Forecasting Records by Maximum Likelihood

RICHARD L. SMITH*

A maximum likelihood method of fitting a model to a series of records is proposed, using ideas from the analysis of censored data to construct a likelihood function based on observed records. This method is tried out by fitting several models to series of athletics records for mile and marathon races. A form of residual analysis is proposed for testing the models. Forecasting consequences are also considered. In the case of mile records, a steady linear improvement since 1931 is found. The marathon data are harder to interpret, with a steady improvement until 1965 with only slight improvement in world records since then. In both cases, the normal distribution appears at least as good as extreme-value distributions for the distribution of annual best performances. Short-term forecasts appear satisfactory, but serious reservations are expressed about using regression-type methods to predict long-term performance limits.

KEY WORDS: Athletics records; Censored data; Generalized extreme-value distribution; Gumbel distribution; Inference for stochastic processes.

1. INTRODUCTION

Athletics records are a subject of evergreen interest, in their own right and as a testing ground for physiological theories (e.g., Lietzke 1954; Lloyd 1966). One aspect is the relation among records achieved at different distances, a subject that has been approached both theoretically (e.g., Keller 1974; Ward-Smith 1985) and empirically (Riegel 1981). This article is concerned with a different aspect—the improvement of records over time. Specifically, I propose methods of model-fitting and forecasting when the available data consist just of a series of records, as illustrated by Figures 1 and 2. Tryfos and Blackmore (1985) derived a method of generalized least-squares analysis, but under the very restrictive assumption that the records are derived from an underlying independent and identically distributed (iid) sequence of random variables. The method proposed here, based on the maximum likelihood principle, is more general, and I consider its application to several athletics series. I propose a probability plotting technique that is useful for assessing a model's fit. I then consider the problem of forecasting, both short-term (Sec. 5) and long-term (Sec. 6), in the latter case disputing the conclusions of Chatterjee and Chatterjee (1982) and Morton (1983, 1984) on ultimate-limit estimation.

Most of the existing mathematical literature on records, and the closely related subject of extremal processes, has been concerned with records in iid sequences. Glick (1978), Galambos (1978, chap. 6), and De Haan (1984) reviewed the area. Recently, attention has been given to records in sequences containing an underlying trend. Yang (1975) considered the effect on record times of an increasing population size. Ballerini and Resnick (1985, 1987) studied records in a sequence consisting of a stationary random sequence superimposed on a linear trend. De Haan and Verkade (1987) considered some cases where (because of either a very long-tailed distribution or a very slow trend) the asymptotic theory is similar to the iid case. On the statistical side, Smith and Miller (1986) proposed a class of models for records from a Bayesian predictive viewpoint. Their models permitted a more general dependence structure than that considered here, but were restrictive in being based on the Gumbel distribution, which does not fit the data particularly well.

2. MODELS

Let Y_n denote the best performance in a particular event in the nth year covered by the data. Depending on the event, this may refer either to a minimum or to a maximum; for definiteness, assume it is a minimum, though we could of course apply an analogous analysis to maxima. Assume Y_n is of the form

$$Y_n = X_n + c_n, \qquad (2.1)$$

where $\{X_n\}$ is iid and $\{c_n\}$ is a nonrandom trend. The records in an N-year period are given by

$$Z_n = \min(Y_1, \ldots, Y_n), \qquad 1 \le n \le N. \qquad (2.2)$$

In most of the analysis assume that the available data consist of $\{Z_n, 1 \le n \le N\}$; that is, the records are available but the underlying sequence $\{Y_n\}$ is not. This is a natural starting point for athletics data, since the sequence $\{Z_n\}$ is easily extracted from a table of records whereas the sequence $\{Y_n\}$ is not. There are also other contexts in which data may be available only in the form of records—for example, the use of historical information in hydrology or certain problems in materials testing (Glick 1978). Note that I am only considering the best performance in each year. Thus, if the record is broken more than once during the year, only the latest (best) record is used in the analysis.

Suppose we have a parametric model in the form $c_n = c_n(\beta)$, $X_n \sim f(x; \theta)$, where f is a continuous density and β and θ are (vector) parameters. The situation may be thought of as a censored-data problem in which the value of Y_n, in a nonrecord year, is censored at the existing value. Then the likelihood function of (β, θ) based on

* Richard L. Smith is Professor of Statistics, Department of Mathematics, University of Surrey, Guildford GU2 5XH, England.

Journal of the American Statistical Association
June 1988, Vol. 83, No. 402, Applications and Case Studies

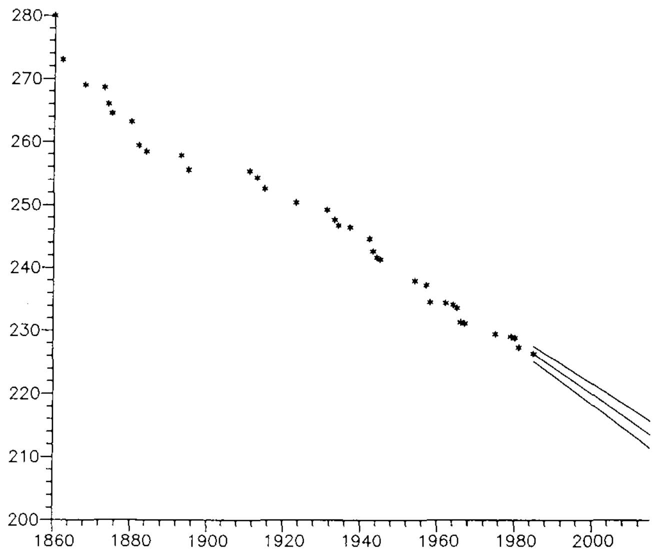

Figure 1. Plot of Mile Records (seconds) Against Year of Achievement for 1860–1985. On the right are predictions of annual best performance (not records) for the next 20 years, with 95% confidence bands computed by the method described in Section 5, based on Model A fitted to 1931–1985 data.

