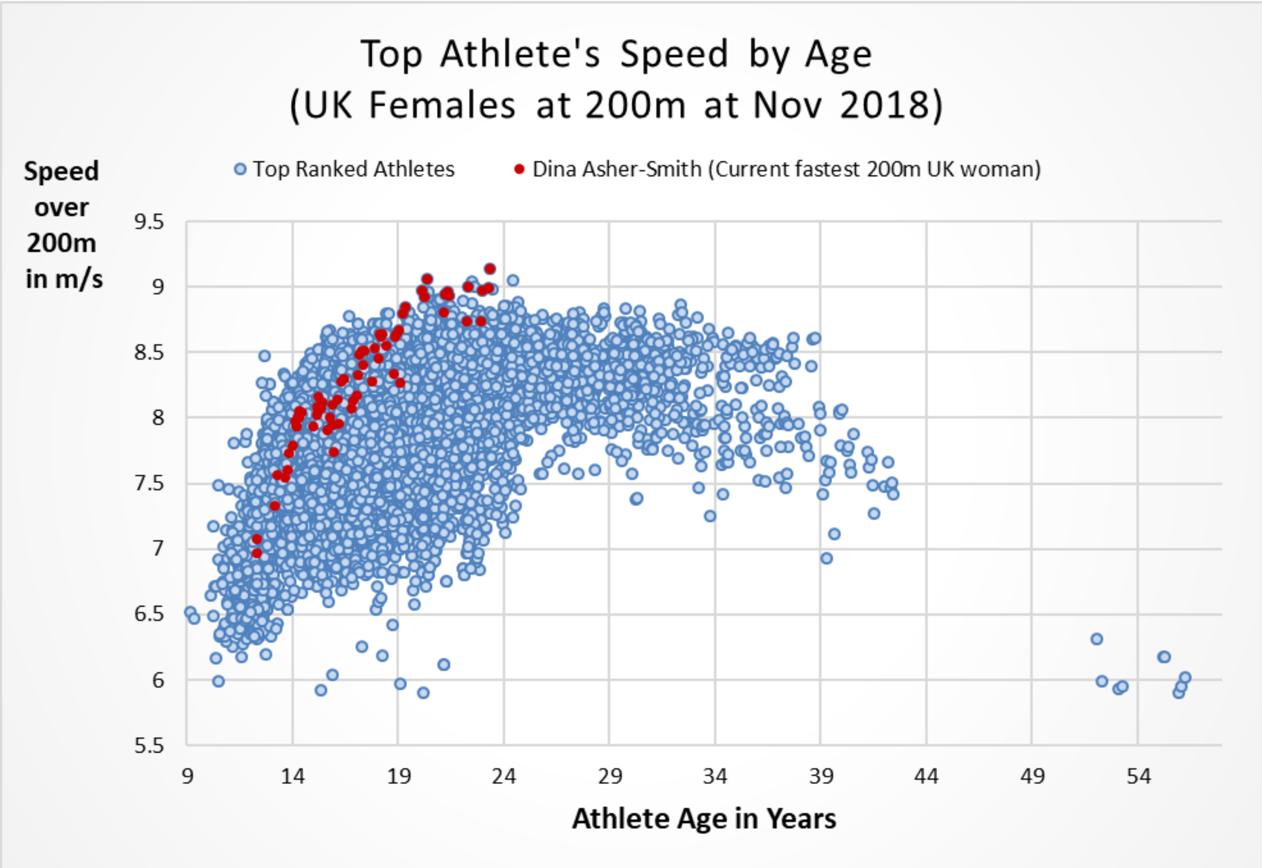
INM430-Principles of Data Science Coursework

Can future top athletes be identified when young?

Date: December 2018

This report should be read in conjunction with the related Jupyter notebook on the City website:
https://smcse.city.ac.uk/student/aczd032/inm430/INM430_Notebook_Mike_Cluley.html



Contents

This report should be read in conjunction with the related Jupyter notebook on the City website:
https://smcse.city.ac.uk/student/aczd032/inm430/INM430_Notebook_Mike_Cluley.html

1 Motivation

1.1 Domain

The selected domain is track and field (t&f) athletics. In the UK taxpayers fund sporting bodies to develop national teams and companies sponsor events and individual athletes. Per Hoye et al in 2009 [Ref1] the process of selecting and hiring future elite performers is partly based on the assessment of individual performances at a young age.
This raises the interesting question: “Can future top athletes be identified when young?”
In the UK the time horizon for official funding is 4 years. UK Sport allocates funds from the government and National Lottery. It makes an investment decision every 4 years. In the 2017-2020 funding cycle, 112 athletes were assessed to have a “credible chance” at winning a medal at either the 2020 or 2024 Olympic Games. UK Sport awarded £4m is awarded direct to 112 athletes and £23m for their training and other support.

The table 1 illustrates part of the challenge above showing the stats for top-ranked athletes 2016 vs 2012 for one event, Women’s 200m.

Athlete Name	2016 Rank or last race Year	Best 2016 Time	Age at 2016 Best Time	2012 Rank	Best 2012 Time	Age at 2012 Best Time	Progress 2012 to 2016
Dina Asher-Smith	1	22.31	20.7	11	23.49	17.9	Faster
Desiree Henry	2	22.46	21.0	6	23.28	16.9	Faster
Jodie Williams	3	22.69	22.9	*2nd in 2011*	-	-	Faster
Kat J-Thompson	4	22.79	23.4	16	23.73	19.6	Faster
Ashleigh Nelson	5	22.96	25.2	*21st in 2011*	-	-	Faster
Amarachi Pipi	6	23.2	20.5	34	24.25	16.6	Faster
Bianca Williams	7	23.27	22.5	13	23.59	18.3	Faster
Louise Bloor	8	23.36	30.5	7	23.36	26.8	Same
Jessica Ennis-Hill	9	23.36	30.4	2	22.83	26.5	Slower
Anyika Onuora	10	23.4	31.6	3	22.93	27.6	Slower
Margaret Adeoye	14	23.57	31.2	4	22.94	27.3	Slower
Shana Cox	29	24.88	31.1	8	23.38	27.2	Slower
Hayley Mills	*24th in 2018*	-	-	9	23.42	23.8	Slower
Sophie Papps	*2014*	-	-	10	23.48	17.7	
Abi Oyepitan	*2012 injury*	-	-	1	22.71	32.4	
Joice Maduaka	*2012*	-	-	5	23.25	38.6	

1.2 Question breakdown into Analysis Objectives and Expected Output

To assess the title question I have broken it down into the following sub-questions to address in this project:

1-At what age do athletes "peak"? What is the age distribution? Has this changed over time?

O1-Average peak age (& STDev) can be calculated from an analysis of All-Time Ranked athletes -It can then be summarised in a table for review.

2- How consistent are athlete rankings over time? What % of a cohort of top athletes in say, 2006, survive through to 2012?

O2-Specific ranked cohorts can be extracted and followed in subsequent year rankings. -Cohort survival can be charted for review.

3- The only published athletic data available relates to event dates, performance times and sometimes an athlete's data of birth. Could the relative age (eg the birth month) within a cohort be used as an additional feature? See literature review for findings in other sports.

O3-The birth months of ranked athletes can be plotted by cohort and charted to check for any trends that could contribute to predicting performance.

4-How well can a formula derived from earlier data be used to predict future rankings?

O4-Athlete performance data with age can be used by a curve fitting function to generate an individual line of best fit. This can be used to predict future performance eg 2018 results for the Top-50 athletes in 2014 can be predicted based on earlier data and compared to actual 2018 results.

1.3 Literature Review

1.3.1 Effect of age on athletic performance

In 2011 Berthelot et al [Ref 1] analysed the performances of international athletes by age see extract in Appendix. This analysis used just one best time per year with an average of 6 races per athlete. Much more data is available for UK athletes.

1.3.2 Relative Age Effect

School and club sports are organised into age groups. In the UK the specific athletics age group cut-off ends on 31st August (it follows the school year). A "relative age effect" has been noted where relatively older children tend to be more successful than younger children in the same cohort. The effect has been noted in several sports even at professional level eg professional Canadian hockey and volleyball, [Ref 4,5,6].

2 Analytical Strategy

2.1 Scope

The scope of this project is restricted to the UK Women's 200m sprint event.

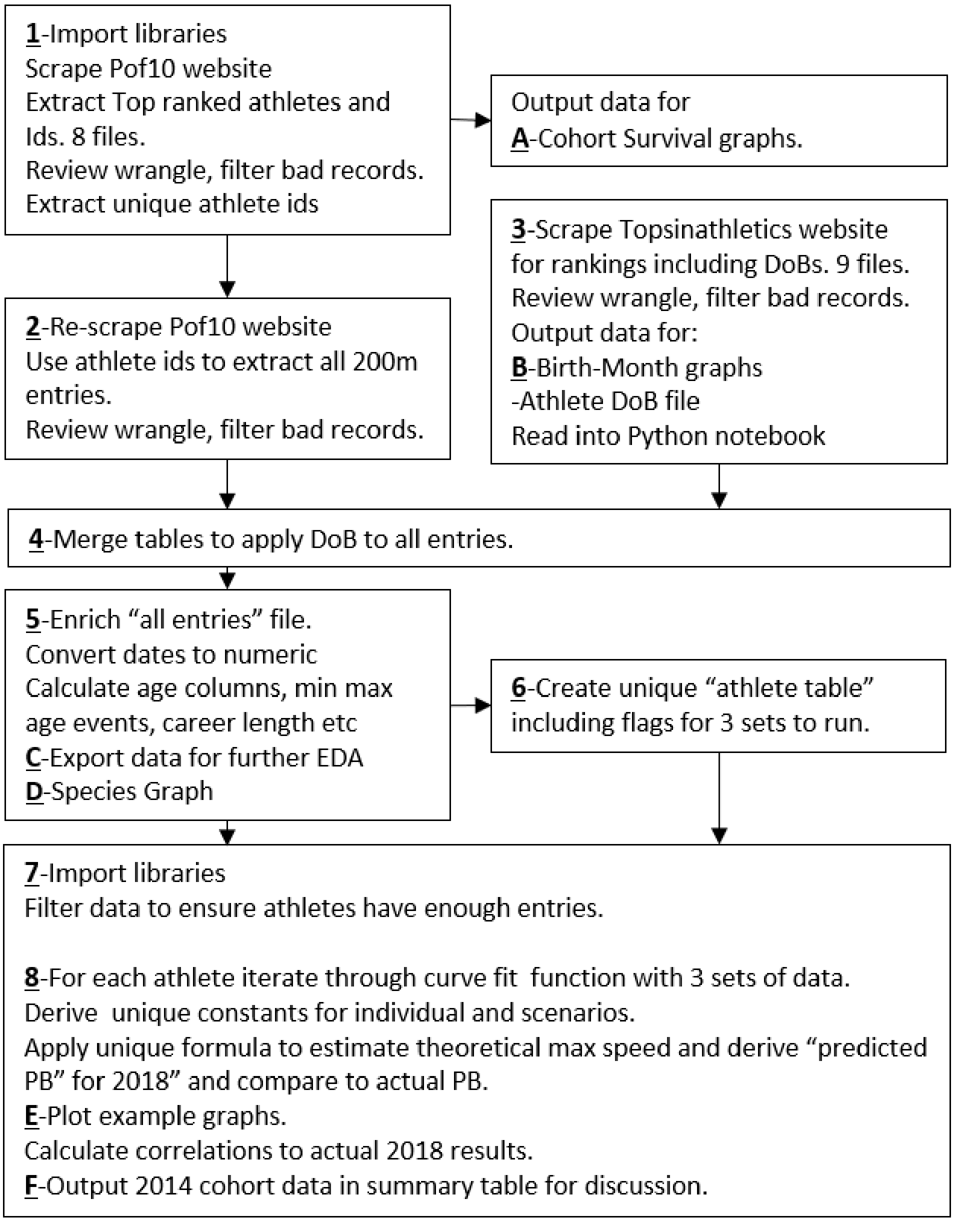
2.2 Data Sources and Ethical Considerations

Since 2006 results from official athletics events have been collated online. This analysis will use data scraped from 2 websites:
1- Power of 10 (https://www.thepowerof10.info)
2- Topsinathletics (http://www.topsinathletics.info)

All data used is available from publicly accessible websites. The names of elite athletes may be used and their approximate ages, but individual birth dates will not be disclosed.

2.3 Planning

The diagram below shows a plan of the work stages required to meet the objectives.



3 Analytical Process (see Jupyter notebook)

Refer to notes in computational notebook City Website link:
https://smcse.city.ac.uk/student/aczd032/inm430/INM430_Notebook_Mike_Culey.html

4 Analytical Findings and Reflections

4.1 Exploratory Data Analysis

Decade	No of Athletes	Average PB Time	Average Age at PB	STDev Age at PB	Average PB Age/ Age at Last Comp	Comment
2018-2010	43	23.27	22.6	3.66	91.94%	Decade not over so figs could change by end 2019
2009-2000	30	23.31	23.4	4.19	83.36%	
1999-1990	31	23.34	24.5	3.41	91.59%	
1989-1980	34	23.41	22.2	3.07	94.33%	
1979-1970	20	23.51	22.1	3.14	97.90%	
1969-1960	7	23.54	20.9	1.24	97.01%	
1959-1950	3	23.70	20.3	2.97	100.00%	
Overall	168	23.37	22.8	3.60	91.89%	

Table 2

- As per discussion in notebook, the overall average age for a PB is 22.8 years, with a standard deviation of 3.6 years.
- Average age and standard deviation are increasing over time increased over time.
- Table 2 also shows that most athletes stop competing soon after they reach their peak speed and decline begins.
- Table 1 on page 1 shows in top-10 ranked athletes in 2012 were aged between 16.6 and 38.6 years a range of 22 years!
- Table 1 shows a significant churn between with older athletes replaced by a new wave of sprinters between 2012 and 2018.
- The species chart (on cover) does confirm the overall trendline for speed performance rising to a peak and then declining following an exponential shape consistent with Moore's formula and previous findings for other athletic events [Ref 1 and see Extract in Appendix]

<u>The large standard deviation indicates that athletes can "peak" at a wide range of ages and applying assuming an athlete is "average" would be unreliable for predicting performance.</u>

4.2 Cohort Survival

The charts on the next page show how the original 2006 top-ranked survived through successive years or were replaced by their peers.

