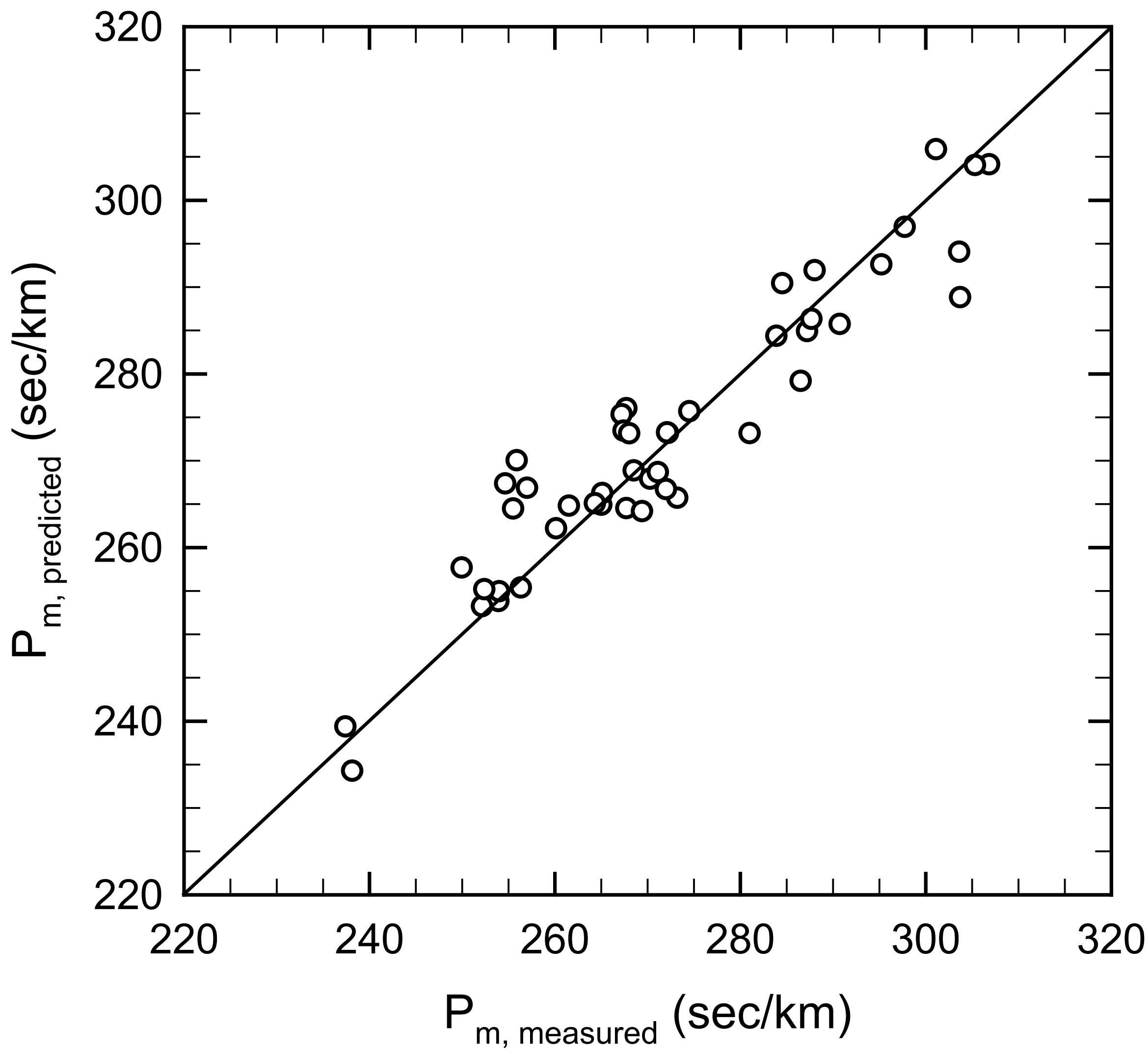

Figure 3. *Mean marathon pace: predicted values against measured values. Solid line: perfect agreement.*

DISCUSSION

The present study converged to a correlation featured by a standard error of estimate typically lower than that encountered in similar studies reported in the literature (Hagan et al., 1981, 1987; Florence & Weir, 1997), despite the heterogeneous physical and training characteristics of the subjects, and developed utilizing only two independent variables, one related to the training volume and the other to the training intensity. The reasons for such a successful result of the regression analysis are probably: (i) the good quality of the sample training data (each runner selected for this study performed his training seriously and with great motivation) and of the achieved race performances (each marathon run by the athlete group has been typically featured by a constant pace throughout the race), (ii) the correct identification of main training indices, (iii) and the choice of a non-linear regression analysis. The main limitation of the regression equation for MPT is its validity range, currently from 167 min (2h, 47min) to 216 min (3h, 36min), which is the range of race finishing time of the sample group. However, it is worth noting that, for example, more than 10 thousand Italian people ran, in 2010, at least one marathon with a finishing time included in the above MPT range; therefore the results of this study are potentially addressed to a relatively large cohort of marathoners.

