Who is the Fastest Man in the World?

Robert TIBSHIRANI

I compare the world record sprint races of Donovan Bailey and Michael Johnson in the 1996 Olympic Games, and try to answer the questions: 1. Who is faster?, and 2. Which performance was more remarkable? The statistical methods used include cubic spline curve fitting, the parametric bootstrap, and Keller's model of running.

KEY WORDS: Sprinting; World record; Curve fitting.

1. INTRODUCTION

At the 1996 Olympic Summer Games in Atlanta both Donovan Bailey (Canada) and Michael Johnson (United States) won gold medals in track and field. Bailey won the 100 meter race in 9.84 seconds, while Johnson won the 200 meter race in 19.32 seconds. Both marks were world records. After the 200 m race, an excited United States television commentator "put Johnson's accomplishment into perspective" by pointing out that his record time was less than twice that of Bailey's, implying that Johnson had run faster. Of course, this is not a fair comparison because the start is the slowest part of a sprint, and Johnson only had to start once, not twice.

Ato Bolton, the sprinter who finished third in both races, was also overwhelmed by Johnson's performance. He said that, although normally the winner of the 100 meter race is considered the fastest man in the world, he thought that Johnson was the now the fastest.

In this paper I carry out some analyses of these two world record performances. I do not produce a definitive answer to the provocative question in the title, as that depends on what one means by "fastest." Hopefully, some light is shed on this interesting and fun debate. Some empirical data might soon become available on this issue: a 150 meter match race between the two runners is tentatively scheduled for June 1997.

2. SPEED CURVES

The results of the races are shown in Tables 1 and 2.

A straightforward measure of a running performance is the speed achieved by the runner as a function of time. The first line of Table 3 gives the interval times for Bailey, obtained from Swiss Timing and reported in the *Toronto Sun* newspaper. These times were not recorded for Johnson. The value 7.7 at 70 m is almost surely wrong, as it would imply an interval time of only 0.5 seconds for 10 m. I contacted Swiss Timing about their possible error, and they rechecked their calculations. As it turned out, the split times were computed using a laser light placed 20 m behind the starting blocks, and they had neglected to correct for this 20 m gap in both the 70 and 80 m split times. The corrected times are shown in Table 3.

The estimated times at each distance shown in Table 4 were obtained manually from a videotape of the races. Here is how I estimated these times. I had recorded the 100 m hurdles race on the same track. Using the known positioning of the hurdles, I established landmarks on the infield whose distance from the start I could determine. Then by watching a video of the sprint races in slow motion, with the race clock on the screen, I estimated the time it took to reach each of these markings.

Table 5 compares the estimated and official split times. After the 40 m mark, the agreement is fairly good. The disagreement at 10, 20, and 30 m is due to the paucity of data and the severe camera angle for that part of the race. Fortunately, these points do not have a large influence on the results, as our error analysis later shows. Overall, this agreement gives us some confidence about the estimated times

Table 1. Results for 1996 Olympic 100 m Final; The Reaction Time is the Time it Takes for the Sprinter to Push Off the Blocks after the Firing of the Starter's Pistol; DQ Means Disqualified

Name	Time	Reaction time
1. Bailey, Donovan (Canada)	9.84	+.174
2. Fredericks, Frank (Namibia)	9.89	+.143
3. Bolton, Ato (Tobago)	9.90	+.164
4. Mitchell, Dennis (United States)	9.99	+.145
5. Marsh, Michael (United States)	10.00	+.147
6. Ezinwa, Davidson (Nigeria)	10.14	+.157
7. Green, Michael (Jamaica)	10.16	+.169
8. Christie, Linford (Great Britain)	DQ	

Wind speed: +.7 m/s

Table 2. Results for 1996 Olympic 200 m Final; "?" Means the Information was Not Available