- As per discussion in notebook, the cohort's original 2006 composition declines, by 2012 less than 20% of the original 2006 remain in the top 100 and 50 [Charts 1,2]
- The Top 10 U20 cohort at 2012 appears exceptional as it has contributed 50% to the All Time top 10 by 2018 [Chart3]

4.2.1 Cohort Survival Charts 2006 to 2012

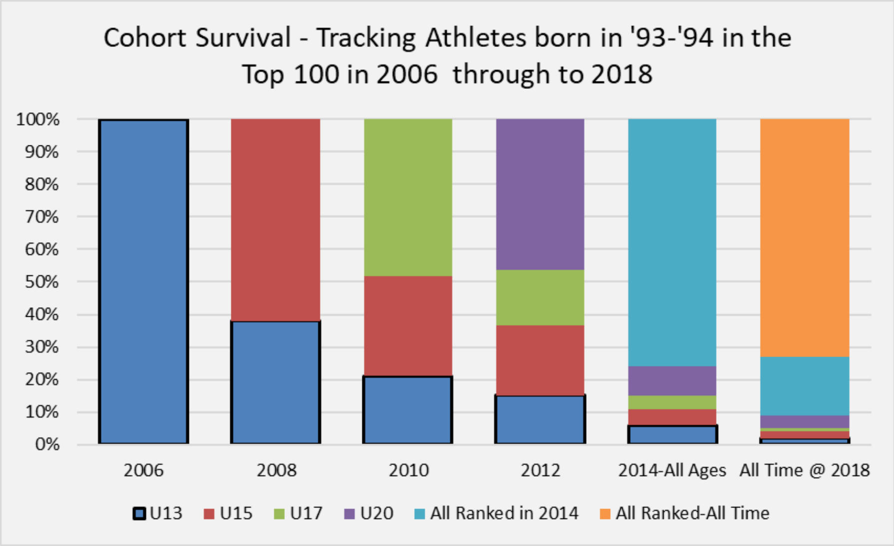


Chart 1

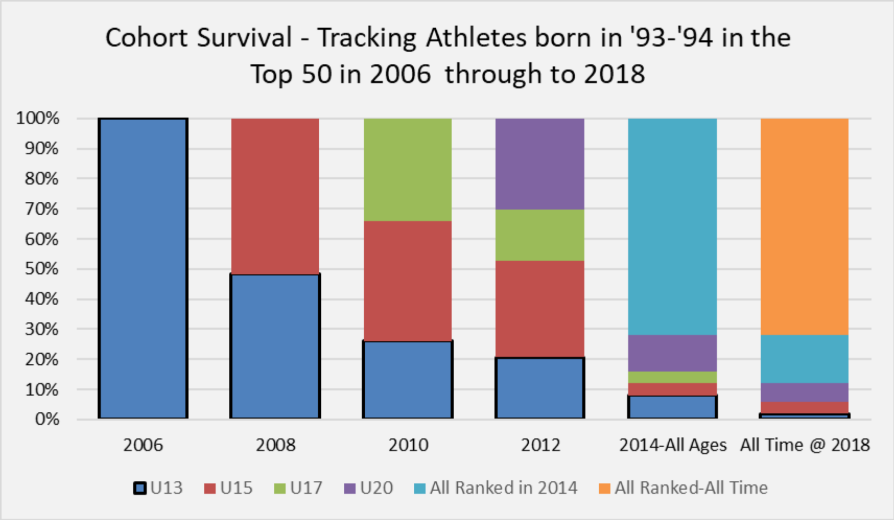


Chart 2

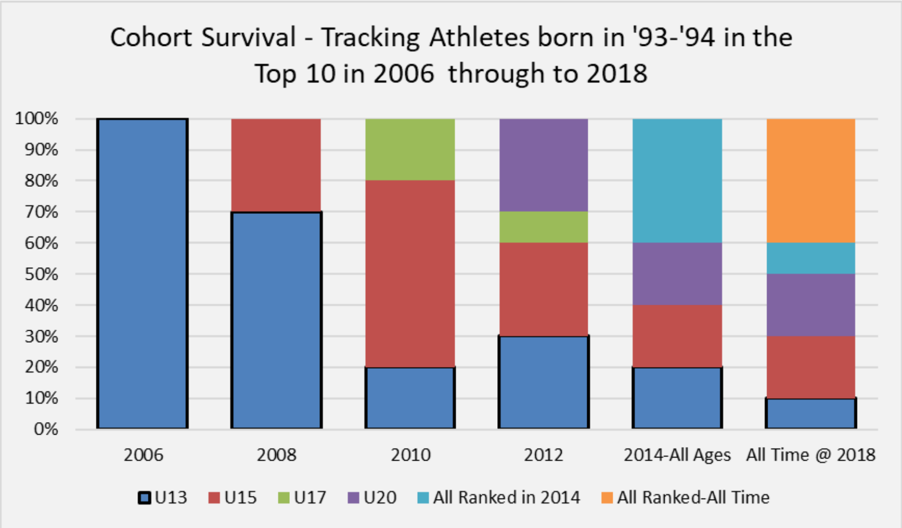


Chart 3

There is a considerable churn of athletes in the top 10, 50 and 100 ranks between the Under 13 in 2006 and Under 20 in 2012. The survival rate of U20 top 10 athletes in 2012 was relatively high through to joining the all age rankings in 2014.
This implies that predicting top athletes from top ranking positions is NOT reliable for athletes in U17 and below. The progress of the U20 cohort studied might be exceptional.

4.3 Relative Age Effect

In the UK sports age group cut-off follows the academic year (eg starts 1st September) until the Under 20 group which starts on 1st January. Per table below England and Wales monthly births do not vary much but there is a seasonal trend with most births in July and the following 3 months.

Table 3: Average births England and Wales 1995 to 2014 (Data Source: Office for National Statistics)

Mth	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
%	8.61	8.60	8.16	8.27	8.30	7.62	8.33	8.04	8.47	8.31	8.72	8.57

However, the distribution of top athlete births varies considerably. The chart below, discussed in the notebook, shows the distribution of birth months for age groups.

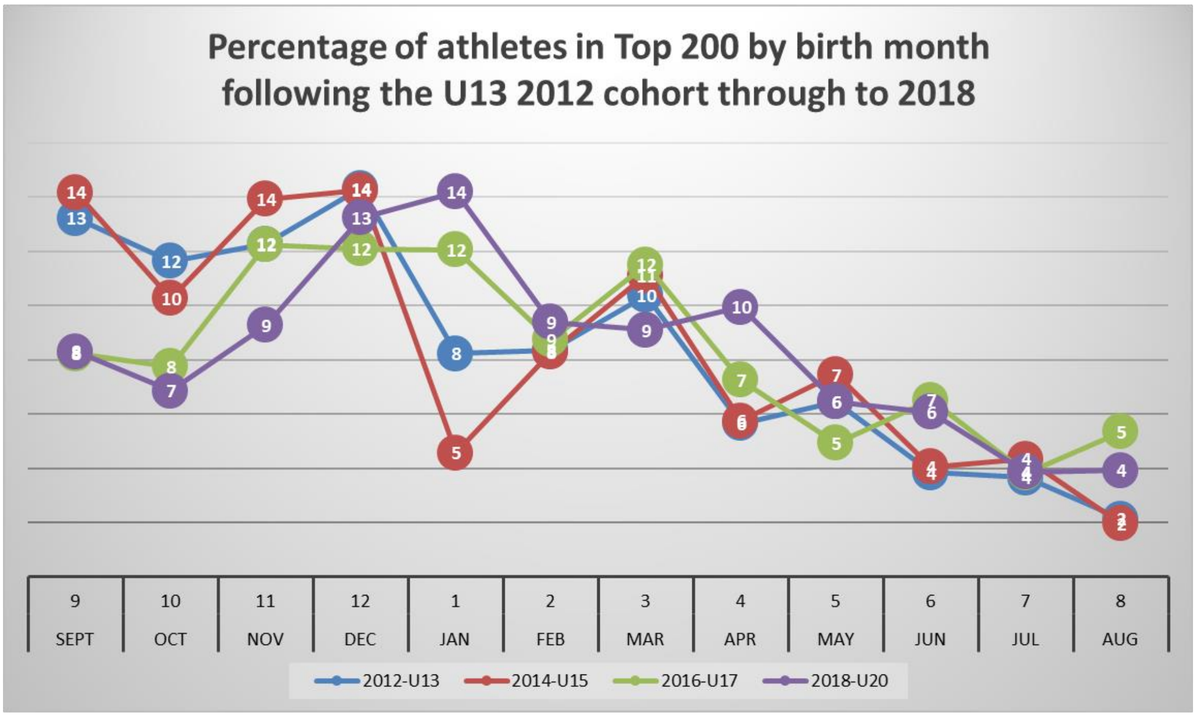


Chart 4

The graph above follows the composition of top 200 U13 cohort from 2012 to 2018.
- September is a "high" month for U13 and U15.
- There is also a peak in in November and December
- Far fewer athletes are born in July.
- But the above variation declines for the older age groups.

A "Relative Age Effect" can be seen with most born top athletes being born in the first 6 months following the age-group cut-off. I decided not to add this as a feature into the current model as the effect decreases with age and a model including this feature might be hard to justify until there is consensus about the effect.

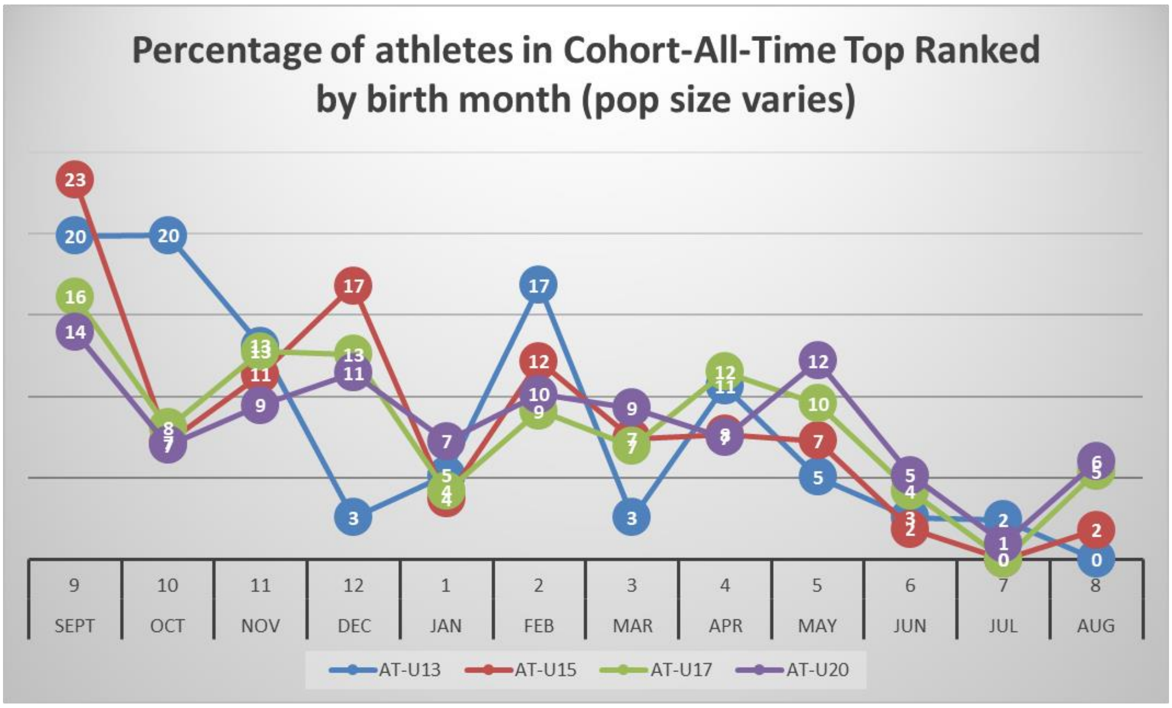


Chart 5 -Other chart from data showing similar effect

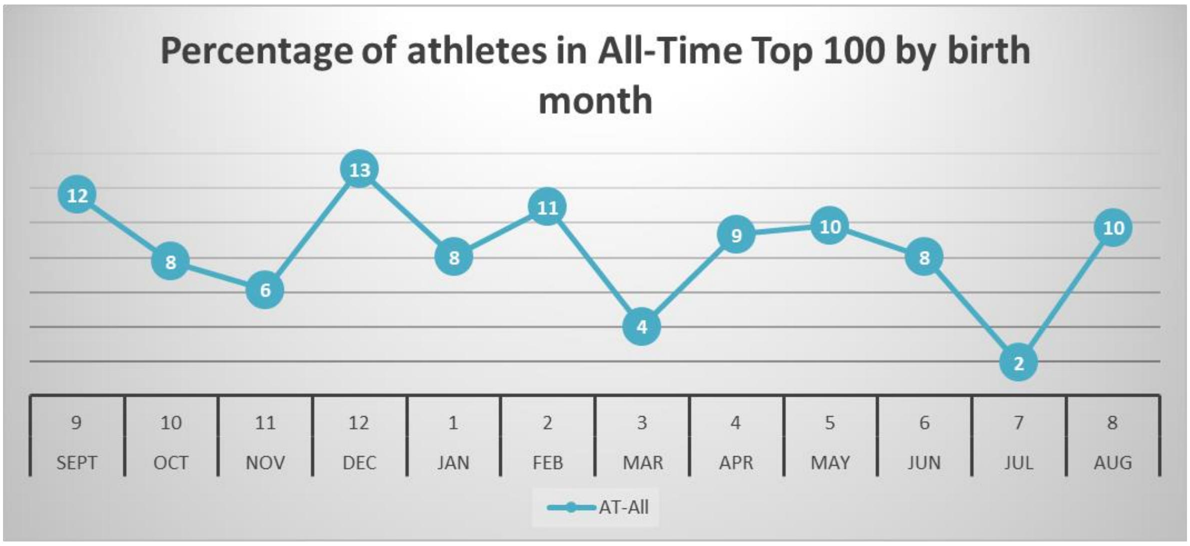


Chart 6 -shows variation decreases with age.

4.4 Modelling Individual Athlete Performance

The results of curve fitting for the top ranked 2015 athletes are shown in Table 4 below. Correlation can be seen visually and by reference to the Spearman coefficients.

		Spearman Correlation						
		0.9410	0.8191	0.4383				
Name	Actual PB Rank @2018	Predicted using data to 2018	Predicted using data to 2016	Predicted using data to 2014	PB Age	No of Events to 2018	No of Events to 2016	No of Events to 2014
Dina Asher-Smith	1	1	2	6	23.3	65	57	51
Jodie Williams	2	2	1	1	20.9	38	30	27
Desiree Henry	3	3	3	4	21.0	49	43	38
Anyika Onuora	4	5	7	10	29.8	35	26	21
Shannon Hylton	5	4	4	32	18.6	34	20	7
Kat J-Thompson	6	7	11	13	22.7	26	22	17
Ama Pipi	7	8	9	7	21.5	49	37	25
Margaret Adeoye	8	6	5	12	28.4	30	19	12
Charlotte McLennaghan	9	13	8	22	17.9	35	21	10
Alisha Rees	10	10	12	28	18.1	58	43	29
Zoey Clark	11	16	25	30	22.6	34	31	18
Hayley Mills	12	11	13	5	23.8	29	19	15
Kimbely Baptiste	13	14	15	16	25.5	39	32	22
Hannah Brier	14	12	6	3	17.3	44	28	13
Sophie Papps	15	9	10	14	19.9	35	30	20
Louise Bloor	16	15	14	11	27.8	34	21	10
Joey Duck	17	17	21	19	26.1	52	39	34
Amy Allcock	18	18	22	27	23.8	52	44	36
Kathryn Christie	19	19	17	8	19.9	57	44	29
Jessica T-Jemmett	20	24	30	31	29.1	59	52	45
Lorraine Ugen	21	21	20	26	23.6	16	13	12
Stacey Downie	22	28	29	9	24.1	32	22	10
Charlotte Wingfield	23	23	26	18	22.5	117	99	82
Hannah Thomas	24	33	31	21	20.4	34	24	10
Mica Moore	25	26	28	17	21.5	32	24	19
Lucy Evans	26	30	32	24	31.7	37	27	17
Laura Wake	27	22	19	15	22.2	47	39	34
Susanna Banjo	28	25	27	20	25.1	42	32	21
Ella Barrett	29	20	16	23	17.2	35	20	9
Amelia Reynolds	30	29	18	25	18.3	42	32	18
Rachel Dickens	31	27	23	2	19.5	28	24	12
Jessica Tappin	32	34	34	33	24.0	28	21	19
Rachel Norris	33	31	33	34	16.6	34	22	10
Abigayle Fitzpatrick	34	32	24	29	23.5	22	15	11

Table 4

The results from this project show that curve-fitting could produce reasonable predictions 2 years out (Spearman coefficient of 0.82) but less well 4 years out (0.44). Though these results could be improved by model refinement and using the data more effectively e.g. taking the best results per individual competition rather than per month.