$Z_1, \ldots, Z_N$ is

$$\prod_{n=1}^{N} \{f(Z_n - c_n(\beta); \theta)\}^{\delta_n}\{1 - F(Z_n - c_n(\beta); \theta)\}^{1-\delta_n}, \tag{2.3}$$

where F is the cdf of f and δ_n is 1 if there is a record in year n, 0 otherwise. The maximum likelihood estimates $\hat{\beta}$ and $\hat{\theta}$ are found by numerical maximization of (2.3). As in classical problems, assume that $(\hat{\beta} - \beta, \hat{\theta} - \theta)$ is approximately normal with mean 0 and covariance matrix given by the inverse of the observed information matrix. For the special but important case of a linear trend, this procedure can be given a rigorous justification in terms of asymptotic theory. (The details are in an unpublished report, available from me.)

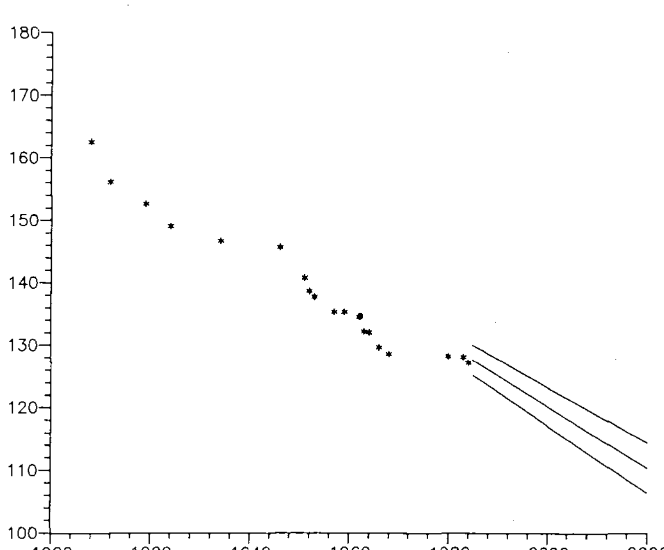

Figure 2. Plot of Marathon Records (minutes) for 1909–1985, With Projections Computed by the Same Method as for Figure 1, Based on Model A for 1909–1985. It is apparent that the improvement has not closely followed a straight line, but none of the alternative models is a noticeable improvement.

Particular distributions of $\{X_n\}$ include the normal distribution; the Type I extreme-value distribution (Gumbel 1958), defined by

$$F(x; \mu, \sigma) = 1 - \exp[-\exp\{(x - \mu)/\sigma\}] \tag{2.4}$$

($x \in \Re$; $\mu \in \Re$, $\sigma > 0$) and referred to here as the *Gumbel distribution*; and the generalized extreme-value (GEV) distribution

$$F(x; \mu, \sigma, k) = 1 - \exp[-\{1 + k(x - \mu)/\sigma\}^{1/k}], \tag{2.5}$$

defined on $\{x : 1 + k(x - \mu)/\sigma > 0\}$ for $\mu \in \Re$, $\sigma > 0$, and $k \in \Re$ [the case $k = 0$ is (2.4)]. This distribution incorporates Gumbel's (1958) Types II and III in a parameterization convenient for estimation (Prescott and Walden 1980, 1983). The rationale for these distributions rests on the definition of Y_n as an annual minimum, though the normal distribution is also a viable choice in practice.

The near linearity of Figures 1 and 2 suggests a linear function

$$c_n(\beta_0, \beta_1) = \beta_0 - n\beta_1, \qquad \beta_1 > 0, \tag{2.6}$$

as an obvious model for the drift term. As alternatives to this, consider a *quadratic-drift model*,

$$c_n(\beta_0, \beta_1, \beta_2) = \beta_0 - n\beta_1 + n^2\beta_2/2, \qquad \beta_1 > 0, \tag{2.7}$$

which I regard as an artificial alternative against which to test (2.6), and an *exponential-decay model*,

$$c_n(\beta_0, \beta_1, \beta_2) = \beta_0 - \beta_1\{1 - (1 - \beta_2)^n\}/\beta_2, \qquad \beta_1 > 0, \; 0 < \beta_2 < 1, \tag{2.8}$$

which is equivalent on reparameterization to models of Chatterjee and Chatterjee (1982) and Morton (1983); this is intended as a more realistic alternative to (2.6), with the feature that the eventual limit $\beta_0 - \beta_1/\beta_2$ exists as $n \to \infty$. Note that (2.8) reduces to (2.6) in the limit $\beta_2 \to 0$.

Thus we have nine candidate models obtained by combining the different distributions and regression curves, as follows:

A: Normal, linear drift
B: Normal, quadratic drift
C: Normal, exponential decay
D: Gumbel, linear drift
E: Gumbel, quadratic drift
F: Gumbel, exponential decay
G: GEV, linear drift
H: GEV, quadratic drift
I: GEV, exponential decay.

