OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS 4754(B)

Applications of Advanced Mathematics (C4)

Paper B: Comprehension

INSERT

Monday **12 JUNE 2006** Afternoon Up to 1 hour

INSTRUCTIONS TO CANDIDATES

- This insert contains the text for use with the questions.

This insert consists of 11 printed pages and 1 blank page.

[Turn over

Modelling athletics records

Introduction

In the 1900 Olympic Games, shortly before world records were first kept, the record time for the marathon was almost exactly 3 hours. One hundred years later, in 2000, the world record stood at 2 hours 5 minutes and 42 seconds; it had been set during the previous year by Khalid Kannouchi of Morocco. At the time of writing this article, the world marathon record for men is 2 hours 4 minutes and 55 seconds, set by Paul Tergat of Kenya.

When will the marathon record fall below 2 hours?

It is clearly not possible to predict exactly when any world record will be broken, or when a particular time, distance or height will be achieved. It depends, among other things, on which athletes are on form at any time. However, it is possible to look at overall trends and so to make judgements about when new records are likely to be set.

Prediction inevitably involves extrapolating beyond existing data, and so into the unknown. If this is to be more than guesswork, it must be based on a suitable mathematical model.

It is reasonable to hope that a general model can be found, one that can be adapted to many athletics events. Such a model will take the form of a formula involving several parameters; these will take different values for different events. The parameter values will take account of the obvious distinction that, whereas records for track events (like the marathon and the mile) decrease with time, those for field events (like the long jump and the javelin) increase.

This article looks at possible formulae for such a model.

The linear model

The simplest type of model is linear and this is well illustrated by the men's mile. The graph in Fig. 1 shows the world record for the mile plotted against the year from 1915 to 2005. Details of these records are given in Appendix A.

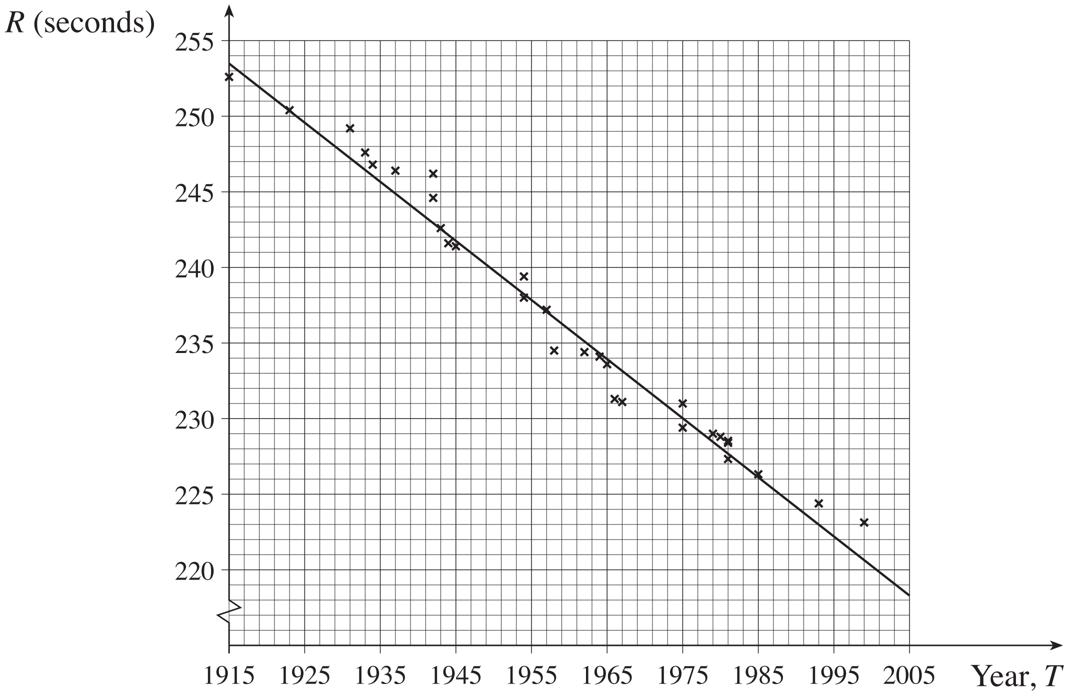

Fig. 1

A line of best fit has been drawn on Fig. 1. Its equation is 25

$$R = 259.6 - 0.391(T - 1900)$$

where

- R is the record time in seconds
- T is the calendar year.