Name	Time	Time at 100 m	Reaction time
1. Johnson, Michael (United States)	19.32	10.12	+.161
2. Fredericks, Frank (Namibia)	19.68	10.14	+.200
3. Bolton, Ato (Trinidad and Tobago)	19.80	10.18	+.208
4. Thompson, Obadele (Barbados)	20.14	?	+.202
5. Williams, Jeff (United States)	20.17	?	+.182
6. Garcia, Ivan (Cuba)	20.21	?	+.229
7. Stevens, Patrick (Belgium)	20.27	?	+.151
8. Marsh, Michael (United States)	20.48	?	+.167

Wind speed: +.4 m/s

Robert Tibshirani is Professor, Department of Preventive Medicine and Biostatistics and the Department of Statistics, University of Toronto, Toronto, Ont., Canada M5S 1A8. The author thanks Trevor Hastie, Geoff Hinton, Joseph Keller, Bruce Kidd, Keith Knight, David MacKay, Carl Morris, Don Redelmeier, James Stafford, three referees, and two editors for helpful comments, Cecil Smith for providing Bailey's official split times from Swiss Timing, and Guy Gibbons of Seagull Inc. for providing the corrected version of the Swiss Timing results. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

Table 3. Official Times at Given Distances for Bailey; The "?" Indicates a Suspicious Time, Later Found to be in Error

Distance (m)	0	10	20	30	40	50	60	70	80	90	100
Original time (s)	.174	1.9	3.1	4.1	4.9	5.6	6.5	7.7?	8.2	9.0	9.84
Corrected time (s)	.174	1.9	3.1	4.1	4.9	5.6	6.5	7.2	8.1	9.0	9.84

Table 4. Estimated Times at Given Distances; Bailey Starts at the 100 m Mark; "+" Denotes Distance Past 100 m: For Example, "+12.9" Means 112.9 m

Distance (m)	0	50	100	+12.9	+40.3	+49.4	+67.7	+76.9	+86.0	+100
Bailey:	.174			2.8	5.0	5.7	7.0	7.8	8.5	9.84
Johnson:	.161	6.3	10.12	11.4	14.0	14.8	16.2	17.0	17.8	19.32

Table 5. Comparison of Official and Estimated Interval Times for Bailey

Distance (m)	0	10	20	30	40	50	60	70	80	90	100
Official	.174	1.9	3.1	4.1	4.9	5.6	6.5	7.2	8.1	9.0	9.84
Estimated	.174	2.1	3.4	4.3	5.1	5.7	6.4	7.2	8.0	8.9	9.84

Table 6. Estimated Times (seconds) for Johnson for Distances over 100 m

	100	110	120	130	140	150	160	170	180	190	200
Johnson	10.12	11.10	12.09	13.06	13.97	14.83	15.61	16.40	17.26	18.23	19.32

for Johnson, and some idea of the magnitude of their error. The speed curves were estimated by fitting a cubic smoothing spline to the first differences of the times, constraining the curves to be 0 at the start of the race. The curves for each runner are shown in the top panel of Figure 1. Because Bailey's 100 m was much faster than Johnson's first 100 m but slower than his second 100 m, it seems most interesting to make the latter comparison. Hence I have shifted Bailey's curve to start at time 10.12 s, and Johnson's time at 100 m.

If Johnson's speed curve always lay above Bailey's, then this analysis would have provided convincing evidence in favor of Johnson because he achieved his speed despite having already run 100 meters. However, Bailey's curve does rise above Johnson's, and achieves a higher maximum (13.2 m/s for Bailey, 11.8 m/s for Johnson). A 95% confidence interval for the difference between the maxima, computing using the parametric bootstrap, is $(-.062, 1.15)$. Hence there is no definitive conclusion from this comparison. The Appendix gives details of the computation of this confidence interval.

