Calculating the Age Factor Function for Competitive Running

Michael Lam

Center for Radiophysics and Space Research, Department of Astronomy, Cornell University, Ithaca, NY 14853
Email: mtl82@cornell.edu

ABSTRACT

Time as a function of age in competitive track and road racing is broadly understood on an individual level but difficult to constrain given the enormous number of factors that enter into performance. Current, publically available large databases of archived race results are rich sources of information for statistical analysis. We look at four different distances over five different races spanning a large range of event type, from middle distance to endurance racing. We calculate the amount of slow down of different groups of runners with age and develop an empirical model for the observed trends.

Subject headings: methods: data analysis – methods: statistical – catalogs

1. INTRODUCTION

Competitive racing is one of the simplest sports to become an active member in though hardly the easiest. Runners can spend enormous amounts of time training compared to the amount of time spent in racing, all with the simple goal of running the as fastest as possible. Improvement comes season after season through hard training but also naturally through physical maturity. Typical training wisdom suggests that one reaches their peak physical shape in young-to-middle adulthood, depending on the race distance, and progressively slow down over the years, though people continue to race to the best of their ability.

The rate of this slow down is not well constrained. Much of this knowledge has been past between runners and common, experiential beliefs persist as a result. Several attempts to determine race time as a function of age have been made. The current, popularly used age graded model, those which allow for comparisons of races between people of two different ages, comes from standards set by World Masters Athletics (WMA, originally World Association of Veteran Athletes), recognized by the International Association of Athletics Federations (IAAF) governing organization as the regulating body for masters athletics (Jones 2011, Fair 2003). WMA maintains updated age graded tables used by many event organizers for guaranteed entry qualification times in larger races and even in overall results, including the USA Track & Field Masters Championships.

The WMA tables have been compiled by several different people who all use similar methods. They compile some combination of single age world records, that is, the fastest times run in an event by a person of a given age. Thus, for any given age, there is only one person of each gender who holds the single age world record. Forming a curve of time run versus age, they attempt to fit a functional form to this curve and characterize it. The results range in success, typically form more of a lower bound for the data, and miss wildly at fitting the very young and very old ages. This method also suffers from a bias in choosing only the top individuals in the world for a given age for computing statistics. Even if the curve was fit perfectly, it cannot be assumed that this will apply to runners slower by any non-small amount. The assumption that each single curve will apply statistically to everyone else is a broad leap. We look to formulate a similar Age Factor Function (note that the WMA analysis refers to this quantity as 1/[Age factor]) applicable to runners of all levels.

2. OBSERVATIONS

We avoided discerning statistical trends by tracking individuals over time and instead looked at large numbers of individuals. For a given individual runner, changes in training plans, injuries, and even other external life effects can cause dramatic differences in times over years (e.g. sleep, work, etc.). Even perfectly healthy individuals who undergo the same training plan every year would be subject to varying conditions in the race, whether internal or external. The statistical uncertainties would overwhelm any visible trends. Computationally, this would also be a much more difficult task as well, as one would have to cross-reference entries with each other to find suitable candidates for testing, converting an $\mathcal{O}(n)$ problem into at least an $\mathcal{O}(n^2)$ problem. Thus, we look for large data sets for the greatest statistical significances.

We collected data on five different races. Because sponsorship of races changes frequently, we use the colloquial names constant throughout their histories. Since recent results from races are often available online, we collected all data via internet archived results except for one. While there are many different races results available, we looked for specific properties with certain selection criteria to minimize variations.

For optimal statistics, we searched for races with large numbers of competitors to improve estimates. Additionally, larger races typically accommodate participnts of all levels, removing the potential bias mentioned previously. Thus, races must have existed for a long enough period of time to have gathered popularity over the years. Bigger races, however, suffer from what we will refer to as the gun time discrepancy. For small races, one's net time from starting line to finish line will roughly equal the gun time from the start of the race to when one crosses the finish line as the time it takes to get to the starting line is small. Automatic timing, typically in the form of electronic chips in larger road races, allows for easy calculation of the net time, which is the physically significant time for a racer. Many races do not utilize chips due to the increase in costs and even many large races do not use them, resulting in inaccurate timing and a higher reliance on self-timing from the runner. Thus, to increase the number of samples while maintaining that they are the physically relevant net times, we need races that have been going on for a long time, started with smaller numbers, but then switched early to chip timing once the race became significantly larger.

In addition, races outdoors can be held under highly variable conditions, so we looked for as best consistency as possible on a year-to-year basis, ruling out races in certain climates or at certain times of the year, though we did not correct for any possible condition changes between years.