In the Normal case, assume a mean of 0 and a variance of σ^2 (unknown), the mean being absorbed into c_n. In the Gumbel and GEV cases, $\mu = 0$ in (2.4) or (2.5).

The models are fit by numerical maximum likelihood. Two algorithms were used: (a) the NAG routine E04CGF, which is a quasi-Newton algorithm with numerically approximated derivatives, and (b) a modified Newton–Raphson algorithm for the Gumbel and GEV cases, using analytic first- and second-order derivatives. The second algorithm is an extension of the method of Prescott and Walden (1983). The basic Newton–Raphson procedure is modified by including a check that the log-likelihood increases at each iteration (if not, a shorter step length is taken) and by switching to the direction of steepest ascent if the Newton step is not defined (Hessian matrix not negative definite) or if it fails to result in an increase of the log-likelihood. Hosking's (1985) algorithm employs the same principles, but only for iid data. In the GEV case, this procedure should converge to a local maximum if one exists, but there is no guarantee that one does. This difficulty was observed by Rockette, Antle, and Klimko (1974) in the closely related case of the three-parameter Weibull distribution; it was analyzed theoretically by Smith (1985). I used a variety of strategies in my algorithms to determine the starting points of the iterations. In difficult cases several different starting points were tried, though experience suggests that the choice of starting point is not critical when the aforementioned modifications of the Newton–Raphson algorithm are employed.

Theoretical justification of maximum likelihood may be given, in the case of a linear trend only, by asymptotic arguments. The structure of the problem, with censoring determined by past values of the series, identifies it as an "inference for stochastic processes." Arguments of Ballerini and Resnick (1985) may be adopted to prove that the process is ergodic; asymptotic normality then follows from the general result of Sweeting (1980). (The details are in an unpublished report, available from me.)

The analysis proceeded by fitting these models to some athletics series. In addition to testing the models against each other, it is possible to fit them to different portions of the same series and thus gain some indication of how the models are changing with time. In Section 3, I summarize some of these procedures' results.

3. ANALYSIS OF RECORDS DATA

The preceding models were fitted to the mile and marathon data of Figures 1 and 2, part of which were tabulated by Smith and Miller (1986). In addition, for the mile-race data, a complete set of annual best performances was available (Ballerini and Resnick 1985), and the results were analyzed for comparison with the results for mile-record data. In the latter case, of course, there is no censoring and the model-fitting is simple regression analysis by maximum likelihood under the three error distributions.

Over short portions of the data we would expect the linear drift (2.6) to be satisfactory; for reasons to be explained shortly, it was found that the normal distribution is the most suitable among the three distributions considered. Therefore, I fitted Model A to different portions of the series; the results are summarized in Tables 1–3. For ease of comparison, the year n was taken relative to a fixed origin, which was arbitrarily fixed as 1930 for the mile data and 1908 for the marathon data. The standard errors in Tables 1–3 are calculated from the observed information matrix.

Inspection of Tables 1 and 2 confirms the visual picture from Figure 1, of a rapid improvement in times over the last 40 years of the nineteenth century, followed by a period of very little improvement to about 1930, followed by fairly steady improvement since then. There is some change in β_1 between the periods 1931–1959 and 1960–1985, though the change is not nearly as great as over the preceding periods and is barely significant. Over the whole series (1860–1985), a linear drift does not fit particularly well. In fact a cubic regression fits this series much better than any of the models considered in Section 2, though, as the main reason for this is the slow improvement from 1900 to 1930, it would not seem reasonable to attach too much significance to this for the purpose of predicting

Table 1. Mile Best Performance: Model A Fitted to Various Portions of the Series

	Years					
Parameter	1860–1985	1931–1985	1860–1894	1895–1930	1931–1959	1960–1985
β_0*	249.6	251.0	235.0	252.9	251.2	246.5
	(.3)	(.6)	(2.6)	(.9)	(1.0)	(1.9)
β_1	.37	.44	.62	.24	.44	.34
	(.01)	(.02)	(.05)	(.04)	(.04)	(.04)
σ	3.3	2.3	2.9	2.7	2.7	1.7
	(.2)	(.2)	(.3)	(.3)	(.3)	(.2)

NOTE: Units are seconds. Standard errors are in parentheses.
* Standardized to a 1930 base.

Table 2. Mile Records: Model A Fitted to Various Portions of the Series

	Years					
Parameter	1860–1985	1931–1985	1860–1894	1895–1930	1931–1959	1960–1985
β_0*	250.3	250.0	237.2	252.9	250.6	244.8
	(.5)	(.6)	(3.8)	(1.3)	(.9)	(1.3)
β_1	.38	.42	.59	.17	.44	.32
	(.01)	(.02)	(.07)	(.05)	(.05)	(.03)
σ	3.4	1.7	3.2	1.7	2.0	1.0
	(.4)	(.3)	(.7)	(.6)	(.5)	(.2)

NOTE: Units are seconds. Standard errors are in parentheses.
* Standardized to a 1930 base.