We note that the estimate of 13.2 m/s for Bailey's maximum speed differs from the figure of 12.1 m/s reported by Swiss timing. Bailey's estimated final speed is 12.4 m/s versus 11.5 m/s reported by Swiss timing. This size of discrepancy is not unexpected because the interval times are only given to within .1 of a second. When a sprinter is running at top speed, he covers 10 m in approximately .8 s, giving a speed of $10/.8 = 12.5$ m/s. Now if each of the interval times are off by .05 s, then the estimated speed ranges from $10/.9 = 11.1$ m/s to $10/.7 = 14.3$ m/s.

Who would win a race of say 150 meters? Here is a simple-minded approach to the question. Bailey's speed at the 100 m mark was 12.4 m/s, and his speed was decreasing by only .036 m/s every 10 m. Johnson's estimated time at 150 m was 14.83 s, as given in Table 6. In order to beat that time Bailey would need "only" to maintain an average speed of more than 10.02 m/s for another 50 m. Of course, it is not clear whether he could do this. In the next section we appeal to a parametric model to perform the necessary extrapolation.

For interest, in the bottom panel of Figure 1 we compare Bailey's curve to that from Ben Johnson's 1987 9.83 s world record race (he was later disqualified for drug usage). They achieved roughly the same time in quite different ways: Ben Johnson got a fast start, and then maintained his velocity; Bailey accelerated much more slowly, but achieved a higher maximum speed.

3. PREDICTIONS FROM KELLER'S MODEL

Keller (1973) developed a model of competitive running that predicts the form of the velocity curve for a sprinter using his resources in an optimal way. Here we use his model to predict the winner of a 150 m race.

According to Keller's theory, the force $f(t)$ per unit mass at time t, applied by a sprinter in the direction of motion, may be written as

$$f(t) = \frac{dv(t)}{dt} + \frac{v(t)}{\tau} \qquad (1)$$

where $v(t)$ is the velocity and τ is a damping coefficient. This is just Newton's second law, where it is assumed that the resistance force per unit mass is $v(t)/\tau$.

Keller estimated τ to be .892 s from various races. Excellent overviews of Keller's work are given by Pritchard (1993) and Pritchard and Pritchard (1994).

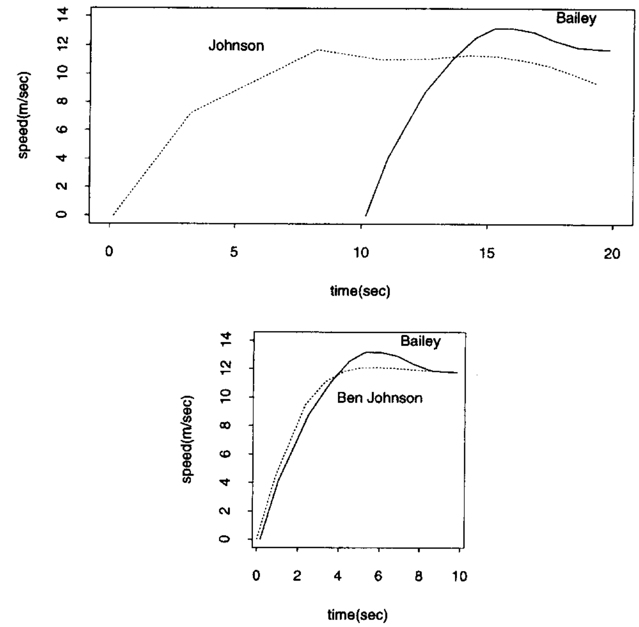

Figure 1. Top Panel: Estimated Speed Curves for Bailey and Johnson. Bailey's curve has been shifted to start at time 10.12 s, Johnson's time at 100 m. Bottom panel: estimated speed curves for Bailey and Johnson from the latter's 1987 world record race.