In the Results, we briefly discuss the variations found between the different races. We did not correct for potential weather changes and their effects on performance nor did we take into account the difficulty in the courses, though we looked for flat ones for better times and bigger field sizes. In actuality, the times themselves do not wholly matter as only the changes with age are the quantities of interest, and thus the deviations from ideal race conditions are unimportant so long as they are consistent.

2.1. *Fifth Avenue Mile*

The New York Road Runners (NYRR) Fifth Avenue Mile boasts an extremely competitive field in a mostly flat, straight-line mile in New York City from East 80th Street to East 60th Street along Fifth Avenue. It is held in late September with typically moderate weather conditions. Starting in 1981, NYRR web archives start at 1986 but only have the top few finishers. We used Open results from 2000-2013, with a total sample of N=36229.

2.2. *Hartshorne Mile*

Held in late January in Ithaca, NY, the Hartshorne Masters Mile, hosted by the Finger Lakes Running Club (FLRC), is the only indoor race data set we obtained. It is run on a flat, 200 m indoor track and is exclusive to Masters age groups (40+) for men though includes a sub-Masters catagory for women (30-39). The race is run in many different sections so reported hand timings are accurate. Started in 1968, we have the full archive Men's results and the Women's from 2000, though the results indicate the race having been held since at least 1991. After clipping several missing or poorly formatted finishing times, we have N=1421 for men and N=285 for women.

2.3. *Syracuse Festival of Races 5k*

This is one of the fastest, record-eligible road race courses in the world, with a 30 ft elevation gain over the whole course. Both facts lead to extremely competitive race conditions, with a course record of 13:27[CITE]. It is held in early October, usually with excellent racing conditions. We have the full results from 1993-2012 except for 1995 as there are no ages included with the results. Based upon the changes in results formats, we believe that chip timing began in 2004, though the number of runners in the years prior was not large enough for the gun time discrepancy to cause a significant shift in results, so we still use these data. For men, N=6843, and for women, N=5753.

2.4. *Boilermaker 15k*

Started in 1978, this is now the largest 15k in the US, capped at 15,000 entries this year. Held in early July, the race conditions are usually not ideal for distance running, though are typically consistent. It is not entirely flat but only ranges in 350 ft over the entire course. We have all available online results, from 2003-2012, though we did not use 2008 results as only gun time results are posted. In total, N=92215.

2.5. *New York City Marathon*

One of the most famous marathons in the world and one of the six World Marathon Majors races, the New York City Marathon is the premier event of the NYRR. Held on the first Sunday of every November since 1970 except in 2012 due to the aftermath of Hurricane Sandy, it typically has quite variable weather and is not flat, largely due to bridge crossings between boroughs. Chip timing began in 1999, and so the Gun Time discrepancy is large for previous years, though breaking the races into several starting waves helps to alleviate the problem. Nonetheless, its enormous size allows was its biggest draw, with N=796577. We found online archives from the years 1984, 1987, 1989, and 1995 corrupted, resulting in much fewer data. Note also that the NYRR does not report times above 10 hours, so the results are chopped off on this extreme. The statistics, however, do not suffer severely as a result of either of these.

3. ANALYSIS

Our primary computed quantity is a two-dimensional histogram of counts of runners as a function of time t in seconds and age a in years. Figure 1 shows this operation on results of the Men's Hartshorne Mile. Each bin has $dt = 5$ s and $da = 1$ yr. Rather than normalize along each age bin and correct against zero counts, we computed the cumulative distribution function (CDF) for each age. This allows us to look at competitors of all different levels for a given age, assuming that someone from the k-th percentile remains at roughly the same level compared to other runners, on average. Note the important distinction that we can only say that a person is in the k-th percentile for runners in a given race, not of people overall. Thus, our usage of "median" refers to the median runner in a race and not about the level to which the median human can compete. Due to time restrictions, we divide these data into quartiles and look at the 25th, 50th, and 75th percentiles.

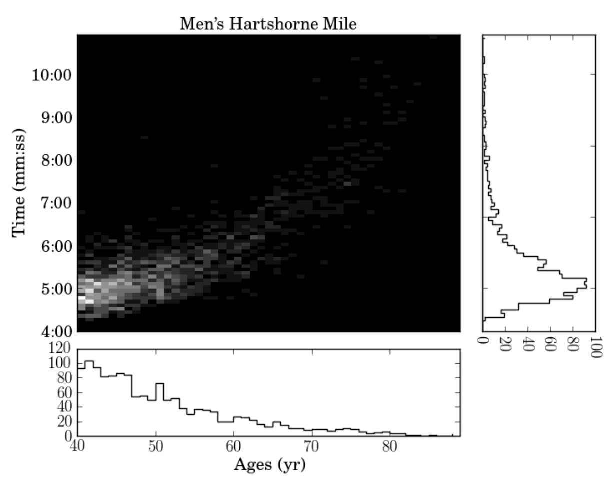

Fig. 1.—: Two-dimensional histogram of runners with respect to time run and age for the Men's Hartshorne Mile. Whiter values indicate increased counts. The subplot on the right shows the cumulative sum going horizontally across bins for a given time and the subplot on the bottom shows the cumulative sum going vertically down for a given age. A smooth trend of time increasing with age is clearly visible.