Table 3. Marathon Records: Model A Fitted to Various Portions of the Series

	Years				
Parameter	1909–1985	1941–1985	1909–1940	1941–1964	1965–1985
β_0*	165.2	164.7	164.5	182.5	140.2
	(1.5)	(2.9)	(2.2)	(5.4)	(3.0)
β_1	.49	.52	.44	.89	.15
	(.03)	(.05)	(.12)	(.11)	(.04)
σ	3.9	3.8	4.0	2.4	1.1
	(.7)	(.7)	(1.4)	(.6)	(.3)

NOTE: Units are minutes. Standard errors are in parentheses.
* Standardized to a 1908 base.

Table 4. *Summary of Fits for the Mile Race: Best Performances of 1931–1985*

	Model								
Parameter	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*
β_0	251.0 (.6)	251.3 (1.0)	251.3 (1.0)	252.8 (.7)	251.7 (1.2)	—*	251.5 (.7)	252.7 (.5)	252.8 (.6)
σ	2.3 (.2)	2.3 (.2)	2.3 (.2)	2.4 (.2)	2.4 (.2)	—*	2.5 (.3)	2.8 (.4)	2.7 (.4)
k	—	—	—	—	—	—*	.47 (.09)	.73 (.17)	.76 (.14)
β_1	.44 (.02)	.47 (.08)	.47 (.08)	.46 (.02)	.36 (.10)	—*	.44 (.02)	.62 (.05)	.62 (.07)
$\beta_2 \times 10^3$	—	6.7 (14)	2.8 (6.0)	—	−1.7 (1.6)	—*	—	3.5 (1.0)	14.4 (6.6)
Negative maximized log-likelihood	124.6	124.5	124.5	131.1	130.6		122.7	120.3	120.7

NOTE: Standard errors are in parentheses.
* No fit.

future performances. Comparison between the analysis based on records and that based on best performances shows generally good agreement, though in some cases the estimated values of σ are significantly higher when based on the best performances. This suggests that there may be some bias in the maximum likelihood estimator of σ in the censored case.

A similar analysis for the marathon data (Table 3) shows a different pattern. In this case I excluded the year 1908 as an obvious outlier (see Fig. 2), but the analysis then indicates steady improvement from 1909 to 1940, followed by much greater improvement from 1941 to 1964 and then very little improvement since 1965, the difference between the latter values of β_1 being very clearly significant. This is somewhat surprising; although the world record has improved only slightly since Derek Clayton's 1969 record, overall international performances have improved considerably during the same period. The reason may have something to do with the much closer attention given to accurate course measurement during the past 15 years; in any case, marathon courses are not comparable in the same way as running tracks. Incidentally, Clayton's 1969 record is widely disputed because of the lack of adequate course measurement on that occasion. His performance has been included in my analysis, though it would make little difference to the conclusions if it were omitted.

For a more detailed analysis, all nine models from Section 2 were fitted. The initial examination of the mile data suggested a linear improvement from 1931 to 1985, so I concentrate on this period in the subsequent analysis. The linear-improvement model may then be tested against the quadratic and exponential alternatives, as a further indication of whether the improvement really is linear. For the marathon data I analyzed the whole series 1909–1985, though in view of the preceding remarks there must be some doubt about whether it is reasonable to fit a single model to the whole series. The results of these fits are given in Tables 4–6.

In no case was the quadratic or exponential model a significant improvement over the linear model, as judged either by the standard error of the new parameter or by the increase in maximized log-likelihood. This is somewhat surprising in view of the preceding analysis and suggests that, in cases where the linear model does not fit well, it

Table 5. *Summary of Fits for Mile Records, 1931–1985*

	Model								
Parameter	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*
β_0	250.0 (.6)	250.8 (.8)	250.8 (.8)	250.7 (.5)	251.0 (.9)	250.9 (.8)	251.0 (.9)	—*	—*
σ	1.8 (.3)	1.6 (.3)	1.6 (.3)	1.3 (.2)	1.3 (.2)	1.3 (.2)	2.4 (.9)	—*	—*
k	—	—	—	—	—	—	.47 (.20)	—*	—*
β_1	.42 (.02)	.52 (.07)	.51 (.07)	.43 (.02)	.46 (.07)	.46 (.07)	.43 (.02)	—*	—*
$\beta_2 \times 10^3$	—	1.6 (1.1)	7.1 (5.2)	—	.54 (1.3)	2.3 (5.4)	—	—*	—*
Negative maximized log-likelihood	58.9	57.9	58.0	60.6	60.5	60.5	58.1		

NOTE: Standard errors are in parentheses.
* No fit.

Table 6. Summary of Fits for Marathon Records, 1909–1985

Parameter	Model A	B	C	D	E	F	G	H	I
β_0	165.2 (1.5)	165.4 (1.9)	165.4 (1.9)	166.5 (1.0)	166.2 (2.1)	—*	—*	—*	—*
σ	3.9 (.7)	3.8 (.7)	3.8 (.7)	2.6 (.5)	2.6 (.5)	—*	—*	—*	—*
k	—	—	—	—	—	—*	—*	—*	—*
β_1	.49 (.03)	.51 (.10)	.51 (.10)	.50 (.03)	.48 (.10)	—*	—*	—*	—*
$\beta_2 \times 10^3$	—	.3 (1.2)	1.1 (4.8)	—	−.2 (1.1)	—*	—*	—*	—*
Negative maximized log-likelihood	73.7	73.6	73.6	75.2	75.2				

NOTE: Standard errors are in parentheses.
* No fit.