(This equation was calculated using a standard statistical technique.) 30

The straight line clearly provides quite a good model for the record time between the years 1915 and 2005. However, it will not continue to do so for ever. For a sufficiently large value of T, the value of R will become negative, which is clearly impossible.

While the record time becoming negative shows that the linear model needs to be refined or replaced, there are also positive times that are quite unrealistic. Over the years, training methods have improved, as have running techniques and conditions, and no doubt this process will continue. However, there is a level of performance that will never be achieved by a human; for example, it seems highly unlikely that any human will ever run a mile in a time of 2 minutes. So, somewhere between the present record and 2 minutes, there is a certain time that will never quite be achieved, a lower bound for the record time. A good model needs to incorporate this idea.

The simple exponential model

In Fig. 2, such a lower bound is represented by the horizontal asymptote, $R = L$. You would expect the record time to approach the asymptote as a curve and this is also illustrated in Fig. 2.

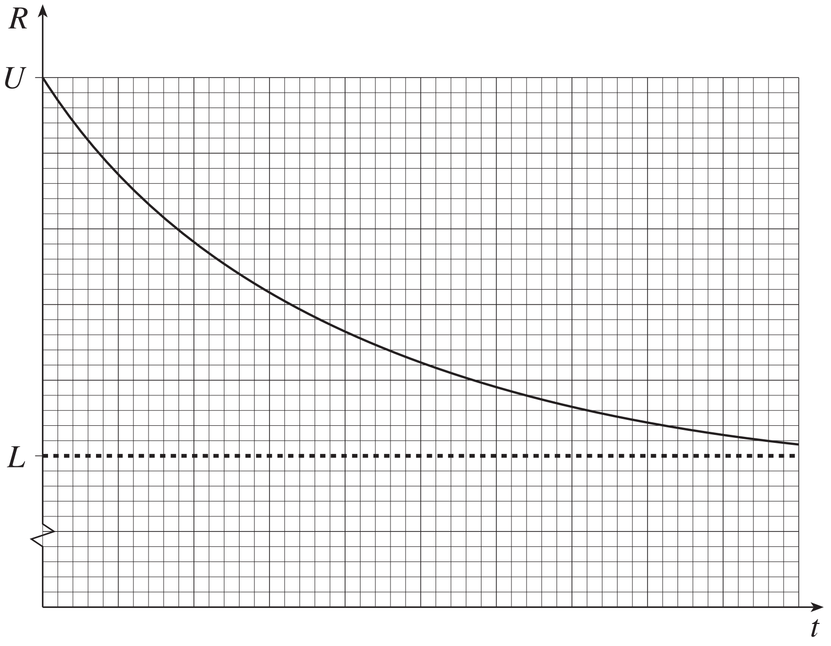

Fig. 2

The data for the mile (illustrated in Fig. 1) could correspond to a part of such a graph before it had flattened out. However, if you look again at Fig. 1, you may think that a gentle curve is more appropriate than the straight line, particularly for the more recent records.

A possible equation for such a model for the record time has the form

$$R = L + (U - L)e^{-kt}$$

where

- U (Upper) is the initial record
- L (Lower) is the value at the asymptote, as illustrated in Fig. 2
- t is the time that has elapsed since records began
- k is a positive constant.

Notice the distinction in this article between t and T. The symbol T has already been used in line 26.

- T denotes the calendar year (so for the present year $T = 2006$).

In this model the record time obeys a simple exponential law and so in this article it is referred to as the *simple exponential model*.

Applying the simple exponential model to the men's marathon

The graph in Fig. 3 shows the record times for the men's marathon from 1908, when world records began, to 2005. In this case the record time, R, is measured in minutes. (Details of the performances are given in Appendix B.) A "curve of best fit", in this case drawn by eye, has been superimposed.