4.5 Reflections and Future Work

1. The age for peak speed varies widely between athletes. Applying an average peak age to each athlete would be an unreliable predictor.

2. There is considerable turnover in the ranks below the under 20 age group. Simply picking fast sprinters at age 13 is unreliable.

3. Some top rank athletes do not compete at 200m every year or appear consistent. On reflection this is probably due to their “main” event being 100m which traditionally has a higher status compared to 200m. The interaction with other sporting events has not been considered but it might explain additional variance in results in a similar way to heats in large competitions (eg an athlete may not perform so well at 200m because they have a 100m race in proximity). This effect could be factored in with further work.

4. Some of the churn seen at lower ranks maybe due to athletes changing their main event. Further analysis of events could test this.

5. The average age of athletes is getting older and careers longer, this may be due to greater incentives to remain compete due to increased support compared to, say, 50 years ago.

6. The Relative Age effect is a potential extra feature, but further work would be required to test its significance or even whether any relationship is causal. Another theory posits that children in the womb during summer months benefit from extra vitamin D [mentioned in Refs 5 & 6]. Further research on southern hemisphere born athletes could be used to test this.

7. Whilst there is a Relative Age Effect extra consideration could be given to summer born children when selecting teams to prevent demotivation. For example, team selection in large schools could be broken down into 6-month age groups.

8. Berthelot et al published similar modelling results with greater R-squared correlation (0.99 compared to 0.94 using data to 2018) – though they used the an average of just 6 events (1 per year) per athlete compared to this analysis using the best event per month.

9. The curve fitting model is promising but needs refining to improve predictive powers over a longer term to be useful and predict speeds past an athlete’s peak. Further work could test simpler formulae for the declining component, but a current problem is lack of data to test any formula.

4.6 Potential business and scientific value

1. Selecting athletes by an algorithm or machine learning is a worthwhile goal. If it worked it would be a cheap, transparent, unbiased and hence fair process. This would free up the cost of selection committees to be spent on support for a greater number of athletes.

2. Monitoring performance might also identify potential drug cheats for increased monitoring.

3. Athlete data is publicly accessible. A possible use could be in a motivational app: athletes could key in their results to see how they compare to their peers and to elite athletes when they were the same age. This information could also motivate summer born children who could compare their result to averages for their age rather than their relatively older school year.

5 References (Exclude from word-count)

1 - Exponential growth combined with exponential decline explains lifetime performance evolution in individual and human species. Berthelot, G., Len, S., Hellard, P. et al. AGE (2012) 34: 1001. https://0-doi-org.wam.city.ac.uk/10.1007/s11357-011-9274-9

2-Hoye R, Smith A, Nicholson M, Stewart B, Westerbeek H (2009) Sport management: principles and applications, 2nd Edition. Oxford, Elsevier Butterworth-Heinemann
3-Funding details - http://www.uksport.gov.uk/sports/olympic/athletics

4-Grondin S, Deschaies P, Nault LP. Trimesters of birth and school output [in French]. Apprent Social 1984; 16: 169-74

5-Annual Age-Grouping and Athlete Development - A Meta-Analytical Review of Relative Age Effects in Sport Cobley, S., Baker, J., Wattie, N. et al. Sports Med (2009) 39: 235. https://0-doi-org.wam.city.ac.uk/10.2165/00007256-200939030-00005

6-(19.12.2014). "New Sports Medicine Data Have Been Reported by Researchers at University of British Columbia (Athletic Performance and Birth Month: Is the Relative Age Effect More than just Selection Bias?)". Health & medicine week (1531-6459), p. 3389.

7-How popular is your birthday? Office for National Statistics https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/articles/howpopularisyourbirthday/2015-12-18

8-N.U.T.S. National Union of Track Statisticians https://www.nuts.org.uk/

9-The Association of GB Athletics Clubs Fact Files http://www.britishathleticsclubs.com/?page_id=269
Fact File 61 -Drop Out Rates

6 APPENDICES (Exclude from word count)

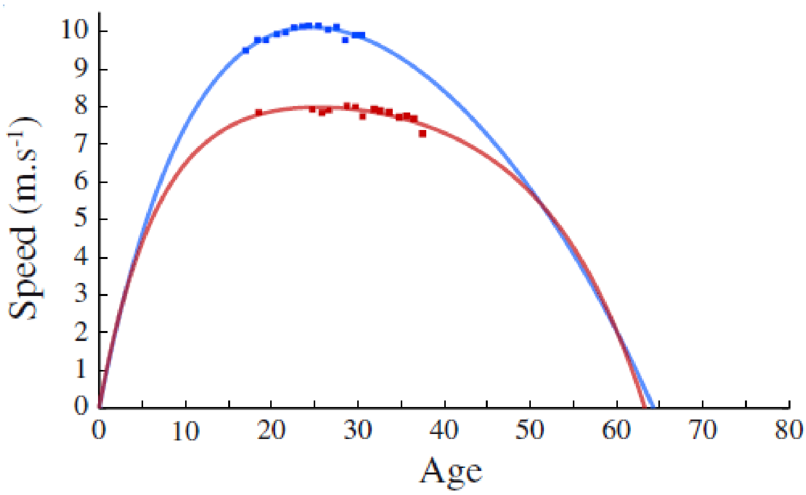
6.1 Extracts from Berthelot's paper

6.1.1 Data

For each year of an athlete's career, when several performances were established, only the best one was kept. The average number of performance points per career was 6.20±1.36 in track and field

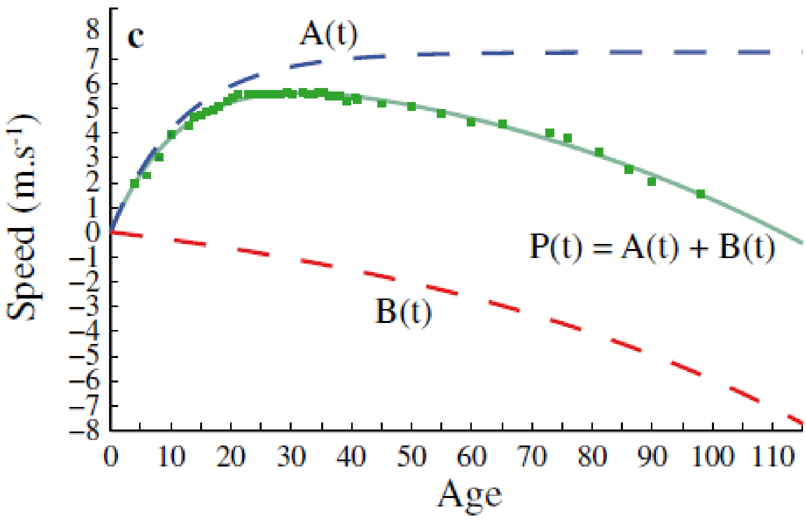
6.1.2 Individual Performance

Fig. 1 The model applied at the individual scale (athletes' and chess players' careers). The model is adjusted to two careers in two track and field events: the 100 m straight (blue men career: Ato Boldon; adjusted R^2=0.99 and peak=24.63 years old and the 400 m in track & field (red women career: Sandie Richards; R^2= 0.99 and peak=25.37).



6.1.3 Species plotting

The model applied at the species scale. For each age, the maximum performance among the studied careers is gathered. Fig C …The marathon event (men) is fitted (R^2=0.99 and peak=31.61).



6.1.4 Moore's Formula:

$$P(t) = a \times (1-e^{bt}) + c \times (1-e^{dt})$$

6.2 Overall Athlete Drop Out Rates

Below are figures of competing women from the Tops-in-athletics database. It shows a high turnover of club membership with more than 50% leaving after 2 years and nearly 95% after 10 years.

	Start	By 2016	Lost	Dropout %
1 Year Dropout -2015:	9,146	5,663	3,483	**38.08%**
2 Year Dropout -2014:	12,452	5,659	6,793	**54.55%**
3 Year Dropout -2013:	14,735	4,782	9,953	**67.55%**

Women Under 13 to Under 20 in "topsinathletics" database in 2006

	2006	2012	2014	2016
Cohort in DB	14,530	1,912	1,294	825
Lost from -2006		12,618	13,236	13,705
Dropout %		**86.84%**	**91.09%**	**94.32%**

Source: The Association of GB Athletics Clubs Fact Files - Fact File 61 -Drop Out Rates http://www.britishathleticsclubs.com/?page_id=269

6.3 Files exported from notebook used to summarise data

1-Rank_ExtractToSort – Cohort survival charts

2-BirthMonth-Review – Birth-month graphs

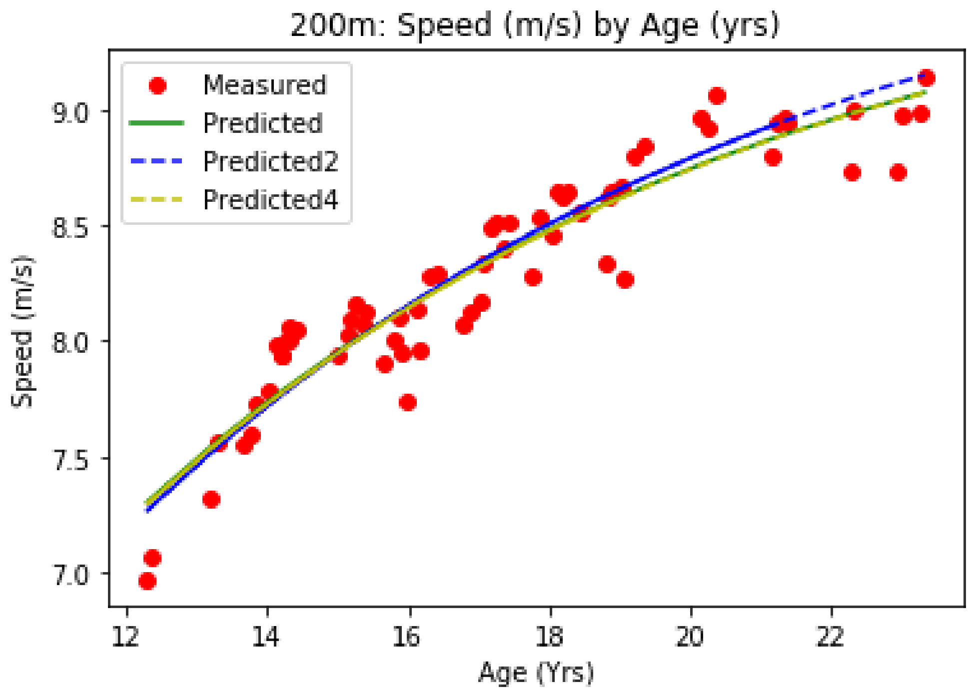
3-EDA-01-Analysis - table for All-time averages

4-OverallGraph – All entries for "species" graph

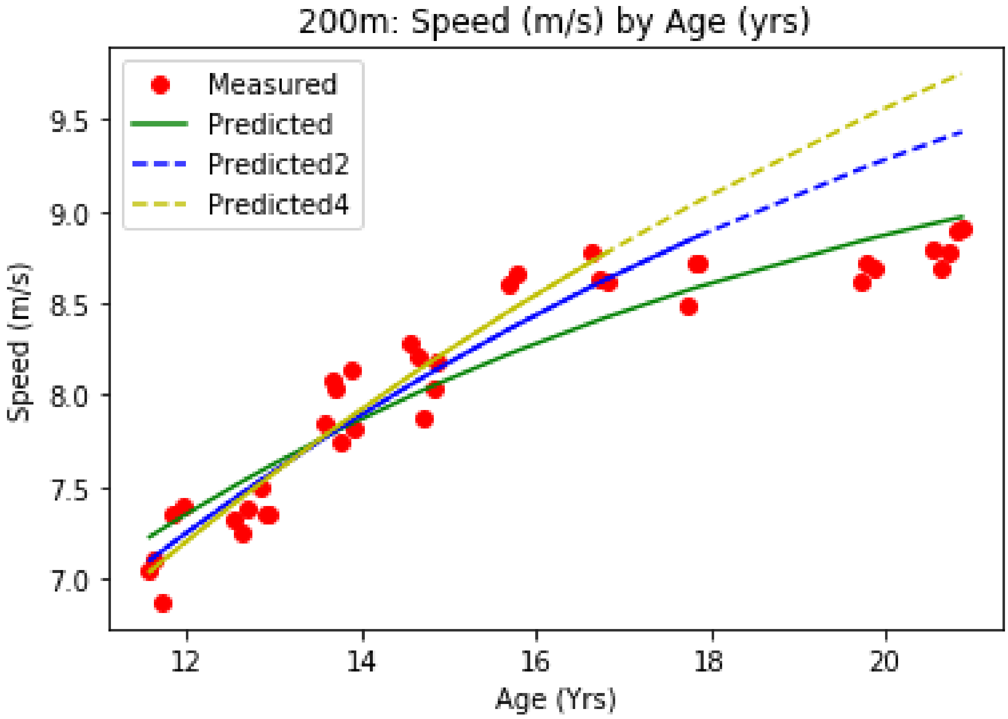
5-Ath_curves2014-Review – Output table for Spearman Ranking

6.4 Example Curve-fitting Output

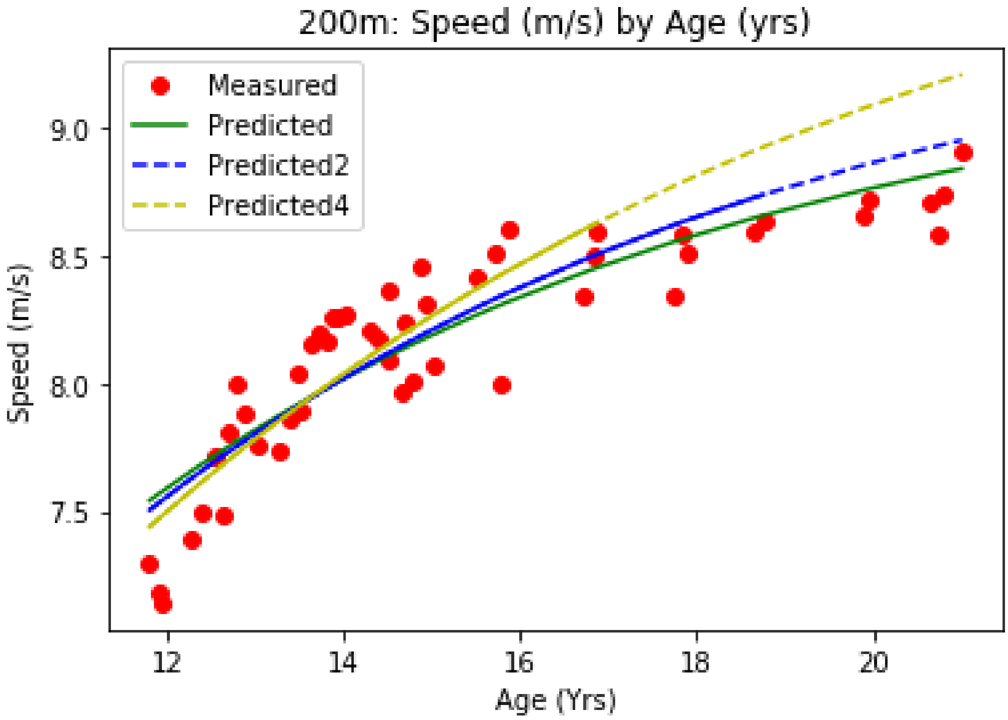
```
46473 Dina Asher-Smith
Optimal parameters
[ 9.80687872 -0.11108384]
R^2: 0.889667496139664
[10.00272974 -0.1054579 ]
R^2_2yr:  0.884733814756173
[ 9.81403588 -0.11063626]
R^2_4yr:  0.8895399118025665
```