may be necessary to look for sharp changes in the parameter values rather than the gradual change of slope implied by the quadratic or exponential models. Comparing the distributions, the normal distribution seems clearly superior to the Gumbel, when judged by the maximized log-likelihoods. The GEV distribution appeared the best of the three when it was fitted at all, though in several cases the algorithm failed to converge. It is known that the log-likelihood has no local maximum when $k \geq 1$ and that the asymptotic properties of maximum likelihood estimation fail to hold when $k > \frac{1}{2}$ (see Smith 1985). In this study, I found $\hat{k} > \frac{1}{2}$ in most cases where a fit was found at all, and I believe that the cases where no fit was found correspond to $k > 1$ (so no local maximum exists). The maximum likelihood method thus appears to fail in this case, so the GEV distribution was not pursued further. It should be emphasized, however, that this is because of a failure of the estimation method and not because of any evidence that the GEV distribution does not fit. Alternative estimation methods include the "maximum product of spacings" method (Cheng and Amin 1983; Ranneby 1984) and Bayesian analysis. Smith and Naylor (1987) gave a Bayesian analysis of the three-parameter Weibull distribution, using the computational method of Naylor and Smith (1982).

4. RESIDUAL PLOTS

As in any regression analysis, much may be learned from residual plots. In this case, I define residuals as $r_n = c_n(\hat{\beta}) - Z_n$ ($1 \leq n \leq N$) so that large residuals correspond to good performances, and I let F denote the (common) estimated df of $\mu + c_n(\beta) - Y_n$ (from the model). I consider two plots—(a) a plot of residuals against time and (b) a plot of residuals against expected values. Figure 3 shows a plot of residuals against time (uncensored observations only) for the mile records under Model A. Large residuals correspond to exceptionally good performances; the two largest residuals are derived from the performances of Herb Elliott (1958) and Jim Ryun (1966), who are widely considered to be among the greatest middle-distance runners.

A plot of residuals against approximate expected values was constructed by adapting a method of Aitkin and Clayton (1980). Define

$$v_{ij} = 1 \quad \text{if } r_j < r_i$$
$$= 0 \quad \text{if } r_j > r_i,\ \delta_j = 0$$
$$= F(r_i)/F(r_j) \quad \text{if } r_j > r_i,\ \delta_j = 1,$$

for $1 \leq i \leq N$, $1 \leq j \leq N$, such that $\delta_i = 0$. Loosely, this represents the probability that the true (uncensored) residual in year j is less than that for year i. Then define

$$u_i = \left(\sum_{j \neq i} v_{ij} + \frac{1}{2}\right) \Big/ N$$

for each i such that $\delta_i = 0$. The residuals r_i are plotted against $F^{-1}(u_i)$ for the uncensored years. Ignoring errors

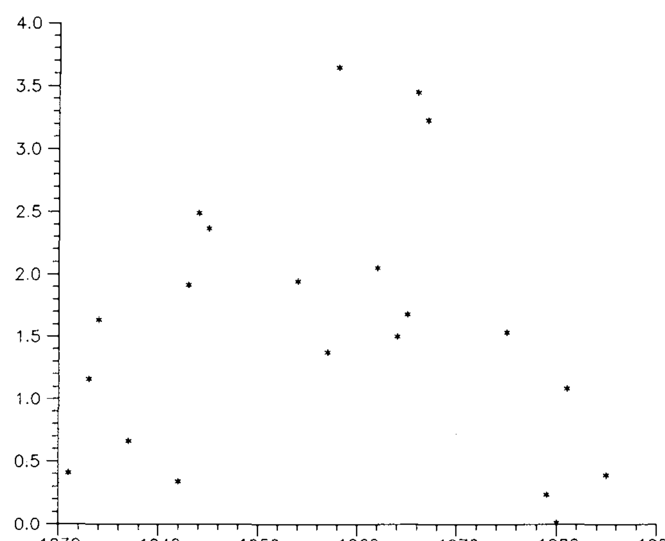

Figure 3. Plot of Residuals (from Model A, 1931–1985) for the Mile-Record Data. Only those residuals corresponding to actual records are plotted; consequently, the residuals are all positive, even though the underlying distribution is assumed normal. The two largest residuals correspond to the performances of Herb Elliott and Jim Ryun; thus there is a case for considering them the overall best runners for the period of the plot.

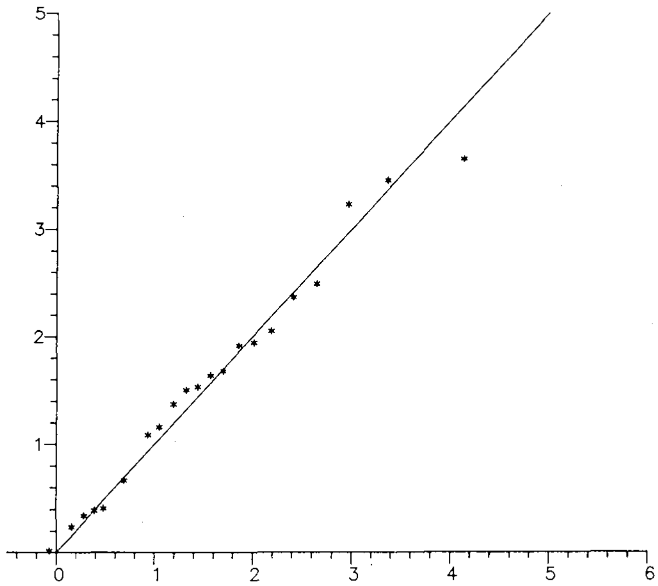

Figure 4. Plot of Residuals (observed vs. expected values) for the Mile Data From Model A (normal distribution). The plot remains close to the diagonal line.