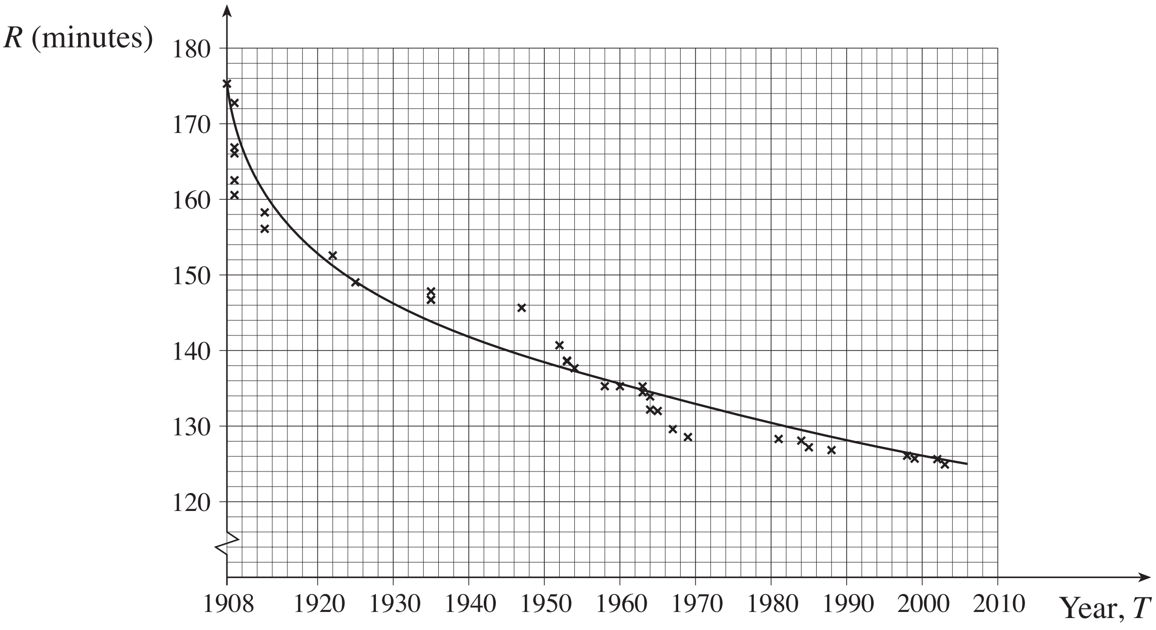

Fig. 3

There are 3 parameters in the equation for the simple exponential model, L, U and k. They will take different values for different athletics events. The values of the 3 parameters can be determined from the coordinates of 3 points on the curve, each point giving rise to one equation. It is easiest to solve the equations if the 3 points chosen correspond to the initial time (i.e. $t = 0$) and two equally spaced subsequent values of t.

For 1908, 1955 and 2002, the curve goes through the points corresponding to

$$\begin{array}{ll} t=0 & R=175 \\ t=47 & R=137 \\ t=94 & R=125.5. \end{array}$$

The first equation is

$$175 = L + (U - L)\,\mathrm{e}^0, \qquad \text{(Equation } \mathbf{1}\text{)}$$

and this can be simplified to give $U = 175$.

The other two equations are as follows.

$$137 = L + (175 - L)\mathrm{e}^{-47k} \qquad \text{(Equation } \mathbf{2}\text{)}$$

$$125.5 = L + (175 - L)\mathrm{e}^{-94k} \qquad \text{(Equation } \mathbf{3}\text{)}$$

Equation **2** can be rewritten as

$$\mathrm{e}^{-47k} = \frac{137 - L}{175 - L}$$

and Equation **3** as

$$\mathrm{e}^{-94k} = \frac{125.5 - L}{175 - L}.$$

Since $\mathrm{e}^{-94k} = (\mathrm{e}^{-47k})^2$, it follows that

$$\frac{125.5 - L}{175 - L} = \left(\frac{137 - L}{175 - L}\right)^2.$$

This equation can be solved to give $L = 120.5$ (correct to 1 decimal place).

Substituting for L, in either Equation **2** or **3**, gives an equation in k. The solution is $k = 0.0254$ and so this model for the marathon record is

$$R = 120.5 + 54.5\mathrm{e}^{-0.0254t}$$

and this can alternatively be written as

$$R = 120.5 + 54.5\mathrm{e}^{-0.0254(T-1908)}$$

where T is the calendar year.

According to this model the 2-hour marathon will never be run.

When Roger Bannister ran the first 4-minute mile in 1954, there was speculation that this represented just about the limit of the capability of the human frame. Now 3 minutes 40 seconds would seem a possibility. So the prediction of the simple exponential model that the 2-hour marathon will never be run feels distinctly unrealistic. This raises questions about the suitability of the model being used.

The general exponential model

A more sophisticated exponential model is given by the equation

$$R = L + (U - L)\mathrm{e}^{-kt^{\alpha}}.$$

In this model, the time t is raised to the power α, where $\alpha > 0$. So this model has 4 parameters, L, U, k and α. In this article it is referred to as the *general exponential model*. The previous simple exponential model was the special case when $\alpha = 1$.