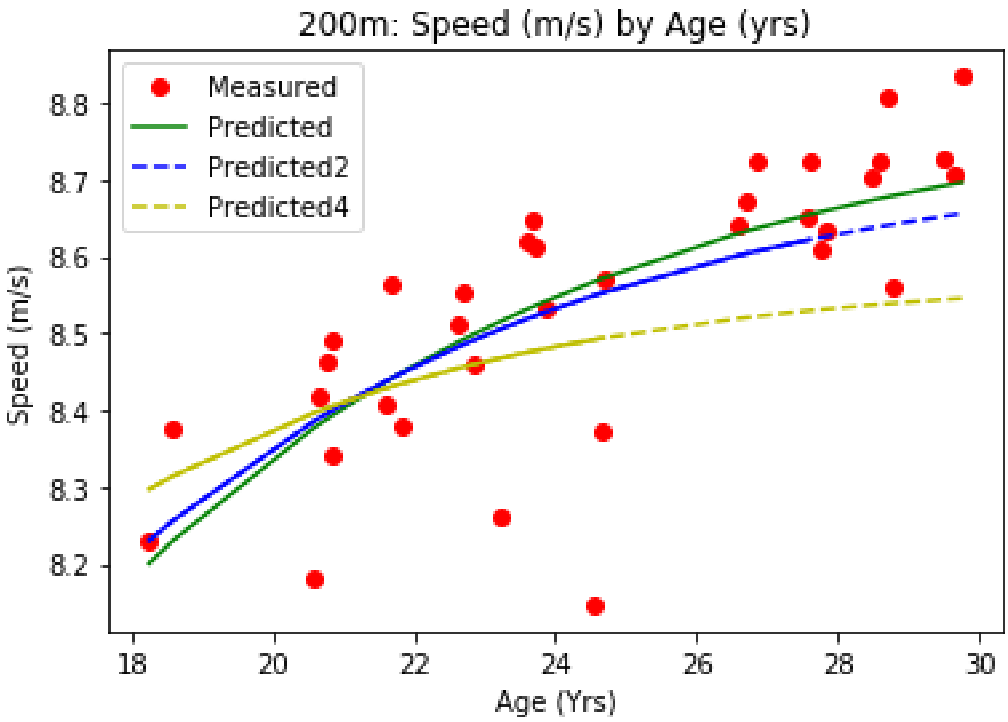
```
26161 Jodie Williams
Optimal parameters
[ 9.90178378 -0.11311771]
R^2: 0.8818954526143798
[11.38349532 -0.08446177]
R^2_2yr:  0.7399333478257812
[12.72310282 -0.06961966]
R^2_4yr:  0.48319093471544017
```



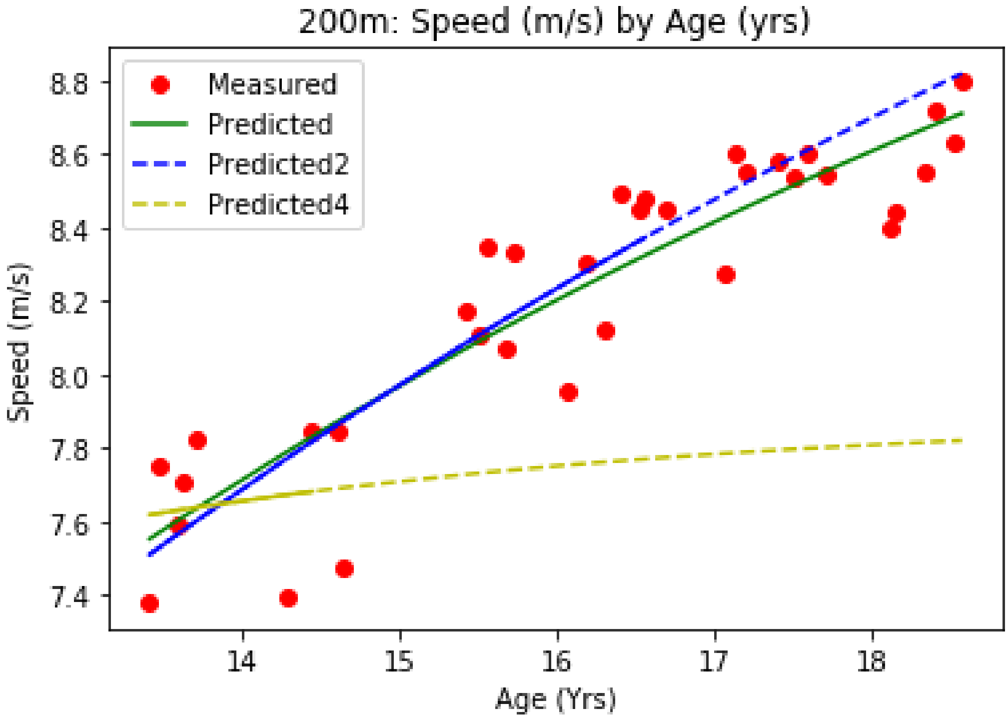
```
46405 Desiree Henry
Optimal parameters
[ 9.32997769 -0.13995001]
R^2: 0.8102598975745361
[ 9.57614121 -0.12968545]
R^2_2yr:  0.7968254790514515
[10.2010208  -0.11067287]
R^2_4yr:  0.6666030271544032
```



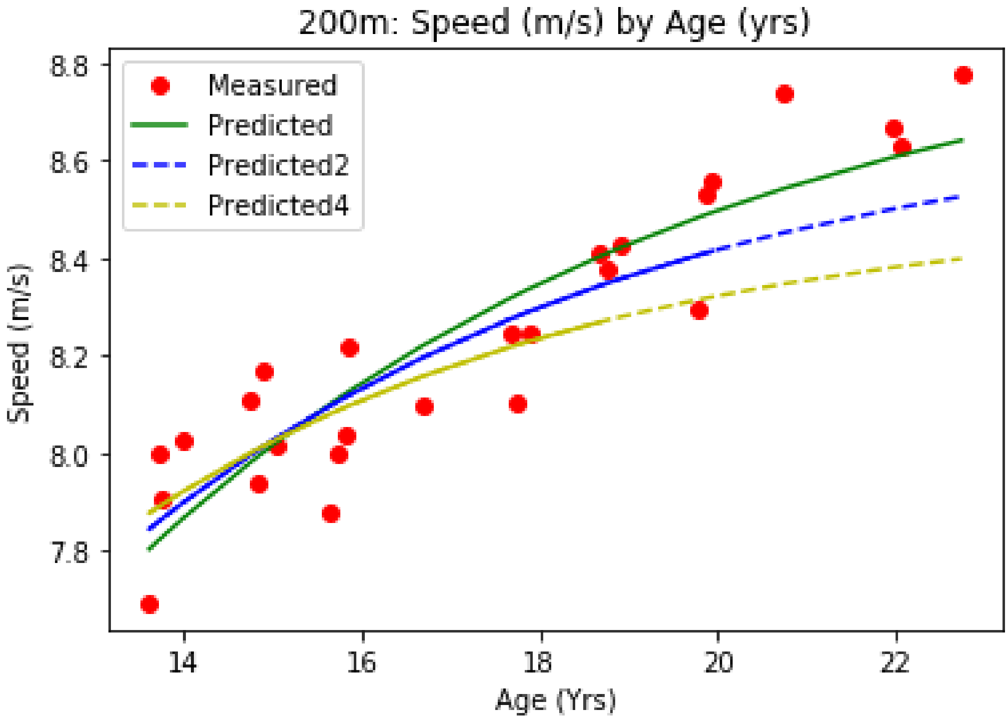
```
21755 Anyika Onuora
Optimal parameters
[ 8.80832005 -0.14654958]
R^2: 0.5369604875206142
[ 8.74024204 -0.15569449]
R^2_2yr:  0.5173583466099984
[ 8.57789646 -0.18736919]
R^2_4yr:  0.2562313651019089
```



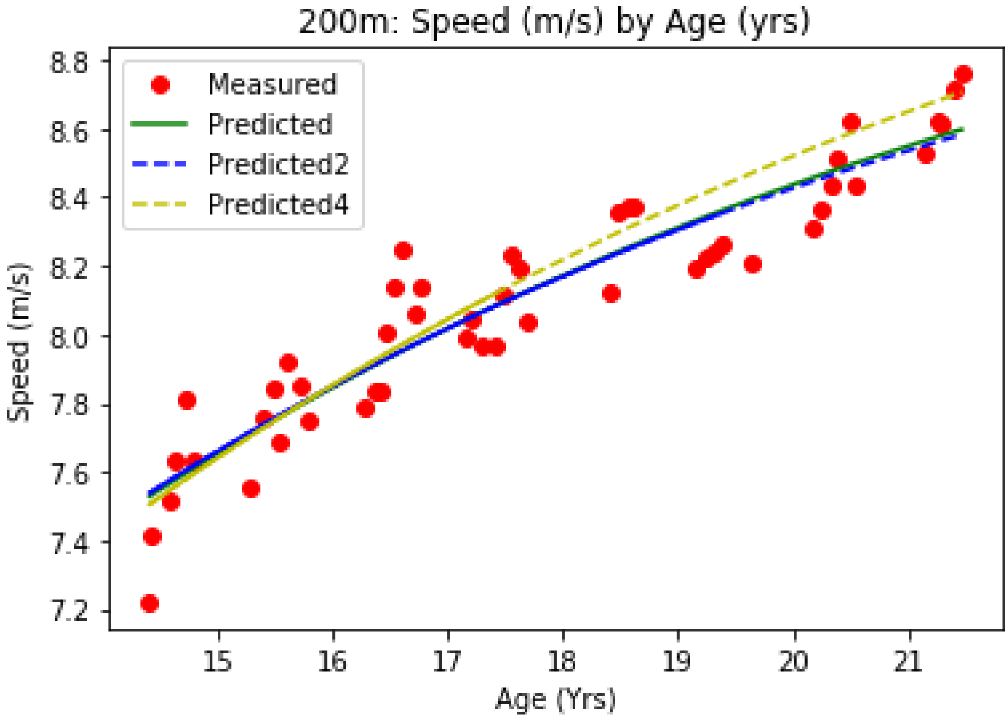
```
72047 Shannon Hylton
Optimal parameters
[10.58122878 -0.09330523]
R^2: 0.8288021222542516
[11.33777483 -0.08099484]
R^2_2yr:  0.8058381093201751
[ 7.89684222 -0.24935309]
R^2_4yr:  -1.14641752408509
```



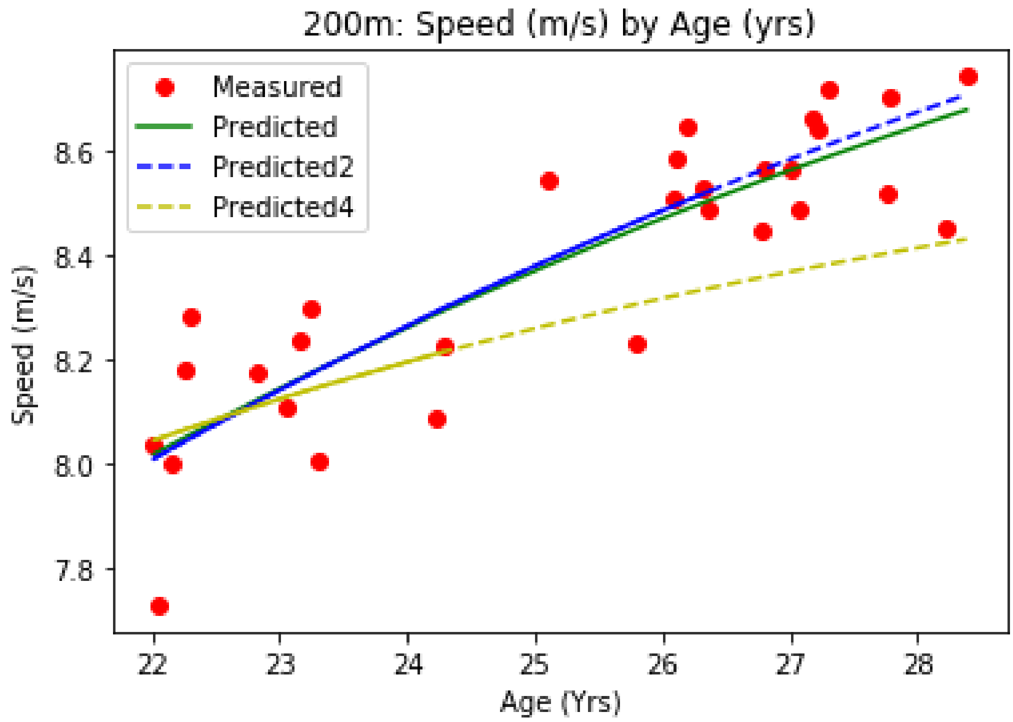
```
21167 Katarina Johnson-Thompson
Optimal parameters
[ 8.91783255 -0.1527176 ]
R^2: 0.8178799744988443
[ 8.71084905 -0.16946619]
R^2_2yr:  0.7770166575541602
[ 8.50914579 -0.19091749]
R^2_4yr:  0.6287552551358889
```



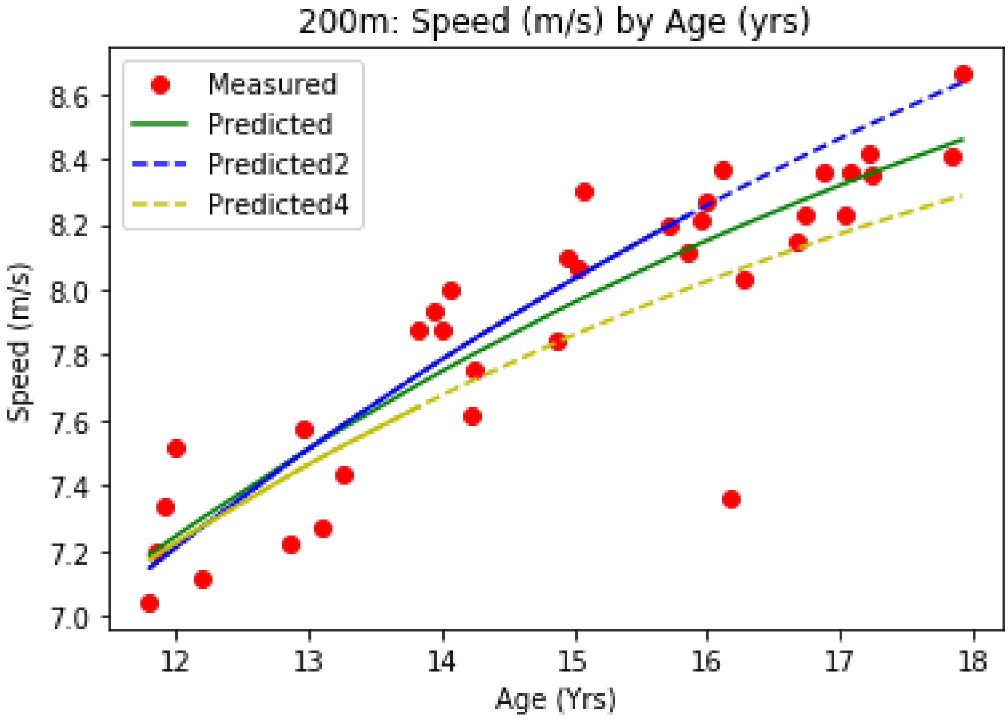
```
47196 Ama Pipi
Optimal parameters
[ 9.53223134 -0.10821131]
R^2: 0.8789887275417857
[ 9.47087814 -0.11028141]
R^2_2yr:  0.8784688591974964
[ 9.90767045 -0.09830843]
R^2_4yr:  0.8524332033231916
```