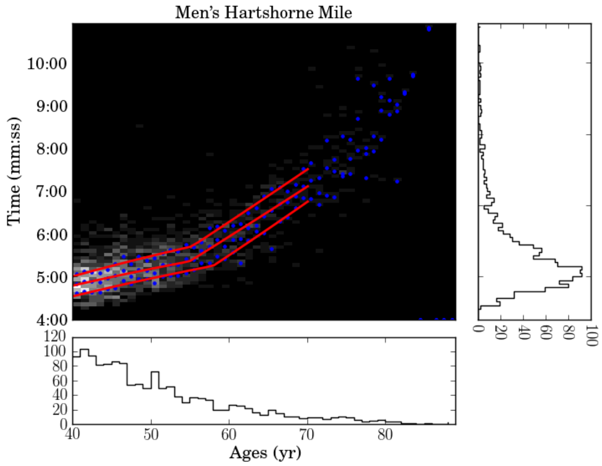

Fig. 2.—: Two-dimensional histogram for the Men's Harshorne Mile. The (blue) points denote each of the three percentiles (25th, 50th, 75th) and the (red) lines indicate the fitted broken power laws.

Finding the unique functional form of the Age Factor Function was non-trivial. To simplify the investigation, we only looked at ages $20 \leq a \leq 70$. Curves of constant percentile traced paths similar to that of the WMA-style age factors curves but beyond this range, the curves were not well constrained given the low number of counts and unpredictability of these runners. While the boundaries could no doubt be pushed slightly on either side, we took a conservative range to continue our analysis.

We found that a simple polynomial fit the data extremely well, even one of low order ($n = 2$). An exponential function also fit reasonably well. We wanted to avoid the arbitrariness of a many parameter model such as in the WMA Jones method of fitting multiple, continuous, low-order polynomials to the curve. As is ubiiquitous throughout nature, we next looked to fitting the curve to a power law form. Replotting these curves on log-space axes, we noticed not a single linear trend as would be indicative of a power law, but two lines joined at a particular cutoff, as indicative of a broken power law. This was true for each of the three different percentile ranges, both genders, and all distances. Therefore, we decided to pursue this form as an empirical fit, without explanation, though this will briefly be discussed later. Figure 2 shows the same Hartshorne Results as before but now with the three different percentiles marked and the broken power law fits. Note that we did not compare residuals from the various methods to find the best approach but we went ahead with the broken power law model given the observed shape.

Changepoint analysis is a useful technique in determining changes in the statistical properties of a noisy function before and after a specific point. In this case, we looked for the change in slope between the two lines in log-space at the value of the to be determined cutoff point. This method was no better than fitting the power law directly, and considering the fitting function found the cutoff point with relative ease, we decided against continuing to refine this method but note its usefulness and speed.

We model a broken power law in log-space as a line $y(x) = mx + b$, with

$$\log t_1 = m_1 \log a + b_1, a \leq a_0 \tag{1}$$

$$\log t_2 = m_2 \log a + b_2, a \geq a_0, \tag{2}$$

where for $i = 1, 2$ marking the younger and older age power laws, respectively, $logt_i$ is the logarithm base 10 of the times of each point in seconds, m_i represents the slope of the line, equivalent to the index of the power law, $loga$ is the logarithm of the age in years, b_i is the y-intercepts of the line. a_0 is the cutoff age, though in log-space we typically compare the quantity $\log a_0$ to $loga$. Since we demand the broken power law to be continuous, we impose the condition

$$\begin{aligned} b_2 &= m_1 \log a_0 + b_1 - m_2 \log a_0 \\ &= \log(t_1(a_0)) - m_2 \log a_0, \end{aligned} \tag{3}$$

thus reducing the number of free parameters to four. Converting these into linear-space, we obtain the functional form of the broken power law as

$$t_1(a) = 10^{b_1} \left(\frac{a}{\text{year}}\right)^{m_1} \quad [\text{s}] \tag{4}$$

$$\begin{aligned} t_2(a) &= 10^{b_2} \left(\frac{a}{\text{year}}\right)^{m_2} \quad [\text{s}] \\ &= 10^{b_1} \, (a_0)^{m_1 - m_2} \left(\frac{a}{\text{year}}\right)^{m_2} \quad [\text{s}], \end{aligned} \tag{5}$$

with t_1 and t_2 representing the piecewise Age Factor Function in units of seconds.