The advantages of this new model are shown by comparing Figs. 4 and 5, for both of which L and U have been given the values 110 and 150 respectively. In Fig. 4, the simple exponential model is illustrated for different values of the parameter k in the family of curves given by the equation

$$R = 110 + (150 - 110)\mathrm{e}^{-kt}.$$

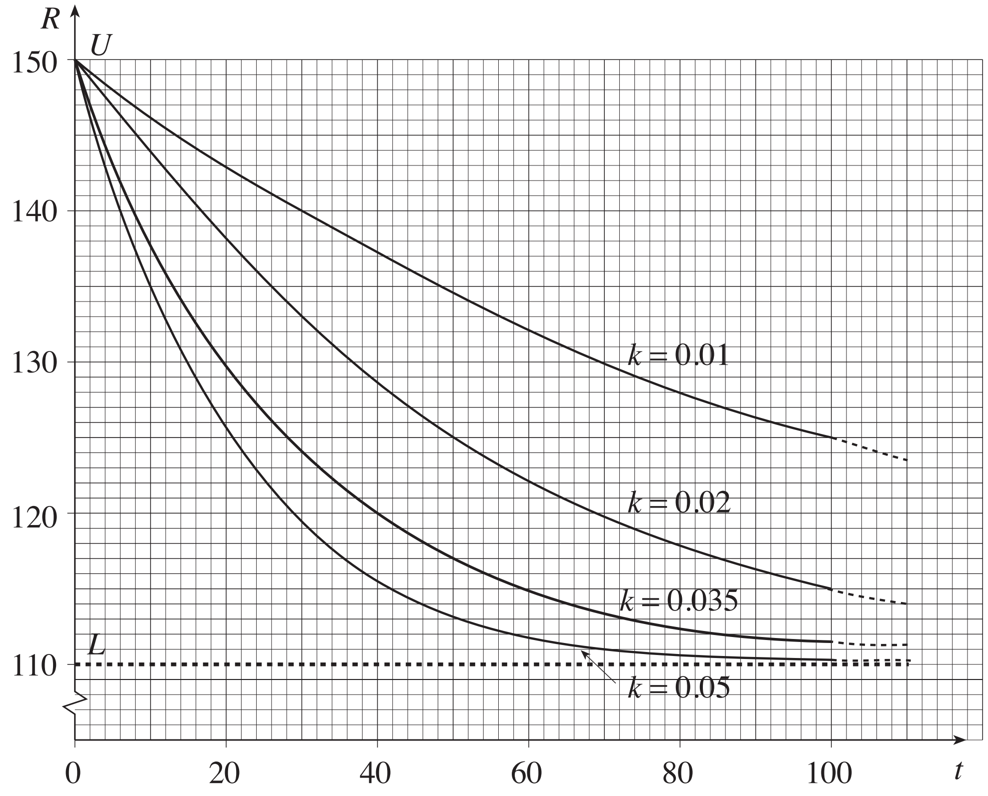

Fig. 4

All the curves approach the asymptote at $R = 110$ in essentially the same manner. (Each curve can be obtained from any other by applying a horizontal stretch.)

It is easy to see that, in Fig. 4, the value of k determines both the initial gradient and the subsequent path of the curve.

For this particular family $$\frac{dR}{dt} = -40ke^{-kt}.$$

When $t = 0$ $$\frac{dR}{dt} = -40k,$$ 115

so that $$k = -\frac{1}{40} \times \text{the initial gradient}.$$

Thus the simple exponential model is completely defined by the starting value, U, the lower bound, L, and the initial gradient.

By contrast the general exponential model allows variation in the shape of the curves. In Fig. 5, there are two curves. Curve A is an example of the simple exponential model and curve B of the general exponential model. Their equations are given by 120

$$\text{A:} \qquad R = 110 + (150 - 110)e^{-0.03t}$$

$$\text{B:} \qquad R = 110 + (150 - 110)e^{-0.134t^{0.5}}.$$

Both of these curves pass through the same initial point (0, 150) and have the same horizontal asymptote $R = 110$. The horizontal asymptote is not shown in Fig. 5; instead the graph has been restricted to smaller values of t to show the differences between the two models more clearly. 125