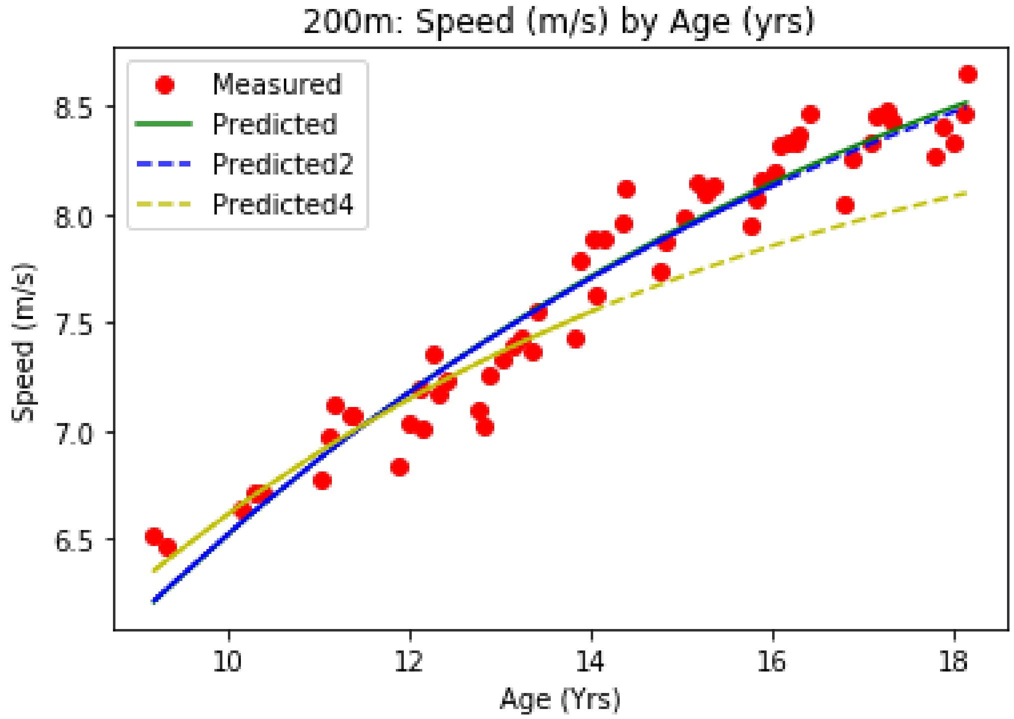
```
45663 Margaret Adeoye
Optimal parameters
[ 9.6595177  -0.08053713]
R^2: 0.7536946649310113
[ 9.80225229 -0.07718603]
R^2_2yr:  0.7494857870897873
[ 8.8006457  -0.11156698]
R^2_4yr:  0.40633905816652516
```



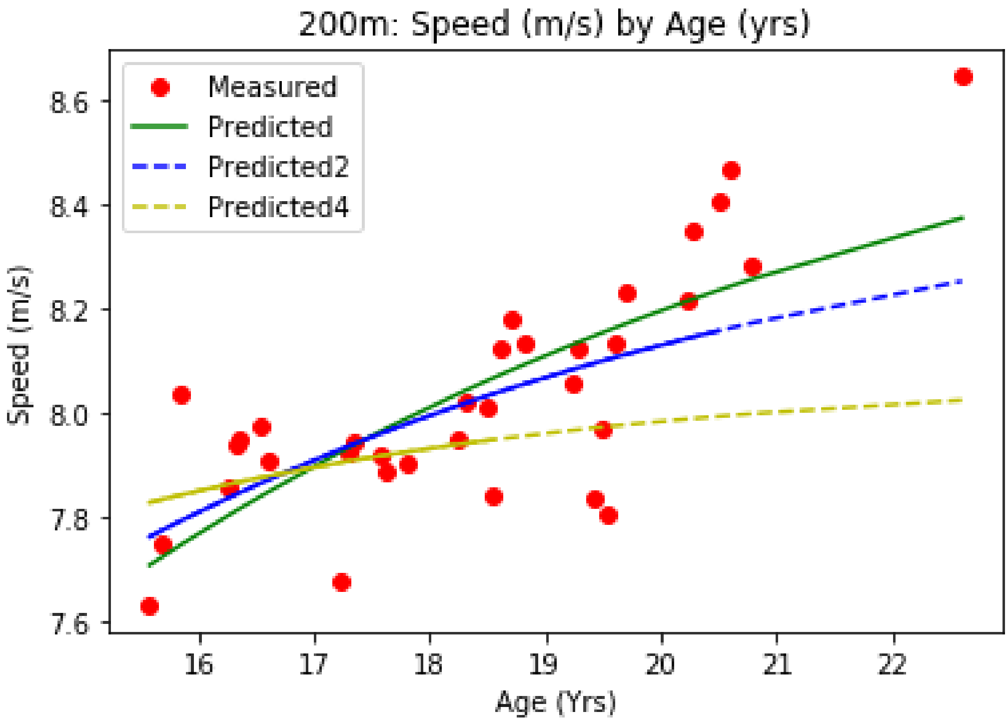
```
46335 Charlotte McLennaghan
Optimal parameters
[ 9.72124966 -0.1139277 ]
R^2: 0.7817412327002776
[10.46273079 -0.09735184]
R^2_2yr:  0.7337915897311793
[ 9.23999309 -0.12685112]
R^2_4yr:  0.7215625243066668
```



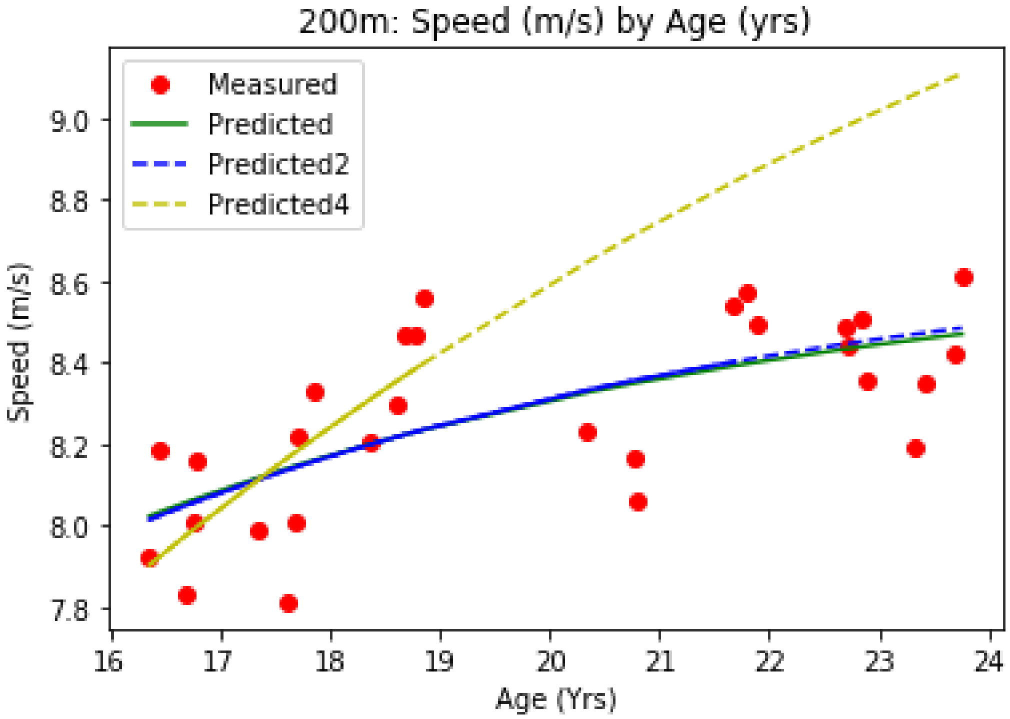
```
85759 Alisha Rees
Optimal parameters
[ 9.97984852 -0.10595    ]
R^2: 0.9332740770690924
[ 9.90661344 -0.10745744]
R^2_2yr:  0.9328215821449853
[ 8.80291307 -0.13929413]
R^2_4yr:  0.7902521814288803
```



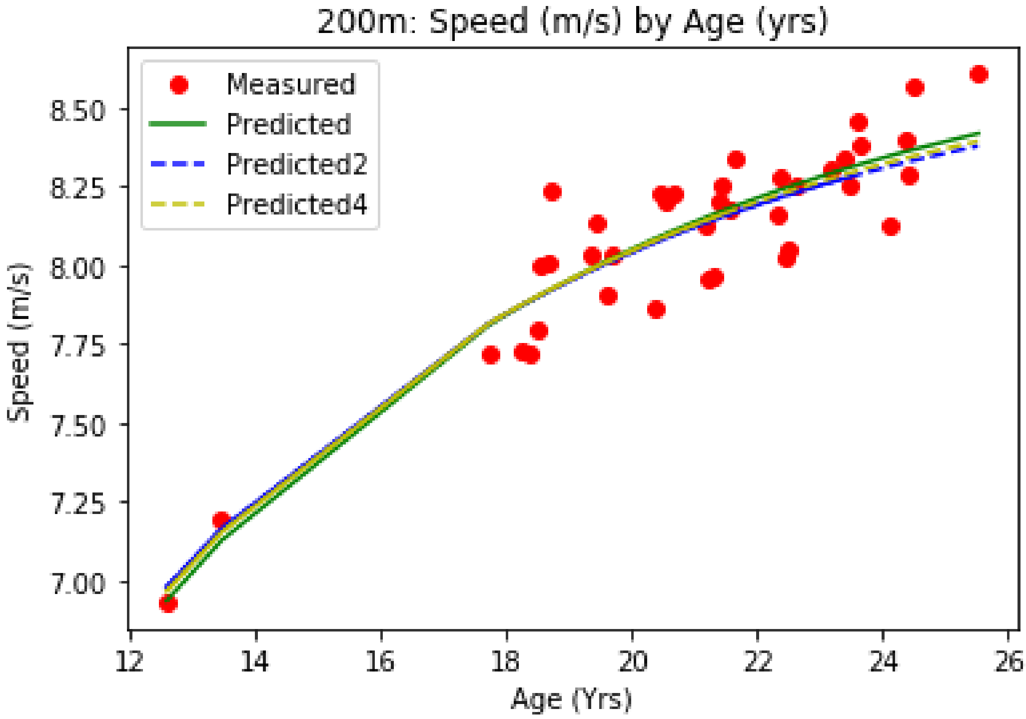
```
86023 Zoey Clark
Optimal parameters
[ 8.79346891 -0.13440341]
R^2: 0.5503430820435296
[ 8.49206225 -0.15768794]
R^2_2yr:  0.5013625060813213
[ 8.07545959 -0.22406761]
R^2_4yr:  0.08245599008686277
```



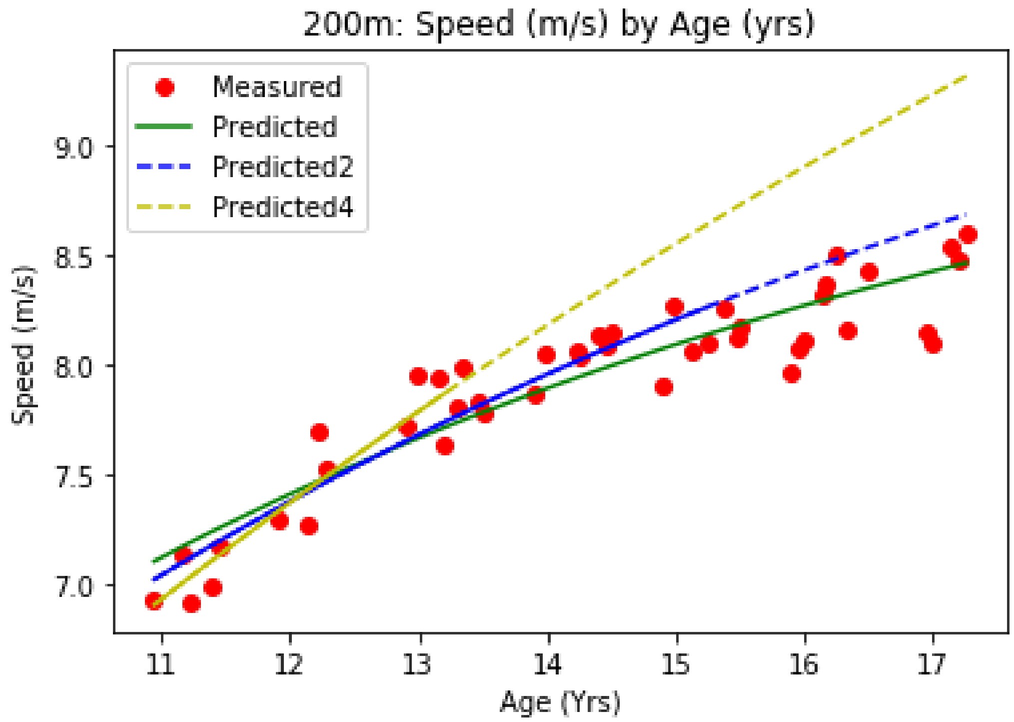
```
19824 Hayley Mills
Optimal parameters
[ 8.66995242 -0.15870073]
R^2: 0.45612728458006957
[ 8.70400363 -0.15500426]
R^2_2yr:  0.4545908974978995
[10.46372575 -0.08605161]
R^2_4yr:  -2.176107108407836
```