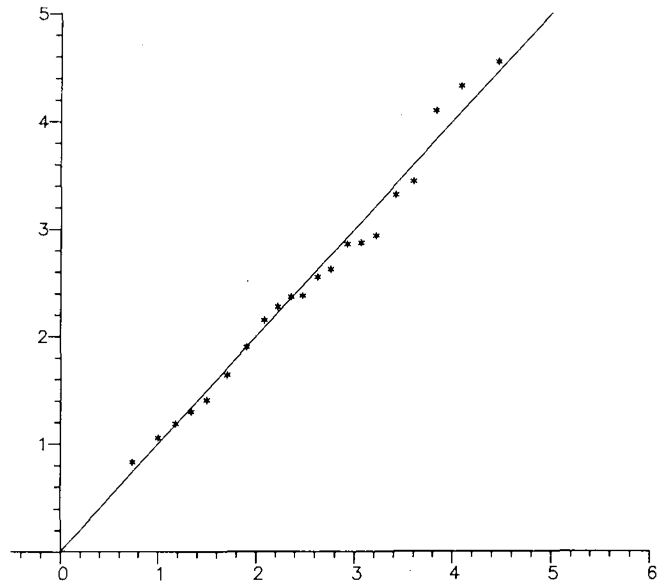

Figure 6. Plot of Residuals (observed vs. expected values) for the Mile Data From Model G (GEV distribution). This appears to be a very good fit, but there are estimation difficulties (see Sec. 3).

of estimation, this should be close to a straight line of unit slope through the origin. This form of probability plot differs from one based on the Kaplan–Meier estimator (Chambers, Cleveland, Kleiner, and Tukey 1983, sec. 6.10) in that the u_i's depend on F. Note that the plot effectively corrects for the censoring. Thus it is not invalidated by the fact that the residuals, as thus defined, are nearly all positive.

This plot is shown in Figures 4–6 for the mile records—Models A, D, and G. In the case of the Gumbel model (Model D), the largest two residuals lie well below their expected values, suggesting that the Gumbel model is a poor fit in the lower tail of the mile-record distribution.

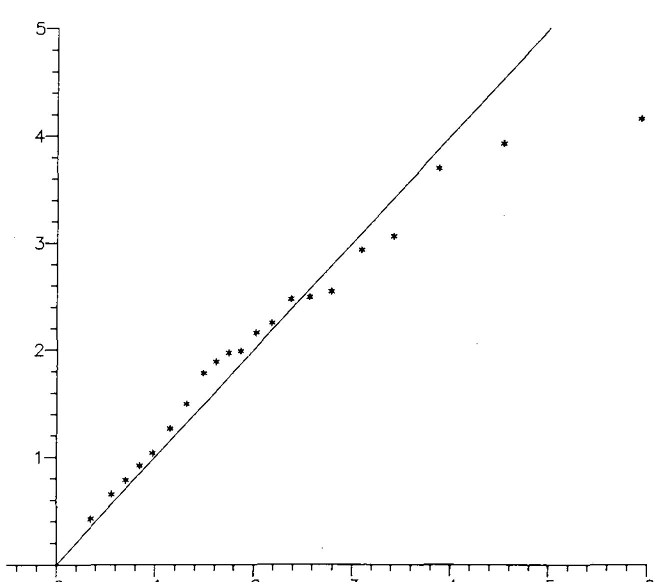

Figure 5. Plot of Residuals (observed vs. expected values) for the Mile Data From Model D (Gumbel distribution). The discrepancy at the right end (equal to the lower tail of the actual distribution) is evident.

The normal and GEV distributions both seem to fit well, however. In view of the difficulties in fitting the GEV distribution, this adds support to the use of the normal distribution in practice. Similar plots for the marathon data showed the same behavior even more strongly.

5. FORECASTING

Suppose we wish to forecast future performances. A simple approach is to estimate some specified quantile of the distribution of the best performance in the required year, substituting the maximum likelihood estimates for the fitted-model parameters. Standard errors may be obtained from the inverse of the observed information matrix by the well-known delta method (Rao 1973, p. 388). As an example, Figure 1 shows the median best performance for the mile for the next 30 years based on Model A fitted to 1931–1985 data. It also includes 95% confidence limits, calculated as estimated ± 2 standard errors. Similar plots are obtained for the other models, though with much larger standard errors (up to 10 seconds) for Models B, C, and E. Figure 2 shows forecasts for the marathon, derived from the same model. Figure 7 shows the median estimate for the mile for all six models. The three models based on linear drift (Models A, D, and F) give virtually identical results, suggesting that the choice of the correct distribution may not be so important for forecasting purposes.

Smith and Miller (1986) took a different approach to forecasting, based on Bayesian predictive distribution. They also discussed the forecasting of records as opposed to annual minima.