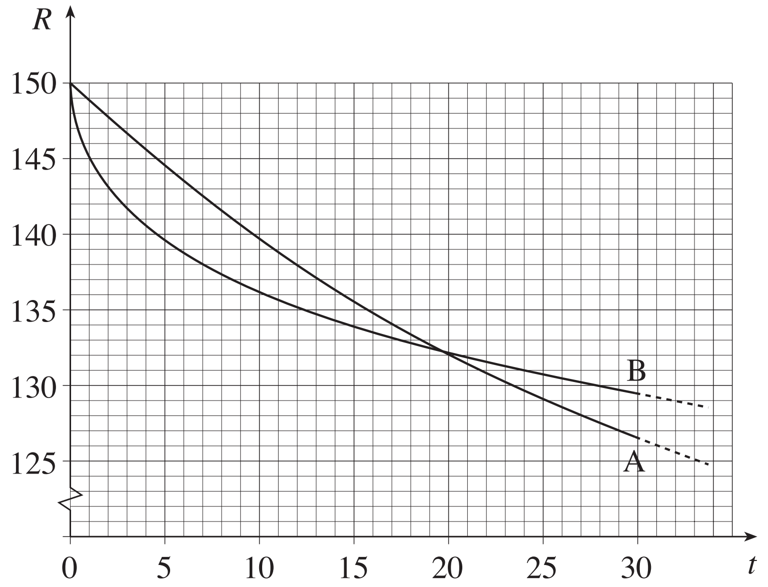

Fig. 5

With the given values for the parameters, according to the general exponential model (curve B) the record times initially fall more quickly than in the simple exponential model (curve A). At about $t = 20$, the two models give the same record time but after that the general exponential model is always further away from the asymptote. 130

The two curves in Fig. 5 are only examples. The values of the parameters were chosen to illustrate the different characteristics of the two models, and have no significance beyond that.

[Turn over

Experience shows that when a new event is introduced, for example the women's marathon in the early 1970s, records tend to decrease very rapidly for the first few years (or, of course, to increase for new field events). It is possible to allow for this in the general exponential model without getting close to the bound unrealistically soon. This is not the case with the simple exponential model.

So the general exponential model, with its 4 parameters, has the flexibility to provide a reasonable model for records.

With this model, it is also possible to address the concern expressed in lines 96 to 97 about the prediction for the men's marathon obtained from the simple exponential model.

For example, the general exponential curve through $(0, 175)$, $(47, 137)$ and $(94, 125.5)$ with $k = 0.0467$ and $\alpha = 0.797$

- has its asymptote at 115 minutes rather than 120.5 minutes
- gives $R = 120$ when $t = 146$; this corresponds to the 2-hour marathon in the year 2054 rather than never.

In Table 6 a number of possible applications of the general exponential model to the men's marathon are listed. They all pass through the same 3 points as before, but have different values for the lower bound, L.

Lower bound, L, for marathon record	**Model**	**Calendar year, T, for 2-hour marathon**
115	$R = 115 + (175 - 115)\mathrm{e}^{-0.0467t^{0.797}}$	2054
110	$R = 110 + (175 - 110)\mathrm{e}^{-0.0579t^{0.706}}$	2045
105	$R = 105 + (175 - 105)\mathrm{e}^{-0.0641t^{0.650}}$	2041
100	$R = 100 + (175 - 100)\mathrm{e}^{-0.0673t^{0.611}}$	2039
95	$R = 95 + (175 - 95)\mathrm{e}^{-0.0686t^{0.582}}$	2037

Table 6

These results show a relationship between the lower bound, L, and the predicted date for the 2-hour marathon. The smaller the lower bound, the sooner we can expect a 2-hour marathon. This finding coincides with common sense.

All the predictions in Table 6 for the 2-hour marathon seem rather cautious. If it happens much sooner, that may well be evidence that an even more sophisticated model is needed. It could even have happened between the time of writing this article and today, when you are reading it.

Finding the parameter values

Table 6 illustrates the versatility of the general exponential model. However, it does not address the question of how you determine the values of the various parameters.

One possible method would be to take a 4th point on the curve, giving 4 equations in the 4 unknowns, U, L, k and α. Apart from the fact that the resulting equations would be very difficult to solve, there is another point to be considered.

The curve in Fig. 3 was drawn by eye and so is not a curve of best fit in a mathematical sense. That would require a statistical technique like that used for the straight line in Fig. 1. This technique is built into curve-fitting software that will find the parameters in the equations of many curves of best fit. Such standard software would work for the simple exponential model but cannot handle the more complicated equation for the general exponential model. So special programming would be needed.