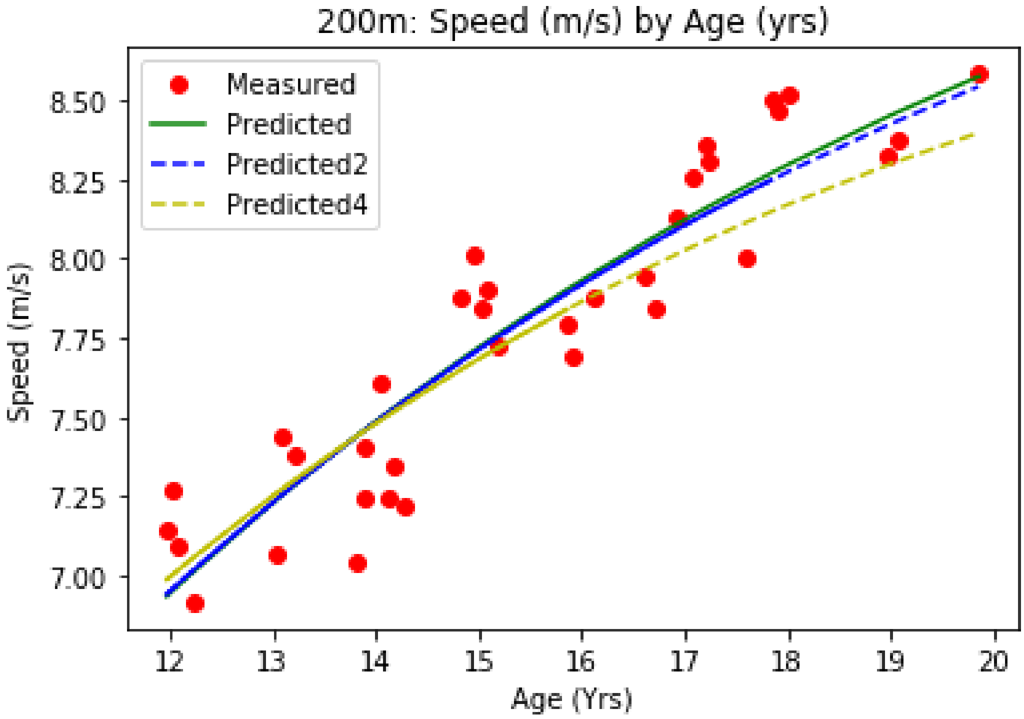
```
49196 Kimbely Baptiste
Optimal parameters
[ 8.78897311 -0.12382347]
R^2: 0.833040299136086
[ 8.70134826 -0.12887997]
R^2_2yr:  0.8280689129567648
[ 8.73222041 -0.12715174]
R^2_4yr:  0.8311750334674058
```



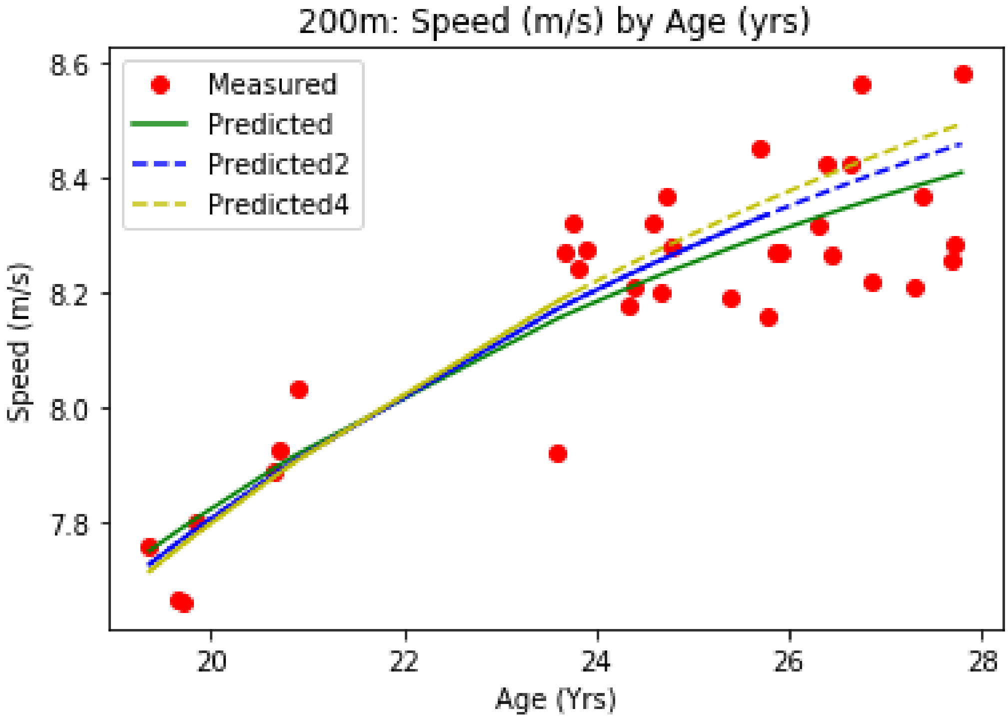
```
74623 Hannah Brier
Optimal parameters
[ 9.62508265 -0.12249794]
R^2: 0.8654514004394253
[10.57636593 -0.09968105]
R^2_2yr:  0.7854945944045739
[14.8206077  -0.05734646]
R^2_4yr:  -0.3482349321541083
```



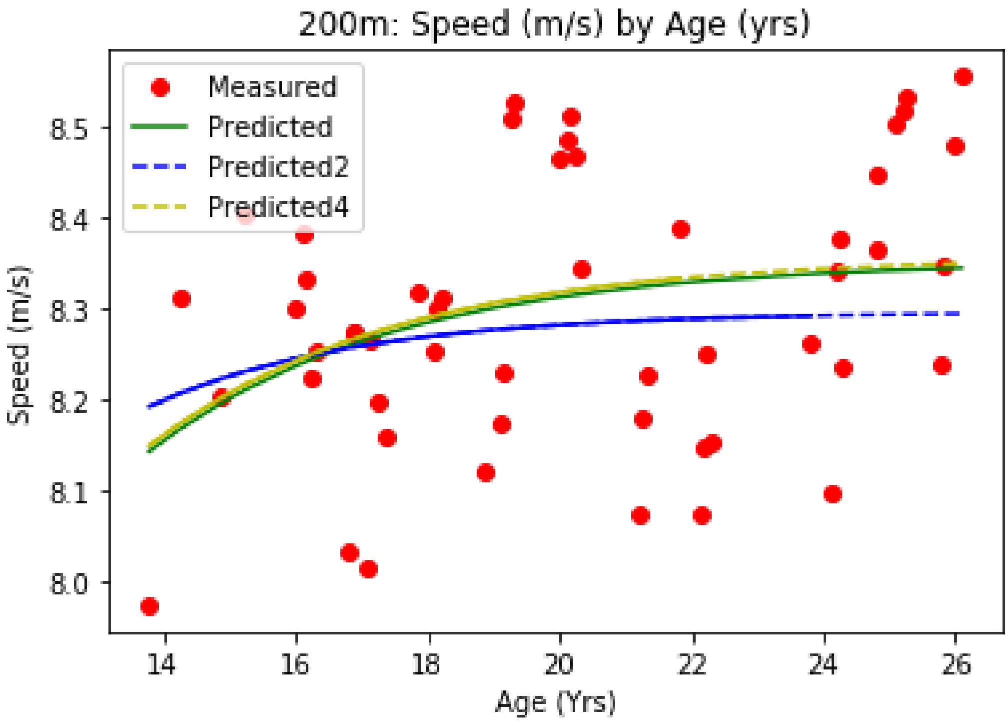
```
27446 Sophie Papps
Optimal parameters
[ 9.94437308 -0.09982665]
R^2: 0.8499226388201775
[ 9.83617908 -0.10218363]
R^2_2yr:  0.8490709544726657
[ 9.35485222 -0.11478141]
R^2_4yr:  0.8233747225114022
```



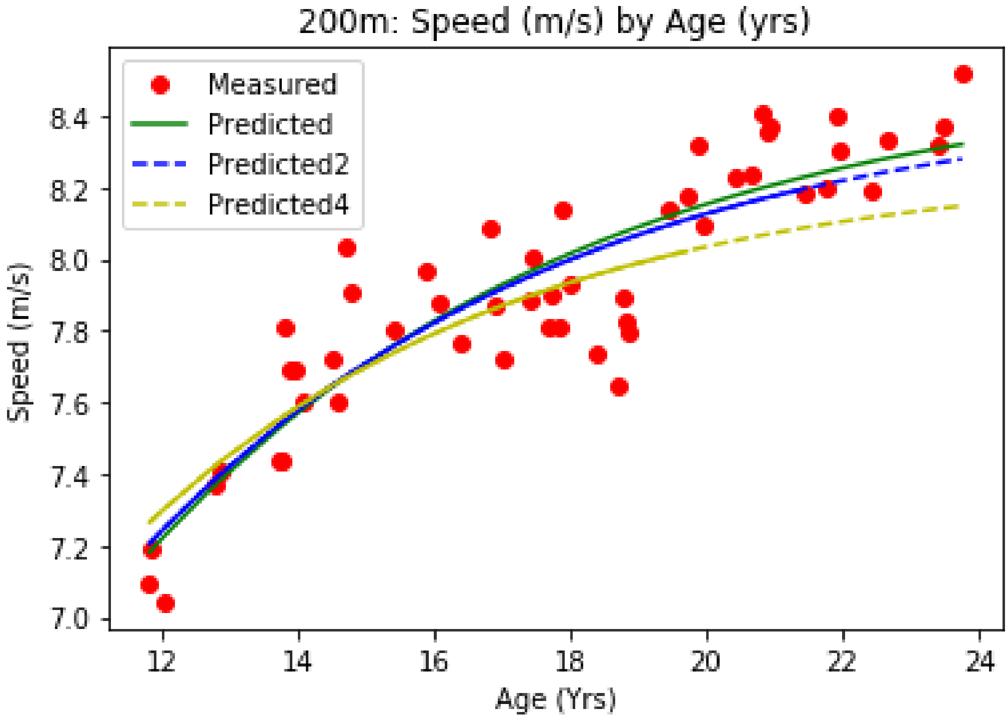
```
19903 Louise Bloor
Optimal parameters
[ 8.85467419 -0.10746578]
R^2: 0.768351035496432
[ 9.00413095 -0.10082752]
R^2_2yr:  0.7486558793252639
[ 9.10955147 -0.09687739]
R^2_4yr:  0.711220500984824
```



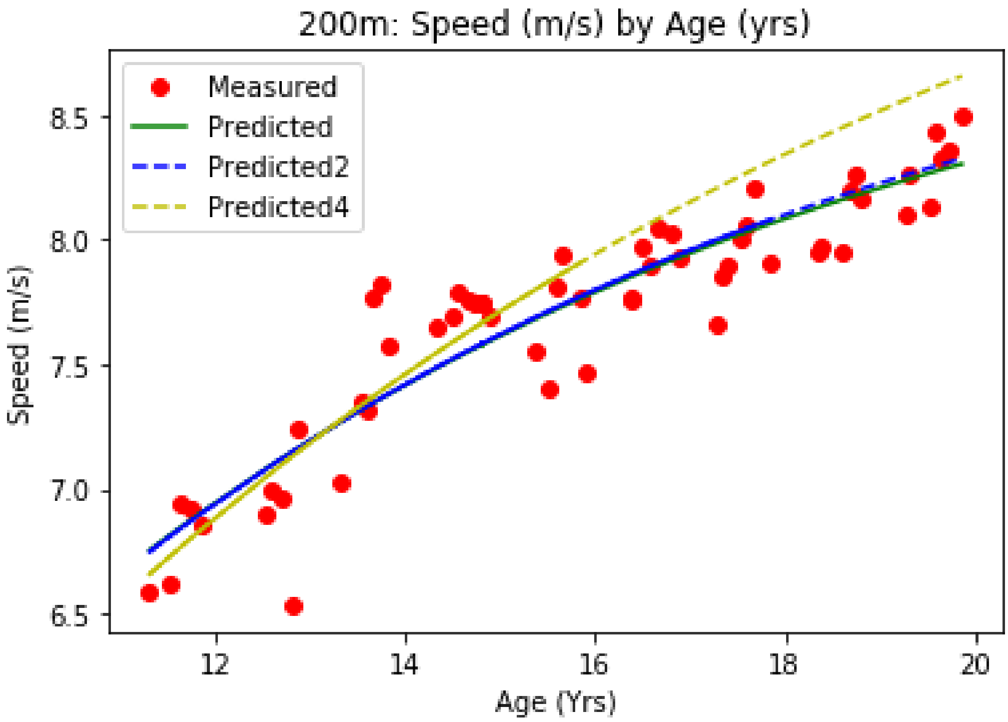
```
20550 Joey Duck
Optimal parameters
[ 8.35192877 -0.26758912]
R^2: 0.1011846707337476
[ 8.29596464 -0.31780543]
R^2_2yr:  0.04837256832428538
[ 8.35676994 -0.26805371]
R^2_4yr:  0.09988894202065857
```



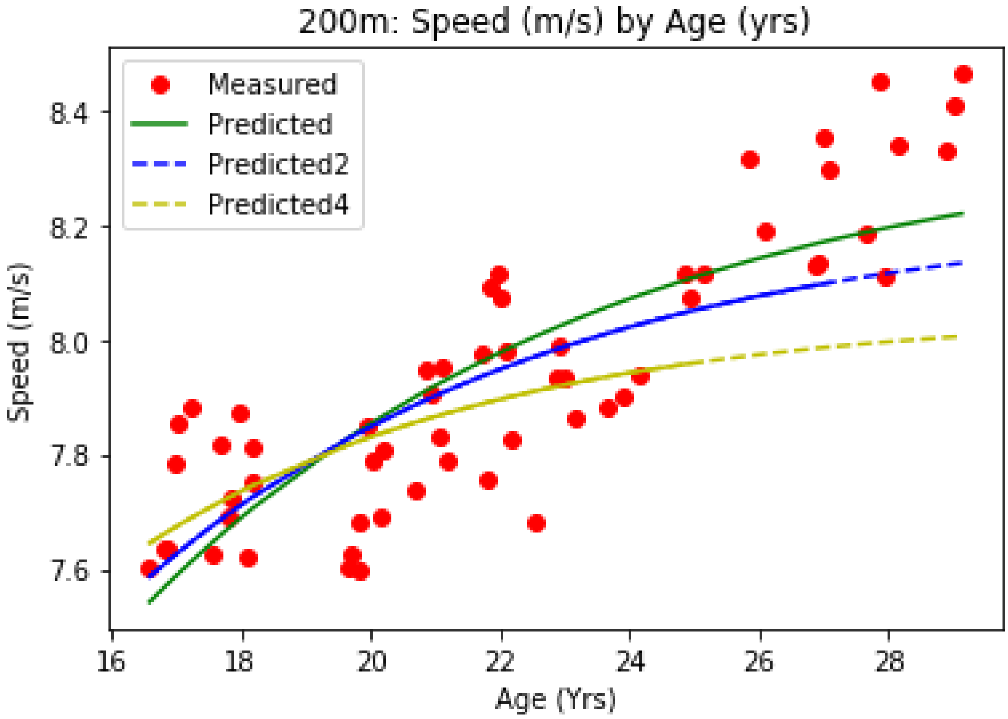
```
17090 Amy Allcock
Optimal parameters
[ 8.53102334 -0.15590737]
R^2: 0.8029357819559273
[ 8.46358373 -0.161167  ]
R^2_2yr:  0.7987227121389862
[ 8.26841178 -0.17829065]
R^2_4yr:  0.7261891649472566
```