Using the Python Programming Language, along with the packages NumPy and SciPy, we fit these curves via the basic non-linear least squares Levenberg-Marquardt Algorithm. From this, we calculated the errors on the parameter estimates and compared them across the different races.

4. RESULTS

We plot all two-dimensional histograms for all of the races in Figures 3 through 12 below. Comments for each race will follow within.

4.1. *Fifth Avenue Mile*

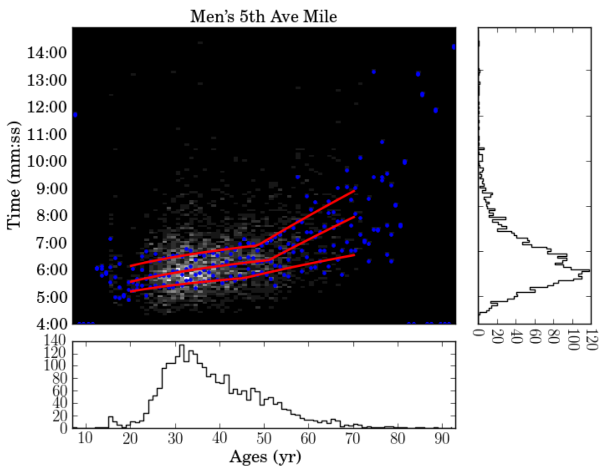

Fig. 3.—: Two-dimensional histogram for the Men's Fifth Avenue Mile.

Table 1
Men's Fifth Avenue Mile Parameters

Percentile	25th	50th	75th
$\log a_0$	1.66 ± 0.05	1.71 ± 0.02	1.68 ± 0.02
a_0 (yr)	45.7 ± 5.3	51.3 ± 2.4	47.9 ± 2.2
b_1	2.35 ± 0.05	2.34 ± 0.06	2.42 ± 0.06
m_1	0.11 ± 0.04	0.14 ± 0.04	0.12 ± 0.04
m_2	0.32 ± 0.07	0.70 ± 0.10	0.69 ± 0.10

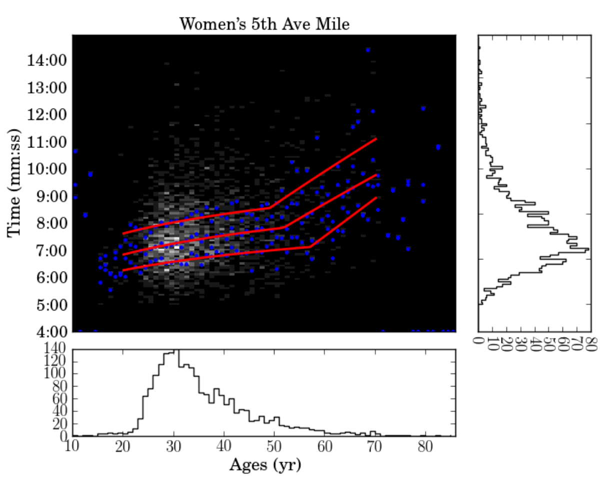

Fig. 4.—: Two-dimensional histogram for the Women's Fifth Avenue Mile.

Table 2
Women's Fifth Avenue Mile Parameters

Percentile	25th	50th	75th
$\log a_0$	1.76 ± 0.01	1.71 ± 0.03	1.69 ± 0.03
a_0 (yr)	57.5 ± 1.3	51.3 ± 3.6	49.0 ± 3.4
b_1	2.42 ± 0.04	2.43 ± 0.09	2.49 ± 0.09
m_1	0.12 ± 0.03	0.14 ± 0.06	0.13 ± 0.06
m_2	1.10 ± 0.23	0.73 ± 0.17	0.73 ± 0.11

4.2. *Harsthorne Mile*

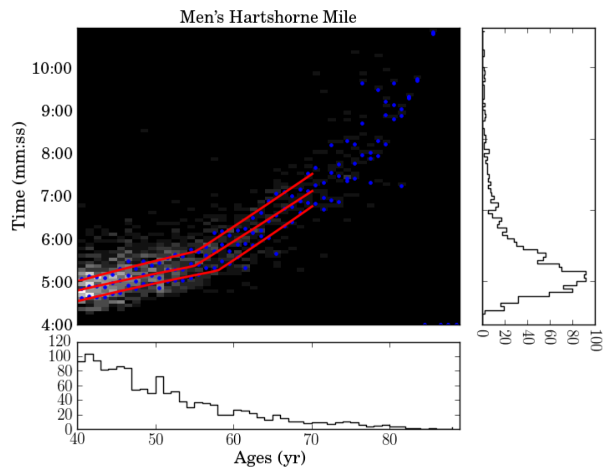

Fig. 5.—: Two-dimensional histogram for the Men's Harshorne Mile, reproduced here for comparison purposes