Starting with assumption (1) and a model for energy storage and usage, Keller shows that the optimal strategy for a runner is to apply his maximum force F during the entire race, leading to a velocity curve

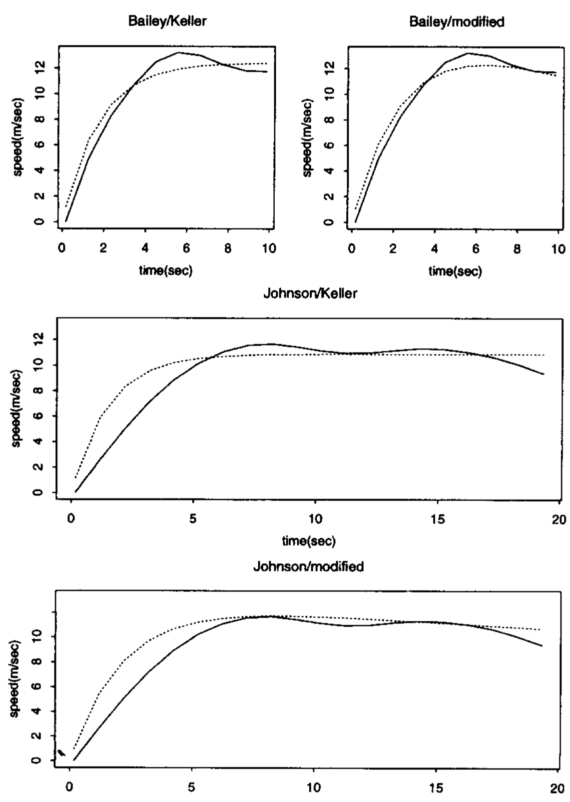

Figure 2. Top Row: Optimal Velocity Curves (Broken) for Bailey's 100 m. The top left panel uses Keller's model (2); the top right panel uses the modified model (4). The middle and bottom rows show the fit of the Keller and modified models for Johnson's 200 m. In all panels the solid curve is the corresponding actual (estimated) velocity curve from the top panel of Figure 1.

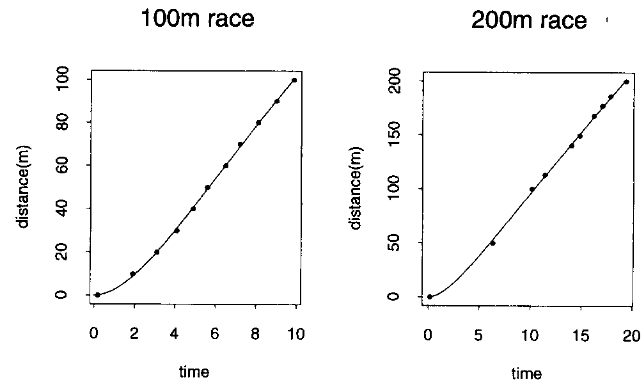

Figure 3. Estimated Distance Curves and Actual Distances (Points) from Least Squares Fit of Model (4).

$$v(t) = F\tau(1 - e^{-t/\tau}). \tag{2}$$

This applies to races of less than 291 m. For greater distances there is a different optimal strategy. By integrating (2) we obtain the distance traveled in time t:

$$D(t) = F\tau^2(t/\tau + e^{-t/\tau} - 1). \tag{3}$$

Figure 2 (top left and middle panels) shows the optimal speed curves for the 100 and 200 m races, with Bailey's and Johnson's superimposed. We used least squares on the (time, distance) measurements to find the best values of τ and F for each runner in equation (3): these were (1.74, 7.16) for Bailey and (1.69, 6.80) for Johnson.

We can use (3) to predict the times for a 150 m race; note that the reaction times must be included as well. The predictions are 13.97 s (Bailey) and 15.00 s (Johnson).