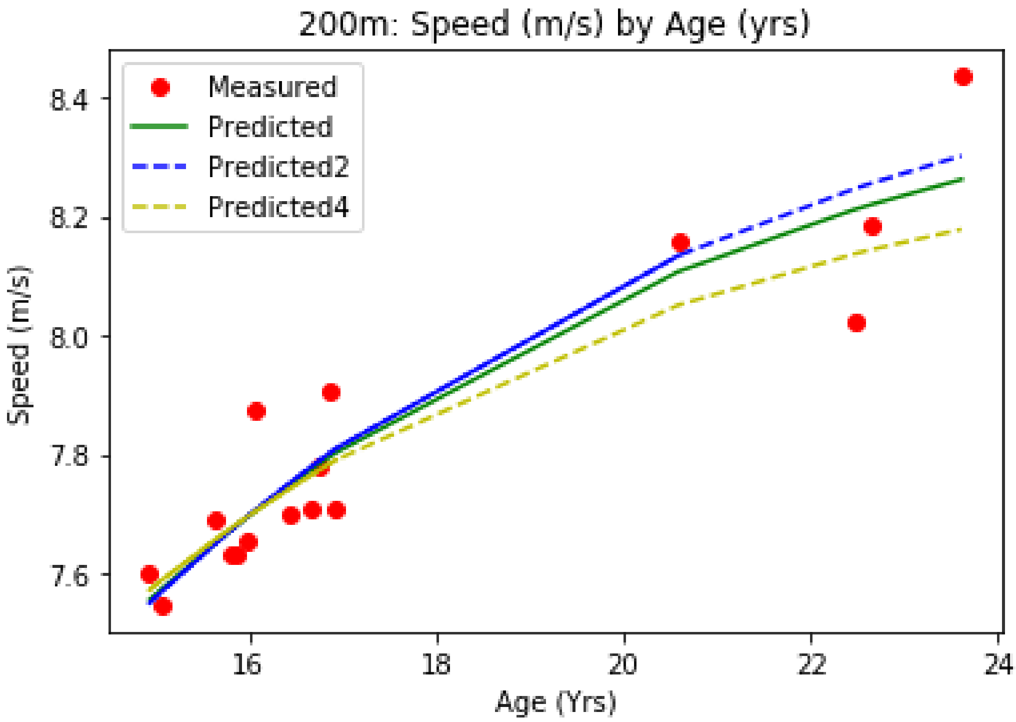
```
20880 Kathryn Christie
Optimal parameters
[ 9.2037759  -0.11708867]
R^2: 0.84736322204872
[ 9.26440545 -0.11528419]
R^2_2yr:  0.8467364342086983
[10.34626918 -0.09122963]
R^2_4yr:  0.6855991102709819
```



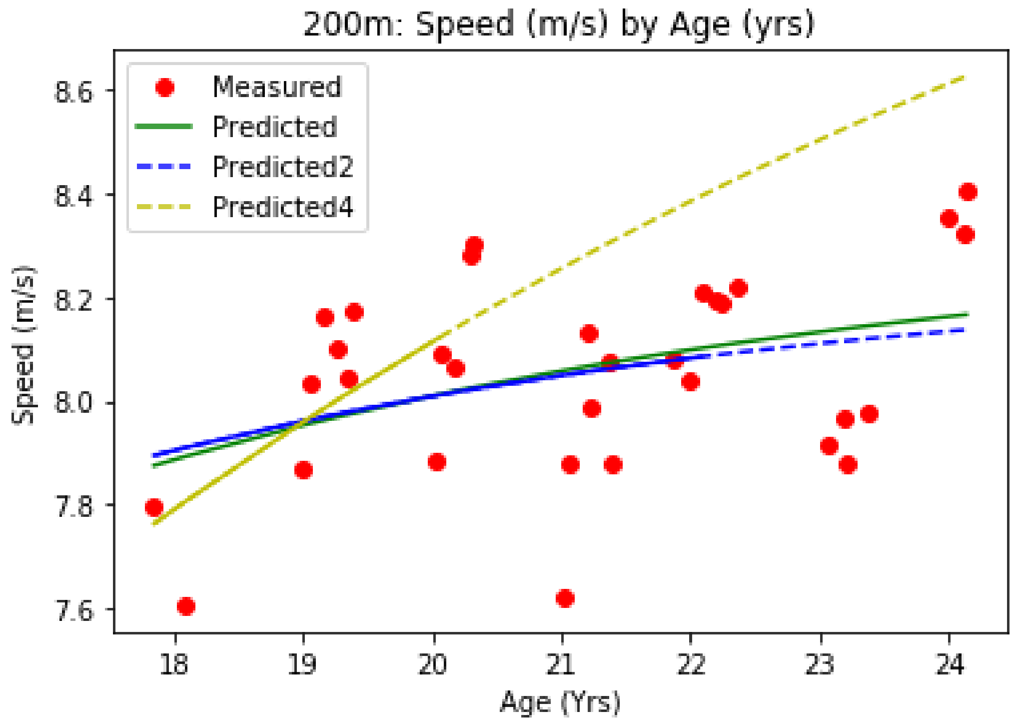
```
20375 Jessica Taylor-Jemmett
Optimal parameters
[ 8.36113571 -0.14001385]
R^2: 0.643768819733837
[ 8.22654206 -0.15387252]
R^2_2yr:  0.6082604929486
[ 8.04760224 -0.18046793]
R^2_4yr:  0.41572072165259
```



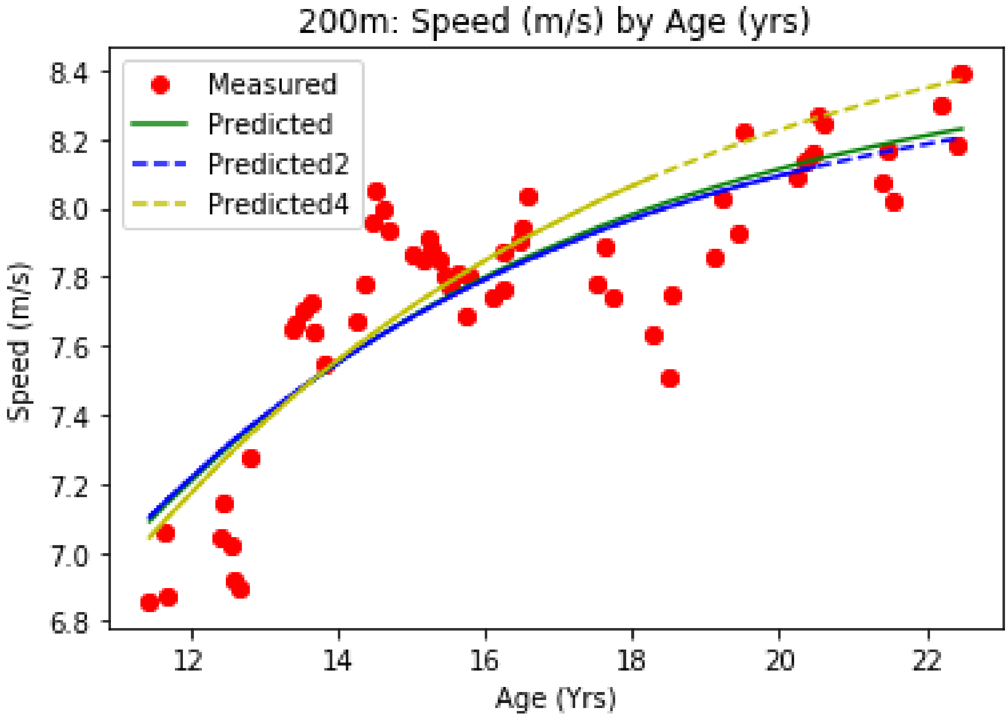
```
37392 Lorraine Ugen
Optimal parameters
[ 8.54469407 -0.14435882]
R^2: 0.8576772747886028
[ 8.61990341 -0.13968615]
R^2_2yr:  0.8524520325028944
[ 8.38929961 -0.15592498]
R^2_4yr:  0.8341859410443128
```



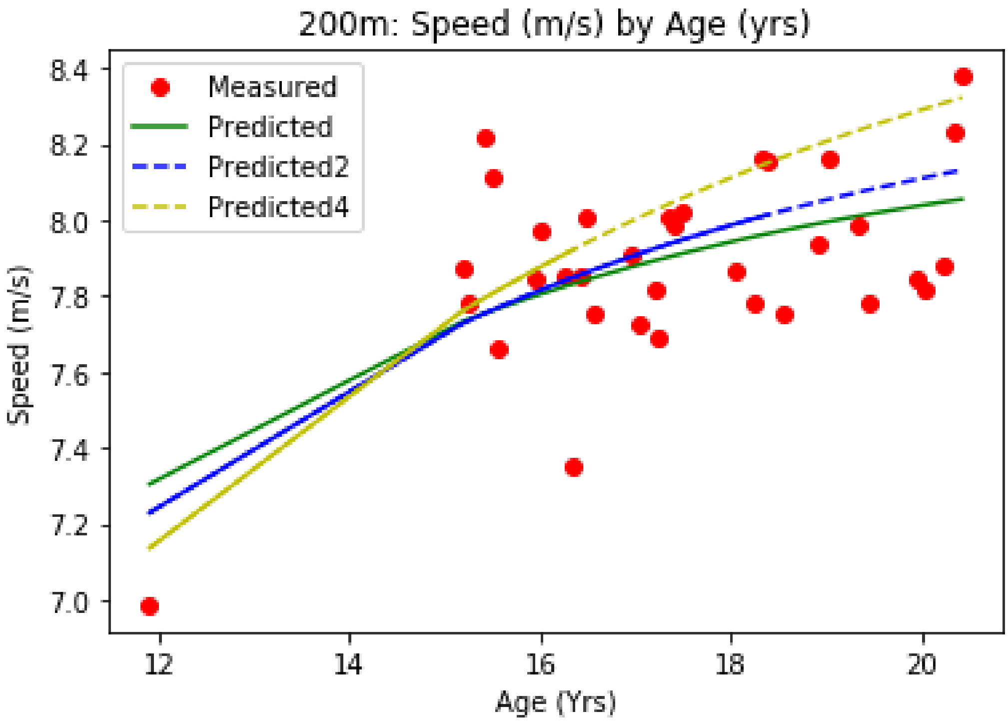
```
20332 Stacey Downie
Optimal parameters
[ 8.32727415 -0.16326824]
R^2: 0.1682978948261067
[ 8.25672405 -0.17524449]
R^2_2yr:  0.1614164717303327
[ 9.7780138 -0.0885258]
R^2_4yr:  -1.6596225760972199
```



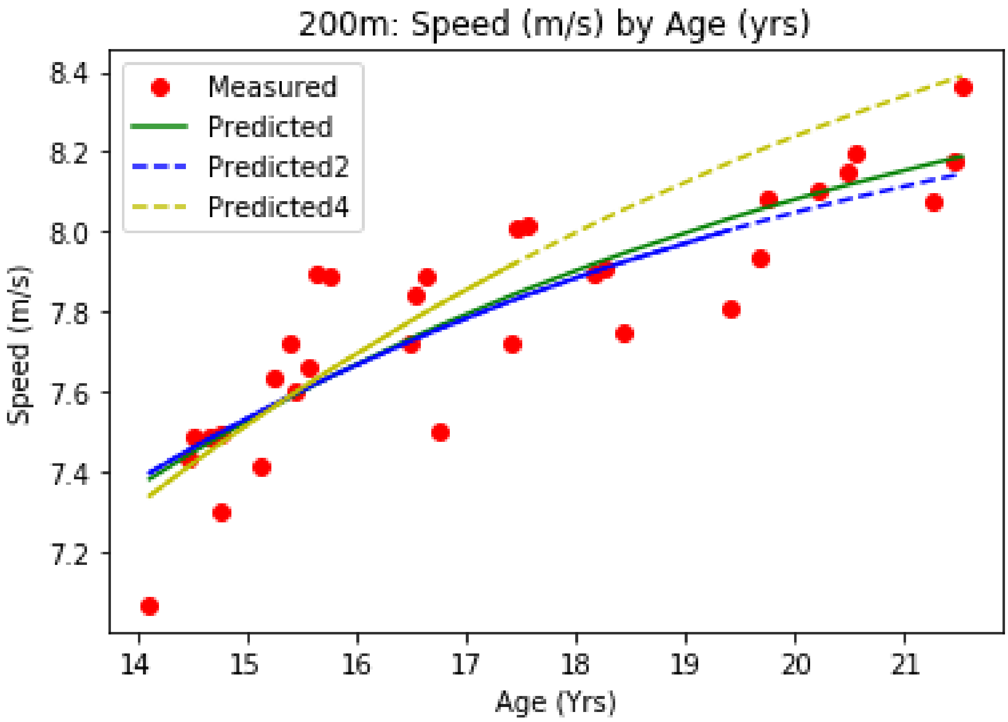

```
31151 Charlotte Wingfield
Optimal parameters
[ 8.46866427 -0.1585592 ]
R^2: 0.7202882159972277
[ 8.4257636  -0.16168558]
R^2_2yr:  0.7190957360604429
[ 8.71294604 -0.14427301]
R^2_4yr:  0.6816611047948946
```


200m: Speed (m/s) by Age (yrs)
Measured
Predicted
Predicted2
Predicted4
Speed (m/s)
Age (Yrs)
8.4
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
12
14
16
18
20
22

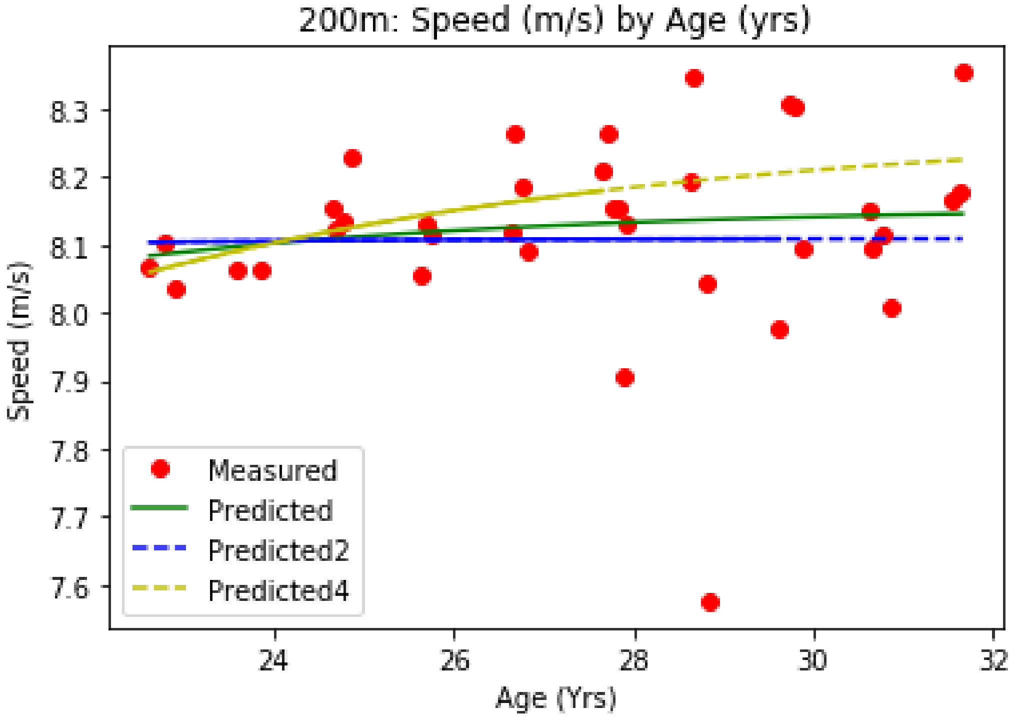
```
20590 Hannah Thomas
Optimal parameters
[ 8.2588716  -0.18125829]
R^2: 0.3527920848121179
[ 8.42930343 -0.16370936]
R^2_2yr:  0.31957132061914817
[ 8.85540963 -0.13765664]
R^2_4yr:  -0.08489506778641087
```