Table 3
Men's Hartshorne Mile Parameters

Percentile	25th	50th	75th
$\log a_0$	1.76 ± 0.01	1.74 ± 0.01	1.74 ± 0.01
a_0 (yr)	57.5 ± 1.3	54.5 ± 1.3	54.5 ± 1.3
b_1	1.81 ± 0.09	1.90 ± 0.07	1.83 ± 0.07
m_1	0.39 ± 0.05	0.35 ± 0.04	0.40 ± 0.04
m_2	1.32 ± 0.12	1.16 ± 0.06	1.15 ± 0.06

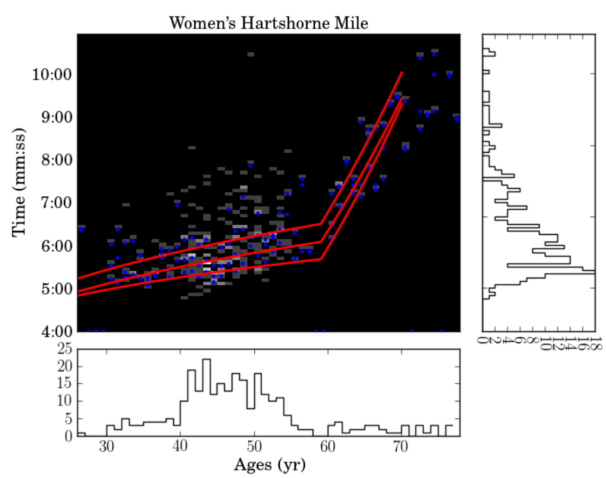

Fig. 6.—: Two-dimensional histogram for the Women's Harshorne Mile.

Table 4
Women's Hartshorne Mile Parameters

Percentile	25th	50th	75th
$\log a_0$	1.77 ± 0.01	1.77 ± 0.02	1.77 ± 0.02
a_0 (yr)	58.9 ± 1.4	58.9 ± 2.7	58.9 ± 2.7
b_1	2.19 ± 0.14	2.11 ± 0.19	2.12 ± 0.19
m_1	0.20 ± 0.09	0.26 ± 0.12	0.26 ± 0.12
m_2	2.87 ± 0.73	2.56 ± 0.94	2.53 ± 0.94

4.3. *Syracuse Festival of Races 5k*

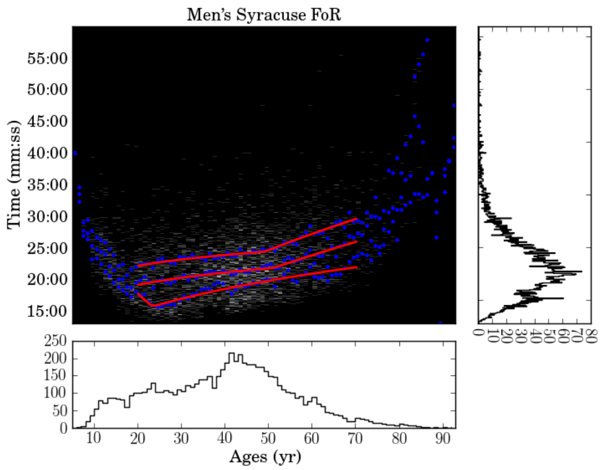

Fig. 7.—: Two-dimensional histogram for the Men's Syracuse Festival of Races 5k.

Table 5
Men's Syracuse Festival of Races Parameters

Percentile	25th	50th	75th
$\log a_0$	1.37 ± 0.01	1.71 ± 0.01	1.69 ± 0.01
a_0 (yr)	23.4 ± 0.5	51.3 ± 1.2	49.0 ± 1.1
b_1	4.01 ± 0.47	2.88 ± 0.2	2.99 ± 0.02
m_1	-0.76 ± 0.35	0.14 ± 0.01	0.11 ± 0.01
m_2	0.30 ± 0.02	0.55 ± 0.04	0.54 ± 0.04

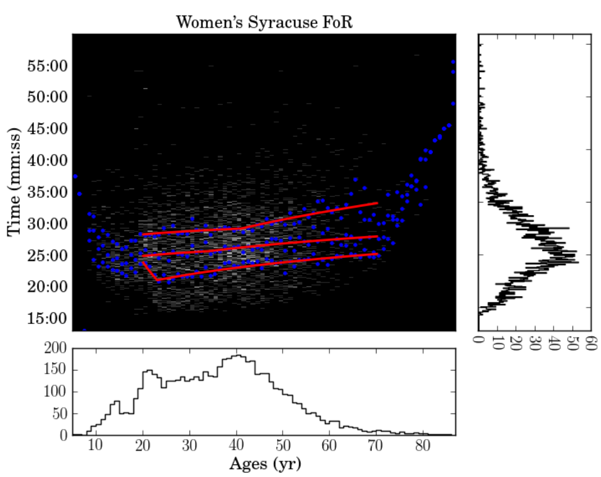

Fig. 8.—: Two-dimensional histogram for the Women's Syracuse Festival of Races 5k.