The same model also predicts a completely implausible 200 m time of 17.72 s for Bailey. One shortcoming of the model is the fact that the velocity curve (2) never decreases, but observed velocity curves usually do. To rectify this it seems reasonable to assume that a sprinter is unable to

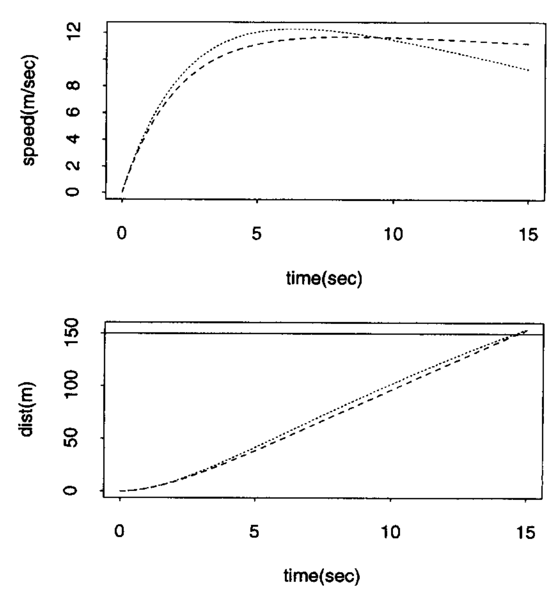

Figure 4. Estimated Optimal Velocity and Distance Curves over 150 m for Bailey (Dotted) and Johnson (Dashed).

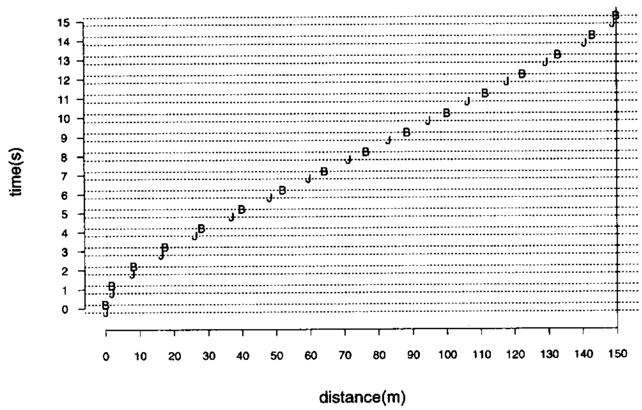

Figure 5. The Predicted Race in 1 s Snapshots. Shown is the Estimated Distance Traveled by Bailey (B) and Johnson (J) at Time = 0 s, 1 s, . . . , 14 s, and at the End of the Race (Time 14.73 s).

maintain his maximum force F over the entire race, but instead applies a force $F - c \cdot t$ for some $c \geq 0$. Using this in (1) leads to velocity and distance curves

$$v(t) = k - ct\tau - ke^{-t/\tau}$$

$$D(t) = kt - c\tau t^2/2 + \tau k(e^{-t/\tau} - 1) \qquad (4)$$

where $k = F\tau + \tau^2 c$. We fit this model to the observed distances by least squares, giving parameter estimates for (τ, F, c) of (2.39, 6.41, .20) and (2.06, 6.10, .05) for Bailey and Johnson, respectively. The fitted distance values are plotted with the actual ones in Figure 3. Note that the estimated maximum force is greater for Bailey than Johnson, but decreases more quickly. Bailey also has a higher estimated resistance.

The estimated curves are shown in the top right and bottom panels of Figure 2. The estimated 150 m times from this model are 14.73 s for Bailey and 14.82 s for Johnson. The latter is very close to the estimated time of 14.83 s at 150 m in the Olympic 200 m race from Table 6.

Figure 4 shows the estimated optimal velocity and distance curves over 150 m from the model, and Figure 5 depicts the predicted race in 1 s snapshots. Bailey is well ahead at the early part of the race, but starts to slow down earlier. Johnson gains on Bailey in the latter part of the race, but does not quite catch him at the end. The estimated winning margin for Bailey is .09 s. The bootstrap percentile 95% confidence interval for the difference is (.03 s, .26 s), and the bias-corrected 95% bootstrap confidence interval is (.02 s, .19 s). One thousand bootstrap replications were used—see the Appendix for details. Figure 6 shows a boxplot of difference in the predicted 150 m times from the bootstrap replications.

Note that this model does not capture a possible change of strategy by either runner in a 150 m race. This might result in different values for the parameters.