6. THE ULTIMATE RECORD

The exponential decay model (2.8) was introduced, in a different parameterization, by Chatterjee and Chatterjee (1982) and applied to predict the ultimate limit of athletic

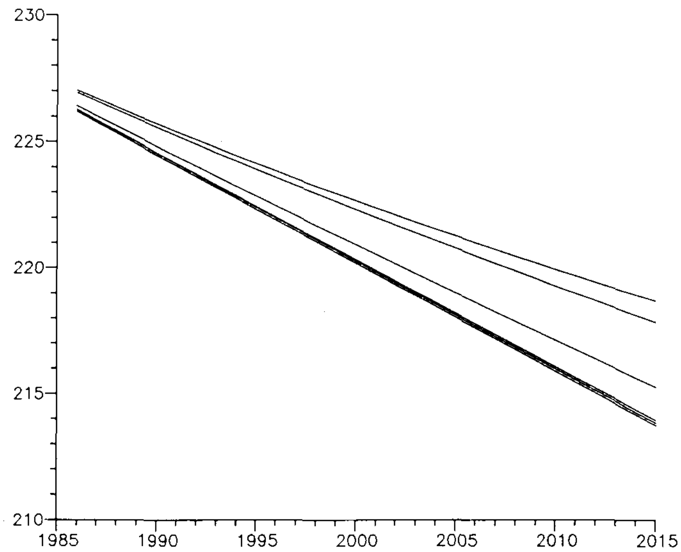

Figure 7. Median Predictions of the Annual Minimum for the Mile Race (1986–2015) From Six Models: B, C, E, G, D, A (from top to bottom).

performances. Subsequent discussion (Maher 1983; Morton 1984; Wootton and Royston 1983) suggested that their fitting method was in error but did not refute the model itself, and Morton (1983, 1984) claimed good results for the same model by a different fitting method.

For the mile-records data, using the estimates reported for Model C in Table 5, we obtain the asymptote $\beta_0 - \beta_1/\beta_2$ as 2 minutes, 58 seconds, with an estimated standard error of 63 seconds. Fitting the same model to the best performance data of Table 4, we can estimate the asymptote as 1 minute, 22 seconds, with a standard error of 395 seconds. In both cases, the estimate looks absurd with a standard error so large as to render the estimate meaningless. By fitting the Gumbel and GEV versions of the same models (Models F and I) we obtain an asymptote of 52.0 seconds (standard error, 409) for Model F applied to mile records, and 3 minutes, 25.5 seconds (standard error, 10.8 seconds), for Model I applied to mile best performances. No estimate was obtained for Model F applied to mile best performances, possibly because of the restriction to $\beta_2 > 0$ and because a local maximum could have been obtained with $\beta_2 < 0$. These estimates were based on the data for 1931–1985, and there were similar wide variations

Table 7. Cramer–Rao Lower Bound for the Standard Deviation of the Estimated Asymptote

θ_3	Estimated asymptote[a] (seconds)	Standard deviation
.06587[b]	82.1	167.9
.2173	200.0	20.0
.2699	210.0	14.3
.3071	215.0	11.9
.3561	220.0	9.7

NOTE: The data are based on Model (6.1) for the mile best-performance data of 1931–1985. θ_1, θ_2, and σ are fixed at maximum likelihood values; θ_3 varies.
[a] That is, $\theta_1 - \theta_2^2/(2\theta_3)$.
[b] Maximum likelihood value.

in applying the same model to other portions of the series and to other series.

It is not clear from these results whether the difficulty arises from the model or from the method of fitting it. To address this point, some Cramer–Rao bounds for standard deviation were calculated, representing lower bounds for any method of estimation. Explicit calculations were done for the case of normal errors with no censoring. Model (2.8) was first reparameterized as

$$\mu_n(\theta_1, \theta_2, \theta_3) = \theta_1 - \frac{\theta_2^2}{2\theta_3}[1 - \exp\{-n\theta_3/(5\theta_2)\}], \quad 1 \leq n \leq N, \quad (6.1)$$

from which the linear case results as $\theta_3 \to 0$. Substituting parameters for the 1931–1985 mile-race data, I set $\theta_1 = 251.2989$, $\theta_2 = 4.7216$, $\theta_3 = .06587$, $\sigma = 2.3284$, and $N = 55$; the estimated asymptote is $\theta_1 - \theta_2^2/(2\theta_3) = 82.1$, with a Cramer–Rao standard deviation of 168. The last value may be compared with 395 obtained earlier from the observed information matrix. It is interesting to vary the value of θ_3 (keeping the other parameters fixed so that the linear fit is unaffected) and thus obtain the Cramer–Rao bound under a variety of assumptions about the true model. Some results of this procedure are given in Table 7. For example, with $\theta_3 = .2699$ (which would correspond to an asymptote of 210 seconds) we obtain a Cramer—Rao bound of 14.3 seconds for the standard error of the asymptote's estimate. This purely hypothetical value gives some indication of what might be possible if the model were correct, though even in this case the standard error is the same order of magnitude as the eventual improvement being predicted.

Thus, even if Model (2.8) or (6.1) were correct, with parameter values that might correspond to reasonable judgment about the long-term improvement in record performances, the standard error of the estimate appears to be too large for the estimate to be valuable. In practice, given the uncertainty about the model's fit and the wide variability of estimates corresponding to different error distributions and different portions of the series, it is questionable whether any useful estimate of the asymptote is possible.

To summarize, although I have not presented any evidence to contradict the exponential-decay model as such, the difficulties of estimating the parameters make it very doubtful that this model can be used to produce meaningful performance estimates in the distant future.