Table 6
Women's Syracuse Festival of Races 5k Parameters

Percentile	25th	50th	75th
$\log a_0$	1.36 ± 0.02	1.53 ± 0.05	1.61 ± 0.05
a_0 (yr)	$22.9.0 \pm 1.1$	33.9 ± 3.9	40.7 ± 4.7
b_1	4.31 ± 0.52	3.10 ± 0.06	3.17 ± 0.06
m_1	-0.89 ± 0.39	0.06 ± 0.04	0.04 ± 0.04
m_2	0.54 ± 0.01	0.12 ± 0.05	0.24 ± 0.05

4.4. *Utica Boilermaker 15k*

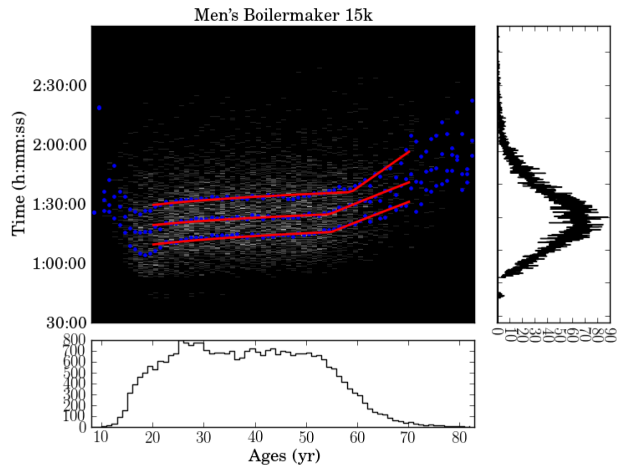

Fig. 9.—: Two-dimensional histogram for the Men's Utica Boilermaker 15k.

Table 7
Men's Utica Boilermaker 15k Parameters

Percentile	25th	50th	75th
$\log a_0$	1.76 ± 0.01	1.74 ± 0.01	1.77 ± 0.01
a_0 (yr)	57.5 ± 1.3	54.5 ± 1.3	58.9 ± 1.4
b_1	3.51 ± 0.02	3.59 ± 0.02	3.65 ± 0.02
m_1	0.09 ± 0.01	0.07 ± 0.01	0.06 ± 0.01
m_2	0.75 ± 0.07	0.69 ± 0.11	1.09 ± 0.11

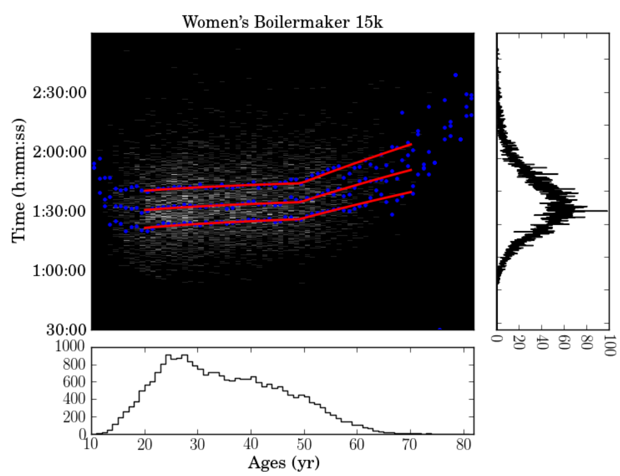

Fig. 10.—: Two-dimensional histogram for the Women's Utica Boilermaker 15k.

Table 8
Women's Utica Boilermaker 15k Parameters

Percentile	25th	50th	75th
$\log a_0$	1.69 ± 0.02	1.69 ± 0.01	1.69 ± 0.01
a_0 (yr)	49.0 ± 2.3	49.0 ± 1.1	49.0 ± 1.1
b_1	3.62 ± 0.03	3.67 ± 0.03	3.73 ± 0.03
m_1	0.06 ± 0.02	0.05 ± 0.02	0.04 ± 0.02
m_2	0.42 ± 0.06	0.45 ± 0.06	0.49 ± 0.06

4.5. *New York City Marathon*

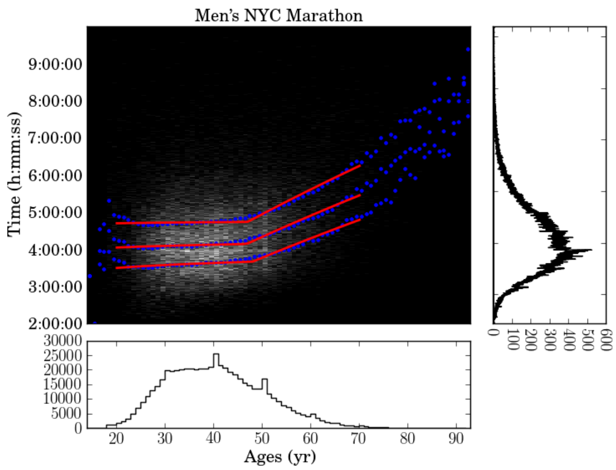

Fig. 11.—: Two-dimensional histogram for the Men's New York City Marathon.