However, the success of such a statistical method depends on the quality of the data. While all the points in Fig. 3 correspond to the records given in Appendix B, and so are correct, they nonetheless all represent unusual occurrences; that is the nature of world records. Some experts believe that, for any athletics event, a better picture is obtained by taking, say, the best five performances each year and constructing a model based on them, rather than relying solely on rare and exceptional occurrences.

Attempts have been made to use such an approach to link sudden large improvements in athletics records to the possible use of performance-enhancing drugs, but so far this work has been inconclusive.

Appendix A Mile records from 1915 (men)

Year	Athlete	Nationality	Time
1915	Taber	USA	4 m 12.6 s
1923	Nurmi	Finland	4 m 10.4 s
1931	Ladoumegue	France	4 m 9.2 s
1933	Lovelock	New Zealand	4 m 7.6 s
1934	Cunningham	USA	4 m 6.8 s
1937	Wooderson	UK	4 m 6.4 s
1942	Hagg	Sweden	4 m 6.2 s
1942	Hagg	Sweden	4 m 4.6 s
1943	Andersson	Sweden	4 m 2.6 s
1944	Andersson	Sweden	4 m 1.6 s
1945	Hagg	Sweden	4 m 1.4 s
1954	Bannister	UK	3 m 59.4 s
1954	Landy	Australia	3 m 58.0 s
1957	Ibbotson	UK	3 m 57.2 s
1958	Elliot	Australia	3 m 54.5 s
1962	Snell	New Zealand	3 m 54.4 s
1964	Snell	New Zealand	3 m 54.1 s
1965	Jazy	France	3 m 53.6 s
1966	Ryun	USA	3 m 51.3 s
1967	Ryun	USA	3 m 51.1 s
1975	Bayi	Tanzania	3 m 51.0 s
1975	Walker	New Zealand	3 m 49.4 s
1979	Coe	UK	3 m 49.0 s
1980	Ovett	UK	3 m 48.8 s
1981	Coe	UK	3 m 48.53 s
1981	Ovett	UK	3 m 48.40 s
1981	Coe	UK	3 m 47.33 s
1985	Cram	UK	3 m 46.32 s
1993	Morceli	Algeria	3 m 44.39 s
1999	El Guerrouj	Morocco	3 m 43.13 s

Appendix B Marathon records (men)

Year	Athlete	Nationality	Time
1908	Hayes	USA	2 h 55 m 18 s
1909	Fowler	USA	2 h 52 m 45 s
1909	Clark	USA	2 h 46 m 52 s
1909	Raines	USA	2 h 46 m 04 s
1909	Barrett	UK	2 h 42 m 31 s
1909	Johansson	Sweden	2 h 40 m 34 s
1913	Green	UK	2 h 38 m 16 s
1913	Ahlgren	Sweden	2 h 36 m 06 s
1922	Kolehmainen	Finland	2 h 32 m 35 s
1925	Michelsen	USA	2 h 29 m 01 s
1935	Suzuki	Japan	2 h 27 m 49 s
1935	Ikenana	Japan	2 h 26 m 44 s
1935	Son	Korea	2 h 26 m 42 s
1947	Suh	Korea	2 h 25 m 39 s
1952	Peters	UK	2 h 20 m 42 s
1953	Peters	UK	2 h 18 m 40 s
1953	Peters	UK	2 h 18 m 34 s
1954	Peters	UK	2 h 17 m 39 s
1958	Popov	USSR	2 h 15 m 17 s
1960	Bikila	Ethiopia	2 h 15 m 16 s
1963	Teresawa	Japan	2 h 15 m 15 s
1963	Edelen	USA	2 h 14 m 28 s
1964	Heatley	UK	2 h 13 m 55 s
1964	Bikila	Ethiopia	2 h 12 m 11 s
1965	Shigematsu	Japan	2 h 12 m 00 s
1967	Clayton	Australia	2 h 09 m 36 s
1969	Clayton	Australia	2 h 08 m 33 s
1981	de Castella	Australia	2 h 08 m 18 s
1984	Jones	UK	2 h 08 m 05 s
1985	Lopes	Portugal	2 h 07 m 12 s
1988	Dinsamo	Ethiopia	2 h 06 m 50 s
1998	de Costa	Brazil	2 h 06 m 05 s
1999	Khannouchi	Morocco	2 h 05 m 42 s
2002	Khannouchi	USA	2 h 05 m 38 s
2003	Tergat	Kenya	2 h 04 m 55 s