7. CONCLUSIONS

The method of maximum likelihood fitting, as developed here, is conceptually simple and may be applied to a wide range of models for records data. The numerical algorithms developed appear good enough to find a local maximum when one exists. There are cases, however, most of them involving the GEV distribution, when no local maximum exists, and in these cases my methods fail. The principal model adopted has been a linear-drift model with normal errors, though over very long periods of time the slope of the linear drift changes. On the other hand, there is no evidence that either the quadratic-drift model or exponential-decay model provides a better fit overall. These models appear to provide a satisfactory basis for short-term forecasting of future performances, but I am extremely skeptical about the use of such methods to predict the long-term limit of record performances.

[*Received June 1986. Revised June 1987.*]

REFERENCES

Aitkin, M., and Clayton, D. (1980), "The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Data Using GLIM," *Applied Statistics,* 29, 156-163.

Ballerini, R., and Resnick, S. (1985), "Records From Improving Populations," *Journal of Applied Probability,* 22, 487-502.

——— (1987), "Records in the Presence of a Linear Trend," *Advances in Applied Probability,* 19, 801-828.

Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A. (1983), *Graphical Methods for Data Analysis,* Boston: Duxbury Press.

Chatterjee, S., and Chatterjee, S. (1982), "New Lamps for Old: An Exploratory Analysis of Running Times in Olympic Games," *Applied Statistics,* 31, 14-22.

Cheng, R. C. H., and Amin, N. A. K. (1983), "Estimating Parameters in Continuous Univariate Distributions With a Shifted Origin," *Journal of the Royal Statistical Society,* Ser. B, 45, 394-403.

De Haan, L. (1984), "Extremal Processes and Record Values," in *Statistical Extremes and Applications,* ed. J. Tiago de Oliveira, Dordecht: D. Reidel, pp. 297-309.

De Haan, L., and Verkade, E. (1987), "On Extreme Value Theory in the Presence of a Trend," *Journal of Applied Probability,* 24, 62-76.

Galambos, J. (1978), *The Asymptotic Theory of Extreme Order Statistics,* New York: John Wiley.

Glick, N. (1978), "Breaking Records and Breaking Boards," *The American Mathematical Monthly,* 85, 2-26.

Gumbel, E. J. (1958), *Statistics of Extremes,* New York: Columbia University Press.

Hosking, J. R. M. (1985), "Algorithm AS215: Maximum-Likelihood Estimation of the Parameters of the Generalized Extreme-Value Distribution," *Applied Statistics,* 34, 301-310.

Keller, J. B. (1974), "Optimal Velocity in a Race," *The American Mathematical Monthly,* 81, 474-480.

Lietzke, M. H. (1954), "An Analytical Study of World and Olympic Records," *Science,* 119, 333-336.

Lloyd, B. B. (1966), "Energetics of Running: An Analysis of World Records," *Advancement of Science,* 22, 515-530.

Maher, M. J. (1983), Letter to the Editor, *Applied Statistics,* 32, 313-314.

Morton, R. H. (1983), "The Supreme Runner: What Evidence Now?" *Australian Journal of Sport Sciences,* 3, No. 2, 7-10.

——— (1984), Letter to the Editor, *Applied Statistics,* 33, 317-318.

Naylor, J. C., and Smith, A. F. M. (1982), "Applications of a Method for the Efficient Computation of Posterior Distributions," *Applied Statistics,* 31, 214-225.

Prescott, P., and Walden, A. T. (1980), "Maximum Likelihood Estimation of the Parameters of the Generalised Extreme Value Distribution," *Biometrika,* 67, 723-724.

——— (1983), "Maximum Likelihood Estimation of the Parameters of the Three-Parameter Generalized Extreme Value Distribution From Censored Samples," *Journal of Statistical Computation and Simulation,* 16, 241-250.

Ranneby, B. (1984), "The Maximum Spacing Method: An Estimation Method Related to the Maximum Likelihood Method," *Scandinavian Journal of Statistics,* 11, 93-112.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications* (2nd ed.), New York: John Wiley.

Riegel, P. S. (1981), "Athletics Records and Human Endurance," *American Scientist,* 69, 285-290.

Rockette, H., Antle, C., and Klimko, L. A. (1974), "Maximum Likelihood Estimation With the Weibull Models," *Journal of the American Statistical Association,* 69, 246-249.

Smith, R. L. (1985), "Maximum Likelihood Estimation in a Class of Nonregular Cases," *Biometrika,* 72, 67-92.

Smith, R. L., and Miller, J. E. (1986), "A Non-Gaussian State Space Model and Application to the Prediction of Records," *Journal of the Royal Statistical Society,* Ser. B, 48, 79-88.

Smith, R. L., and Naylor, J. C. (1987), "A Comparison of Maximum Likelihood and Bayesian Estimates for the Three-Parameter Weibull Distribution," *Applied Statistics,* 36, 358-369.

Sweeting, T. J. (1980), "Uniform Asymptotic Normality of the Maximum Likelihood Estimator," *The Annals of Statistics,* 8, 1375-1381.

Tryfos, P., and Blackmore, R. (1985), "Forecasting Records," *Journal of the American Statistical Association,* 80, 46-50.

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.

Wootton, R., and Royston, J. P. (1983), Letter to the Editor, *Applied Statistics,* 32, 88-89.

Yang, M. C. K. (1975), "On the Distribution of the Inter-Record Times in an Increasing Population," *Journal of Applied Probability,* 12, 148-154.