Table 9
Men's New York City Marathon Parameters

Percentile	25th	50th	75th
$\log a_0$	1.68 ± 0.01	1.67 ± 0.01	1.67 ± 0.01
a_0 (yr)	47.9 ± 1.1	46.8 ± 1.1	46.8 ± 1.1
b_1	4.03 ± 0.02	4.13 ± 0.02	4.22 ± 0.02
m_1	0.05 ± 0.01	0.03 ± 0.01	0.01 ± 0.01
m_2	0.71 ± 0.03	0.68 ± 0.02	0.69 ± 0.02

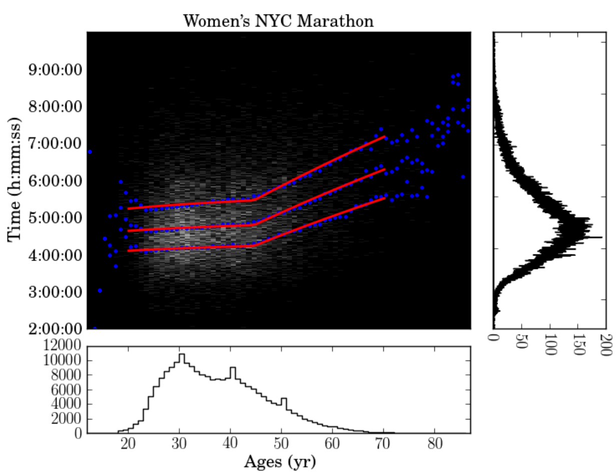

Fig. 12.—: Two-dimensional histogram for the Women's New York City Marathon.

Table 10
Women's New York City Marathon Parameters

Percentile	25th	50th	75th
$\log a_0$	1.65 ± 0.01	1.65 ± 0.01	1.65 ± 0.01
a_0 (yr)	44.7 ± 1.0	44.7 ± 1.0	44.7 ± 1.0
b_1	4.12 ± 0.01	4.17 ± 0.02	4.21 ± 0.02
m_1	0.04 ± 0.01	0.04 ± 0.01	0.05 ± 0.01
m_2	0.58 ± 0.02	0.60 ± 0.02	0.60 ± 0.02

4.6. *Inter-race Comparison*

The results of the parameter fits are shown in Figures 13, 14.

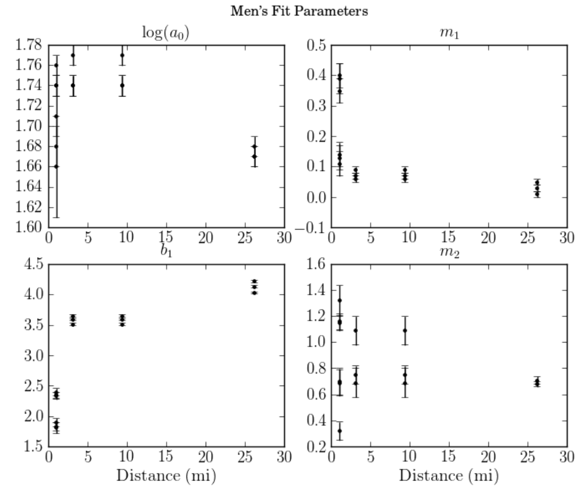

Fig. 13.—: Two-dimensional histogram for the Men's New York City Marathon.

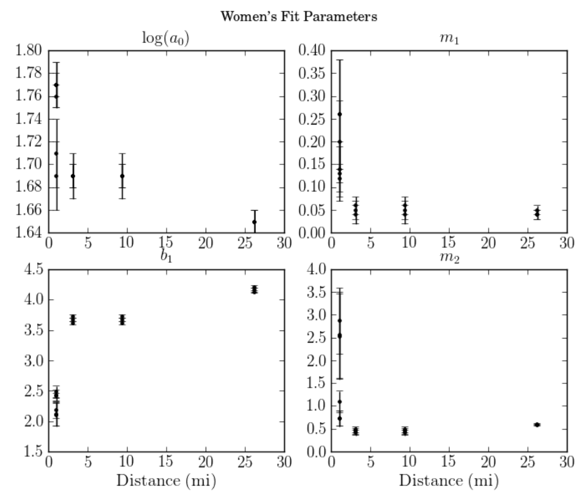

Fig. 14.—: Two-dimensional histogram for the Women's New York City Marathon.