From Keller's theory one can also predict world record times at various distances as a function of F and τ. Keller fit his predicted world record times to the actual ones, for distance from 50 yards to 10,000 m, in 1973. From this he obtained the estimates $F = 12.2$ m/s^2, $\tau = .892$ s. The fit was quite good: for 100 m—9.9 s (actual), 10.07 s (predicted); for 200 m—19.5 s (actual), 19.25 s (predicted). (The world records that Keller reports in 1973 of 9.9 and 19.5 s are questionable. The 100 m record was 9.95 s, although 9.9 s was the best hand-timed performance. The 200 m record was 19.83 s.) It is interesting that at the time, the 100 m record was faster than expected, but the 200 m record was slower. Johnson's performance brings the 200 m world record close to the predicted value. It may be that the 200 m record has been a little "soft," with runners focusing on the more glamourous 100 m race. Note that the predictions do not include a component for reaction time: with Johnson's reaction time of .161 s, the predicted record would be 19.41 s.

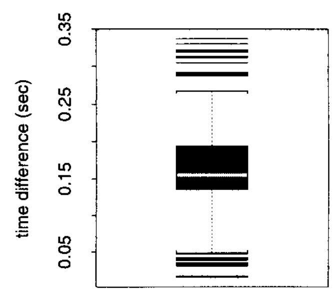

Figure 6. Boxplot of Johnson's Minus Bailey's Predicted 150 m Times from 1,000 Bootstrap Replications.

4. ADJUSTMENT FOR THE CURVE

Johnson's first 100 m (10.12 s) was run on a curve, and Bailey's was run on a straight track. Figure 7 shows the sprint track.

In the previous analysis we ignored this difference. Assuming we want to predict the performance of the runners over a straight 150 m course (the course type for the May 1997 race has not been announced at the time of this writing), we should adjust Johnson's 200 m performance accordingly. Intuitively, he should be given credit for having achieved his time on the more difficult curved course.

What is the appropriate adjustment? The centripetal acceleration running of an object moving at a velocity v around a circle of radius r is $a = v^2/r$. The radius of the circular part of the track is $100/\pi = 31.83$ m. With Johnson's velocity ranging from 0 to 11.8 m/s, his centrepital acceleration ranges from 0 to 4.37 m/s^2. We cannot simply add this acceleration to the acceleration in the direction of motion because the centrepital acceleration is at right angles to the direction of motion. However, he does biological work in achieving this acceleration, and hence spends energy. Unfortunately, just how much energy is expended is

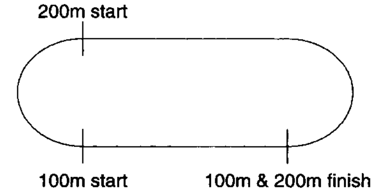

Figure 7. The Sprint Track, Showing Start and Finish Lines for the 100 and 200 meter Races.

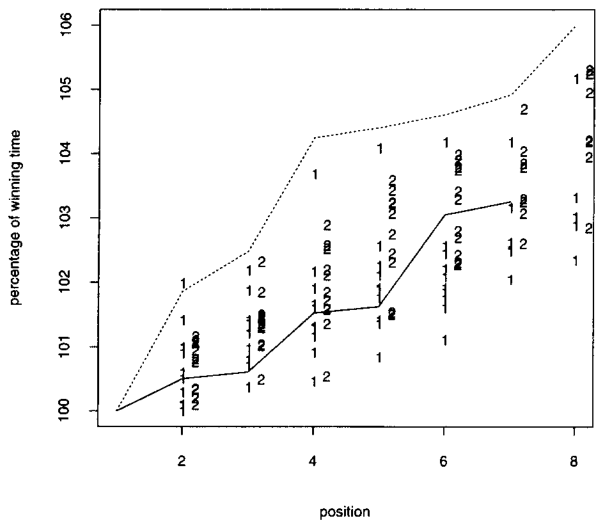

Figure 8. Percentage that Each Runner Achieved as a Function of the Winning Time in the Race for the 100 meter (Solid Curve) and 200 meter (Broken Curve) Races. The "1s" and "2s" correspond to the previous nine Olympic 100 and 200 m finals, respectively.

difficult to measure, in the opinion of the physicists that I consulted. Hence I have not been able to quantity this effect.