The duration of time over which the training measurements were accumulated ranged, in similar studies reported in the literature (Slovic, 1977; Foster 1975, 1983; Grant et al. 1984; Hagan et al. 1981, 1987; Billat et al. 2001), from 8 to 12 weeks. Hagan et al. (1987) showed that regardless of the period of observation (8, 9, or 12 weeks), their equation variables remained unchanged. The present investigation seems to support this consideration; the processing of present data was not affected by the period (from 8 to 11 weeks) over which training data were averaged. The 8-week period was then selected for the regression equation since it involves a lower amount of data to record and process.

The number of subjects tested for this study was 22, lower than that considered by Hagan et al. (1981, 1987), but comparable to the one used by Billat et al. (2001, 2003) and higher than the sample number of Florence & Weir (1997). It should be emphasized that some of the 22 athletes provided training data for more than a single marathon race (up to four) run typically at six months/one year of distance from each other; therefore the number of overall training data (and relevant marathon races run by the 22 athletes) was 46, satisfactory to develop an accurate regression analysis. The sample group is mostly formed by male athletes (21 subjects out of 22), However, the training data (recorded for 3 marathons run) provided by the sole woman of the sample group are accurately correlated by the present analysis. Of course, the accumulation of further training data for a larger sample of women and of athletes, in general, including elite marathoners, could lead in the near future to an even more accurate correlation within a larger field of validity.

Of the several training indices investigated, the analysis suggests that some are poorly correlated to MPT, like the mean distance for each workout, the maximum distance per week, the number of previous marathons, and the number of workout days per week. This result is in agreement with the majority of similar studies reported in the literature. Conversely, the mean distance run per week and the mean training pace emerged as the most important training characteristics affecting the marathon performance. Hagan et al. (1981) included (among the others) these two parameters in their correlation for MPT prediction based on the analysis of male runners training data; in a following study (Hagan et al., 1987), conducted for female distance runners, MPT was correlated to the mean training pace and to the daily (rather than the weekly) workout. It is interesting to note that the maximal aerobic power was not related to MPT for the female sample. A similar study (Bale et al., 1985) on female marathon runners indicated the number of sessions per week, the ectomorphy, the distance run per week and the number of years of training as the best predictors for MPT. Conversely, studies conducted for first-time marathoners (Grant et al., 1984) and for middle-aged and older runners (Takeshima et al., 1995) found only a low or moderate level of correlation between the weekly running distance and the endurance running performance.

A more recent investigation on elite runners (Billat et al., 2003) revealed that the velocity at the high-intensity training is the main factor predicting the 10km-run performance; this finding is in agreement with the results of this study, showing that the mean training pace P, which is related to intensity of the training period, is the best predictor for MPT (lowest SEE and highest r). However, incorporating also the mean weekly distance K is found, in the present analysis, to enhance the predictive accuracy of the developed correlations for MPT; this is probably ascribed to the characteristics of the sample, not including elite runners but amateurs only.

The effect of the athletes' physical characteristics on MPT was not investigated in this study, apart from the body mass index (BMI) of each runner, recorded and included in the analysis. Hagan et al. (1981, 1987) showed that BMI was a better predictor of MPT for experienced female runners than for male runners. Campbell (1985) found that the most important predictors for running speed in half-marathon were the distance per week and BMI. No statistical significance was reached for the relationship of BMI with finish time for both men and women ultramarathon runners was recorded by Hoffman et al. (2010), whereas BMI was shown to account for 10-11% of the variation in finish time in a previous study (Hoffman et al., 2008), where a larger sample size was considered. The present regression analysis indicated that a BMI lower than 23 kg/m^2 (that is quite usual for an experienced long-distance runner) contributes to increasing the accuracy of the regression equations for MPT. Since BMI values of the sample group varied in a too narrow range, no attempt to relate it to MPT was make; however it is argued from results that BMI may (negatively) affect MPT only beyond some critical values (say 23 kg/m^2).

CONCLUSIONS

The results of this investigation, conducted by processing the marathon training data of 22 athletes, showed a high correlation between marathon performance time and a couple of training indices that can be easily calculated by processing training data accumulated over a given period of observation (8 weeks) prior the marathon race. These indices (the mean weekly distance and the mean training pace) condensed the volume and intensity of training respectively, regardless of the type of training programme followed by each athlete. Currently, the validity range of the correlation for MPT is from 167 min to 216 min. The standard error of estimate of the regression equation is 4 min but it goes down to about 3min 30sec when only runners with BMI lower than 23 kg/m^2 are considered.