5. DISCUSSION

5.1. *Fifth Avenue Mile*

These results are fairly typical though the scatter in the curves at large ages in quite high as the numbers of participants rapidly decrease, thus greatly increasing the errors in the least-squares fitting. While the race overall is much larger, there are more participants at this end in the Hartshorne Mile, and thus that race would be more useful

5.2. *Hartshorne Mile*

The curve for the Hartshorne Mile was already fairly visible before in the Men's race and the numbers agree fairly well, though it is interesting to note that the power law indices are much higher than for the counterpart Fifth Avenue Mile. In the Women's race, the low number of participants have clearly skewed the results.

5.3. *Syracuse Festival of Races 5k*

The most problematic piece of these curves are the apparently misplaced cutoff point for the 25th percentile curves for both men and women. While there is an obvious dip in the curve there, this cannot physically be the same cutoff as for every other race but is rather an artifact of the least-squares fitting over the imposed age range. Fixing this age range, or placing weights in the cost function, would prevent this in future analysis.

5.4. *Boilermaker 15k*

We originally attempted to look for the peak racing age for a given race. One test statistic would be to measure the minima of the Age Factor Function curves. For the Men's race, we see an obvious dip in all three percentiles at roughly $a = 19$ yr. This is counterintuitive to commonplace knowledge that runners peak in ability in their early 20s for shorter races and all the way up well beyond 30 for the marathon. For the Boilermaker, while there are roughly 2/3rds the number of runners total at age 20 than at the peak, we suspect that there is a higher occurrence of faster runners around age 20, thus skewing the CDFs and the placements of the percentiles on the histograms. Therefore, we believe this dip not to reflect the true peak age of a runner. Also as a result, we do not believe it prudent to use this test statistic as a measure for the peak age of any given race without further analysis.

5.5. *New York City Marathon*

With such a large sample size, these fits are obviously the best, and it is apparent that there seems to be some underlying structure present in the curves. It is interesting to note the increased numbers of participants at ages 30, 40, 50, and even 60, an unexpected find that suggests an increased desire to run the marathon at these ages, most probably to cross off a lifetime goal.

5.6. *Comparison*

It is interesting to note the cases of agreement and disagreement with the parameter fits. Both mile races provide the largest errors, as do the Syracuse Festival of Races 5k problems noted above. Whether or not the parameters agree over all ranges of distances is still not well determined but qualitatively the power law indices m_1 and m_2 seem to remain relatively constant as a function of race distance and much more so than cutoff age a_0. Whether this is the result of some physical mechanism or the statistical analysis we cannot yet say.

5.7. *Test Case*

As a demonstration of the utility of these models, we compute the time increase per year for a newly Masters-level male athlete ($a = 40$ yr) of approximately median level.

Table 11
Masters Test Case

Race	Median Time (h:mm:ss)	Δt/year (s/yr)	Δt/mile/year (s/mi/yr)
Mile	6:06.7	1.3	1.3
5k	21:11.0	4.4	1.4
15k	1:23:56.7	8.7	0.9
Marathon	3:34:45.6	15.9	0.6

This confirms the popular notion that time increases in this range at roughly 1 s per mile per year.

6. FUTURE WORK

Potentially, we can fit many more numbers of percentile curves from the histograms, providing an enormous expansion upon the single curves produced for the WMA tables. Combining the CDFs for each age and looking at the changes over age could yield the most information.

Over time, runners as a whole become faster as general health improves. As with other corrective factors, we did not adjust the times to account for this. By separating participants per year, we could potentially determine the rate independent again from world records

In Tim Noakes' book Lore of Running (2003), the author briefly references several papers discussing physiological changes with age. He notes that there seems to be a distinct point at roughly age 50 when aerobic capacity and muscle strength drastically change. This provides us with excellent support that our form of the Age Factor Function, a broken power law, is indeed a physically meaningful one.

7. CONCLUSIONS

After obtaining data from five races over four different distances, we analyzed the slow down with age in what we call the Age Factor Function. There are several problems with the analysis as noted though we find the results largely onsistent with experiential information.

We thank various runners of the High Noon Athletic Club, especially Jim Bisogni, for providing historic results to the Hartshorne Masters Mile. We also thank the FLRC, NYRR, the Syracuse Festival of Races, and the Boilermaker organizations for use of their archives.

REFERENCES

Fair, R., 2003, How Fast Do Old Men Slow Down?, http://www.nber.org/papers/w3757
Jones, A., 2011, http://home.roadrunner.com/ alanjones/AgeGrade.html
Noakes, T, 2003, Lore of Running
http://www.world-masters-athletics.org