5. COMPARISON TO OTHER RACE COMPETITORS

In the rest of this paper I focus on the question of which of the two performances was more remarkable. These two races were particularly unique because the same two runners (Fredericks and Bolton) finished second and third in both. This suggests an interesting comparison. Figure 8 shows the percentage that each runner achieved as a function of the winning time in the race for the 100 (solid curve) and 200 meter race (broken curve). Johnson's winning margin was particularly impressive. Also plotted in the figures are the corresponding percentages achieved in the previous nine Olympic games, going back to 1952. (Throughout this analysis I restrict attention to post-1950 races because before that time races were run on both straight and curved tracks, and it was not always recorded which type of course had been used.) There has never been a winning margin as large as Johnson's in a 200 meter Olympic race, and only once before in a 100 meter race. This was Robert Hayes' 10.05 s performance in 1964 versus 10.25 s for the second place finisher. Johnson's margin over the second place Fredericks is also larger than the margin between the winner and third place in all but two of the races.

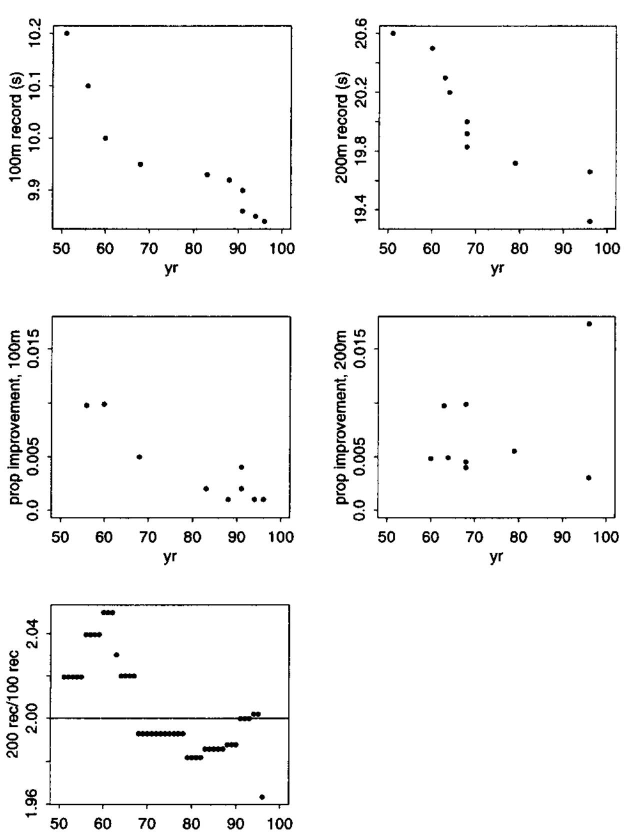

Figure 9. The Top Panels Show the Evolution of the 100 m (Left) and 200 m (Right) World Records. The middle panel shows the proportion improvement of the existing record that was achieved each time in the 100 m (left) and 200 m (right). The bottom left panel shows the evolution of the ratio of the 200 to 100 m world record times.

6. EVOLUTION OF THE RECORDS

The top panels of Figure 9 show the evolution of the 100 and 200 meter records since 1950. The proportion improvements, relative to the existing record, are shown in the middle panels of Figure 9. Johnson's 19.32 performance represented a 1.7% improvement in the existing record, the largest ever. (Tommie Smith lowered the 200 m world record to 20.00 s in 1968, a 1.0% improvement from the existing world record of 20.2 s. However, the 20.2 s value was a hand-timed record: the existing automatic-timed record was 20.36 s, which Smith improved by 1.8%.) If we include Johnson's 19.66 world record in the 1996 U.S. Olympic Trials, then overall he lowered Pietro Mennea 19.72 world record by 2.02% in 1966.