ACKNOWLEDGEMENTS

The author is grateful to the athletes for their availability to be subjects for this investigation, and to Prof. S. Lo Presti, for the technical assistance and suggestions.

DEDICATION

This paper is dedicated to the memory of Dr. Maurizio Magagnini, trainer and dear friend, who passed away on July 7, 2010.

REFERENCES

1. BALE P, ROWELL S, COLLEY E. Anthropometric and training characteristics of female marathon runners as determinants of distance running performance. *J Sports Sci.* 1985; 3:115-126. doi:10.1080/02640418508729741 [Abstract] [Back to text]

2. BILLAT VL, DEMARLE A, SLAWINSKI J, PAIVA M, KORALSZTEIN J-P. Physical and training characteristics of top-class marathon runners. *Med Sci Sports Ex.* 2001; 33:2089-2097. [Full Text] [Back to text]
3. BILLAT VL, DEMARLE A, PAIVA M, KORALSZTEIN J-P. Effect of training on the physiological factors of performance in elite marathon runners (males and females). *Int J Sports Med.* 2002; 23:336-341. doi:10.1055/s-2002-33265 [Back to text]
4. BILLAT VL, LEPRETRE P-M, HEUGAS A-M, LAURENCE M-H, SALIM D, KORALSZTEIN J-P. Training and bioenergetic characteristics in elite male and female Kenyan runners. *Med Sci Sports Ex.* 2003; 35:297-304. [Abstract] [Back to text]
5. CAMPBELL MJ. Predicting running speed from a simple questionnaire. *Br J Sports Med.* 1985; 19:142-144. [Full Text] [Back to text]
6. FAUDE O, KINDERMANN W, MEYER T. Lactate threshold concepts: How valid are they? *Sports Med.* 2009; 39:469-480. doi:10.2165/00007256-200939060-00003 [Back to text]
7. FLORENCE S-I, WEIR JP. Relationship of critical velocity to marathon running performance. *Eur J Appl Physiol.* 1997; 75:274-278. [Full Text] [Back to text]
8. FOSTER C, DANIELS J. Running by the numbers. *Runner's World.* 1975; 10:14-17. [Back to text]
9. FOSTER C. VO_2 max and training indices as determinants of competitive running performance. *J Sports Sci.* 1983; 1:13-22. [Abstract] [Back to text]
10. GRANT AJY, SHARP RH, AITCHISON TC. First time marathoners and distance training. *Brit J Sports Med.* 1984; 18:241-243. [Full Text] [Back to text]
11. HAGAN RD, SMITH MG, GETTMAN LR. Marathon performance in relation to maximal aerobic power and training indices. *Med Sci Sports Ex.* 1981; 13:185-189. [Abstract] [Back to text]
12. HAGAN RD, UPTON SJ, DUNCAN JJ, GETTMAN LR. Marathon performance in relation to maximal aerobic power and training indices in female distance runners. *Br J Sports Med.* 1987; 21:3-7. [Full Text] [Back to text]
13. HOFFMAN MD. Anthropometric characteristics of ultramarathoners. *Int J Sports Med.* 2008; 29:808-811. doi:10.1055/s-2008-1038434 [Back to text]
14. HOFFMAN MD, LEBUS DK, GANONG AC, CASAZZA GA, VAN LOAN M. Body composition of 161-km ultramarathoners. *Int J Sports Med.* 2010; 31:106-109. doi:10.1055/s-0029-1241863 [Back to text]
15. JOYNER MJ, COYLE EF. Endurance exercise performance: the physiology of champions. *J Physiol.* 2008; 586.1:35-44. doi:10.1113/jphysiol.2007.143834 [Back to text]
16. LEGAZ A, MUNGUÍA D, SERVETO JR. Physiological measures associated with marathon running performance in high-level male and female homogeneous groups. *Int J Sports Med.* 2006; 27:289–295. doi:10.1055/s-2005-865628 [Back to text]
17. ROECKER K, SCHOTTE O, NIESS AM, HORSTMANN T, DICKHUTH H-H. Predicting competition performance in long-distance running by means of a treadmill test. *Med Sci Sports Ex.* 1998; 30:1552-1557. [Full Text] [Back to text]
18. SJÖDIN B, SVEDENHAG J. Applied physiology of marathon running. *Sports Med.* 1985; 2:83-99. [Abstract] [Back to text]
19. SLOVIC P. Empirical study of training and performance in the marathon. *Res Quart.* 1977; 48:769-777. [Full Text] [Back to text]
20. TAKESHIMA N, TANAKA K. Prediction of endurance running performance for middle-aged and older runners. *Brit J Sports Med.* 1995; 29:20-23. [Full Text] [Back to text]