The bottom left panel of Figure 9 shows the ratio of the 200 meter world record versus the 100 meter world record from 1950 onward. The ratio has hovered both above and below 2.0, with Johnson's world record moving it to an all-time low of 1.963. The average speed in the 100 m record race was 10.16 m/s, and that for the 200 m race was 10.35 m/s, the fastest average speed of any of the sprint or distance races. A ratio of below 2.0 is predicted by the mathematical model of Keller (1973).

7. CONCLUSIONS

Who is faster, Bailey or Johnson? The answer depends on the definition of "faster," and there is no unique way of comparing two performances at different distances. Our results are inconclusive on this issue:

- It is not fair to compare the average speeds (higher for Johnson) because the start is the slowest part of the race, and Johnson had to start only once.
- Bailey appeared to achieve a higher maximum speed, although the difference in maxima was not statistically significant at the .05 level; Johnson maintained a very high speed over a long time interval.
- Predictions from an extended version of Keller's optimal running model suggest that Bailey would win a

(straight) 150 m race by .09 s. However, they do not account for the fact that Johnson's times are based on a curved initial 100 m.

It would clearly be a close race, and there are a number of factors I have not accounted for. This entire comparison is based on just one race for each runner: consistency and competitiveness come into play any race. Perhaps most important is the question of strategy. Each runner would train for and run a 150 meter race differently than a 100 or 200 meter race. The effect of strategy is impossible to quantify from statistical considerations alone.

Whose performance was more remarkable? Here, Johnson has the clear edge:

- Johnson's winning margin over the second and third place finishers (the same runners in both races!) was much larger than Bailey's, and was the second largest in any Olympic 100 or 200 m final race.
- Johnson's percentage improvement of the existing world record was the largest ever for a 100 or 200 m race. However, the 200 m record might have been a little "soft" because it was well above the record as predicted by Keller's theory.

SOURCES

- David Wallechinsky, *The Complete Book of the Summer Olympics*, Boston, New York, Toronto, London: Little, Brown, 1996.
- Mika Perkimki, csmipe@uta.fi, *World-Wide Track & Field Statistics On-Line*, [http://www.uta.fi/~csmipe/sport/index.html].
- IBM, *Official Results of the 1996 Centennial Olympic Games*, [http://results.atlanta.olympic.org].
- *Track and Field News* (Oct. 1996).

APPENDIX: ERROR ANALYSIS

The data in Table 4 were obtained manually from a videotape of the race, and hence are subject to measurement error. To assess the effect of this error, we applied the parametric bootstrap (Efron and Tibshirani 1993). The maximum amount of error in the times was thought to be around $\pm.05$ s for Bailey's times, and $\pm.20$ s for Johnson's early times and $\pm.15$ s for Johnson's last 100 m times. Therefore, I added uniform noise on these ranges to each time measurement. For each resulting dataset I estimated the speed curves for Bailey and Johnson. This process was repeated 1,000 times.

The observed difference in the maximum speeds was $13.2 - 11.8 = 1.4$ m/s. The upper and lower 2.5% points of the 1,000 observed differences was $(-.062, 1.15)$. The same parametric bootstrap procedure was used for the error analysis of the fit of Keller model, in Section 3, and was used to produce Figure 6.

[Received October 1996. Revised January 1997.]

REFERENCES

Efron, B., and Tibshirani, R. (1993), *An Introduction to the Bootstrap*, London: Chapman & Hall.

Keller, J. (1973), "A Theory of Competitive Running," *Physics Today*, 42–45.

Pritchard, W. (1993), "Mathematical Models of Running," *SIAM Review*, 359–379.

Pritchard, W., and Pritchard, J. (1994), "Mathematical Models of Running," *American Scientist*, 546–553.