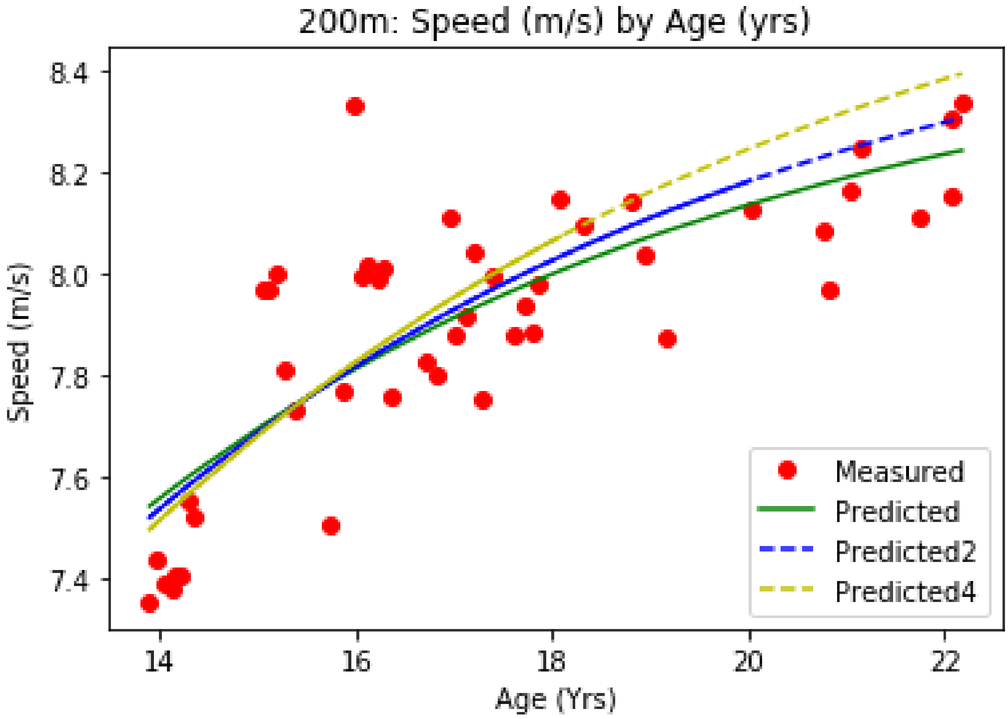
```
30746 Mica Moore
Optimal parameters
[ 8.65065342 -0.13606726]
R^2: 0.7615960899510037
[ 8.54493893 -0.14230191]
R^2_2yr:  0.7560522216031729
[ 9.179121   -0.11392979]
R^2_4yr:  0.6370612281323746
```



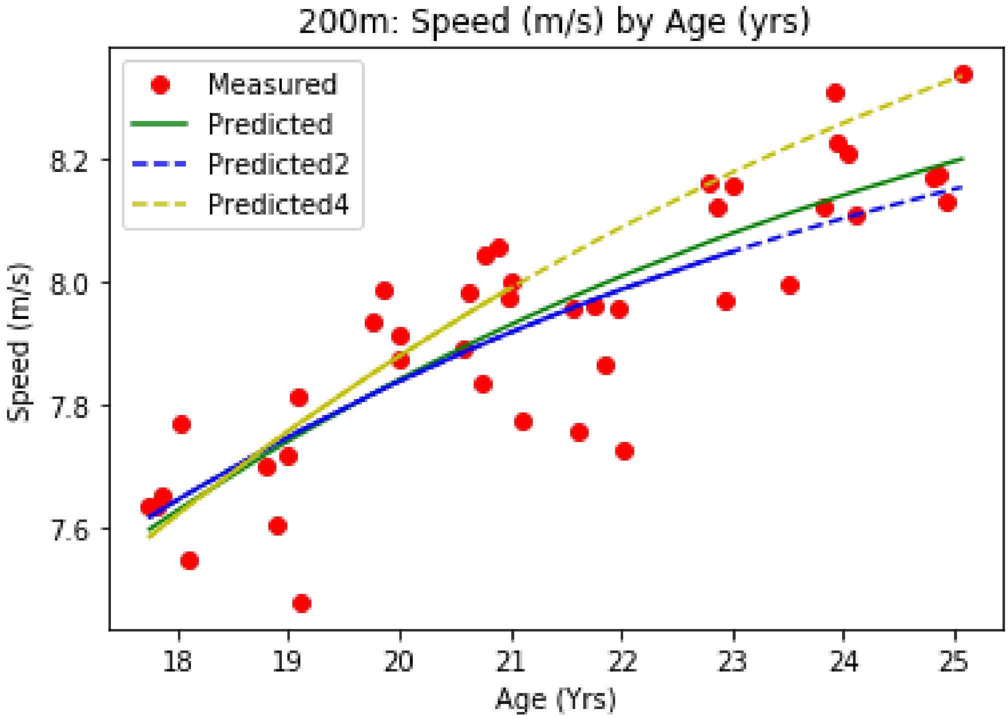
```
34693 Lucy Evans
Optimal parameters
[ 8.15698774 -0.20858309]
R^2: 0.017470639686256506
[ 8.10917935 -0.32318818]
R^2_2yr:  -0.016833162760919285
[ 8.2753684  -0.16135308]
R^2_4yr:  -0.12951712288681771
```



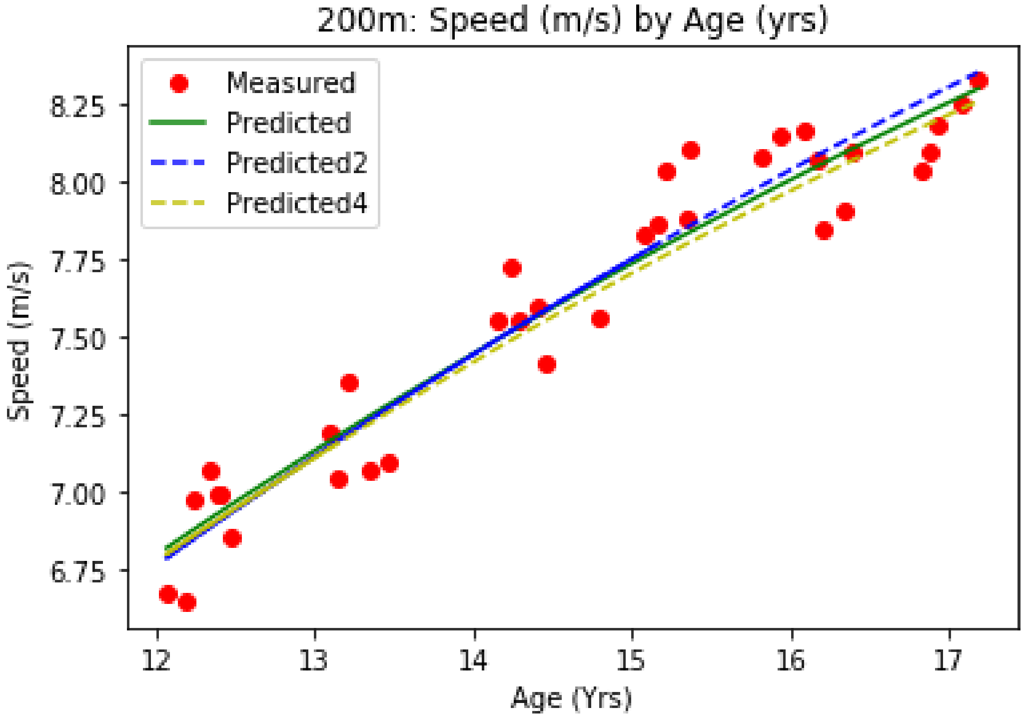
```
8227 Laura Wake
Optimal parameters
[ 8.50799591 -0.15658756]
R^2: 0.643096634034305
[ 8.63905016 -0.14708642]
R^2_2yr:  0.6297570113237725
[ 8.82586603 -0.13619688]
R^2_4yr:  0.5689683198059808
```



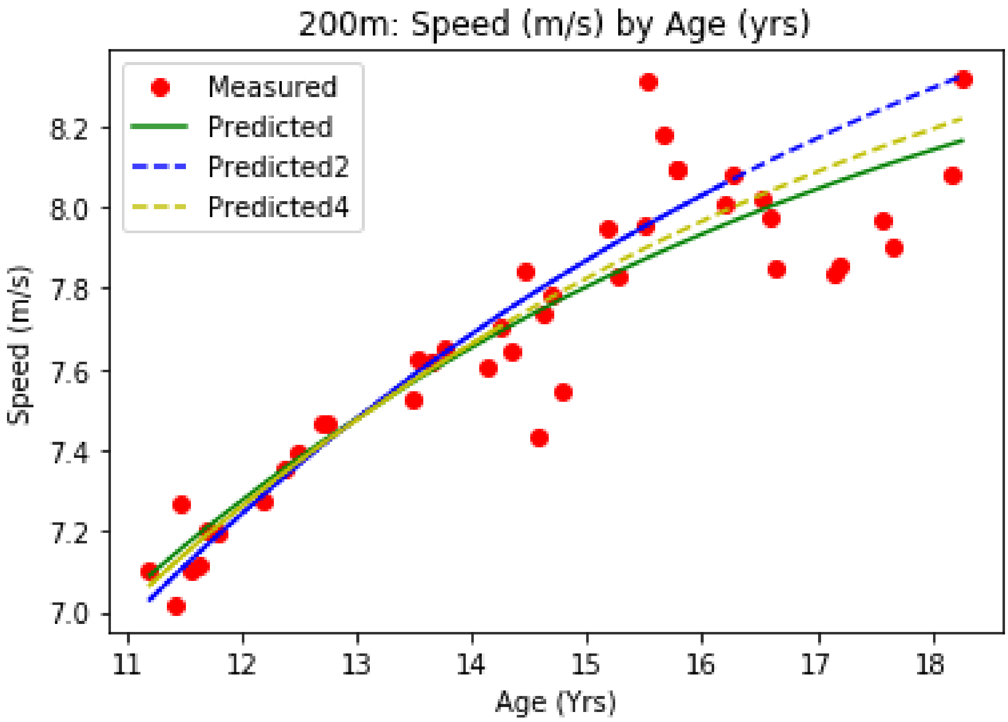
```
47062 Susanna Banjo
Optimal parameters
[ 8.62579253 -0.11982532]
R^2: 0.721331618623952
[ 8.49893429 -0.12765433]
R^2_2yr:  0.708003841561054
[ 8.98128725 -0.10486838]
R^2_4yr:  0.5840932373472034
```



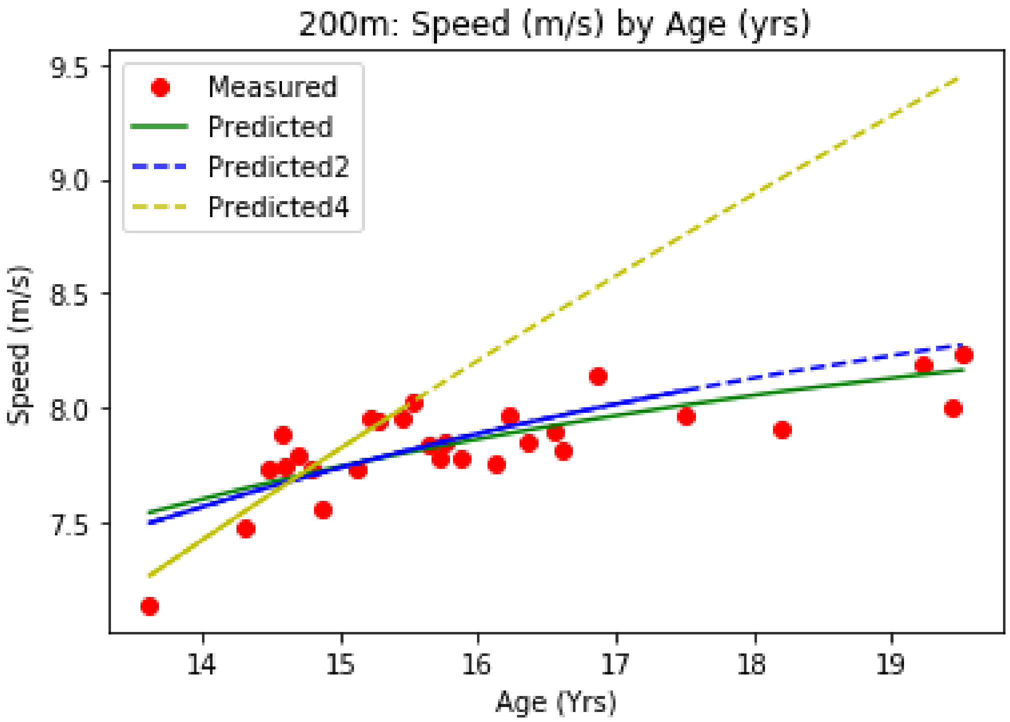
```
90855 Ella Barrett
Optimal parameters
[11.47192743 -0.07474298]
R^2: 0.9253137399727793
[12.09743981 -0.06818745]
R^2_2yr:  0.9219156628948443
[11.33005362 -0.07591655]
R^2_4yr:  0.9215557716271756
```



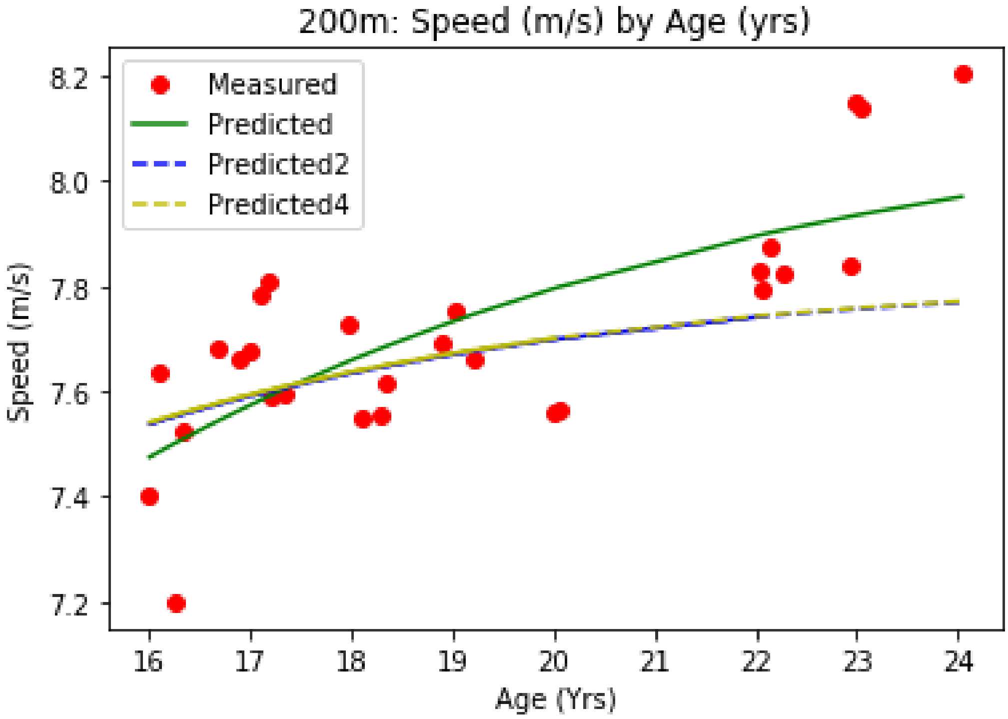
```
83810 Amelia Reynolds
Optimal parameters
[ 8.73667558 -0.14920192]
R^2: 0.8392671057232672
[ 9.18465105 -0.12961203]
R^2_2yr:  0.7854811461829747
[ 8.88619169 -0.14172412]
R^2_4yr:  0.8329992105348951
```



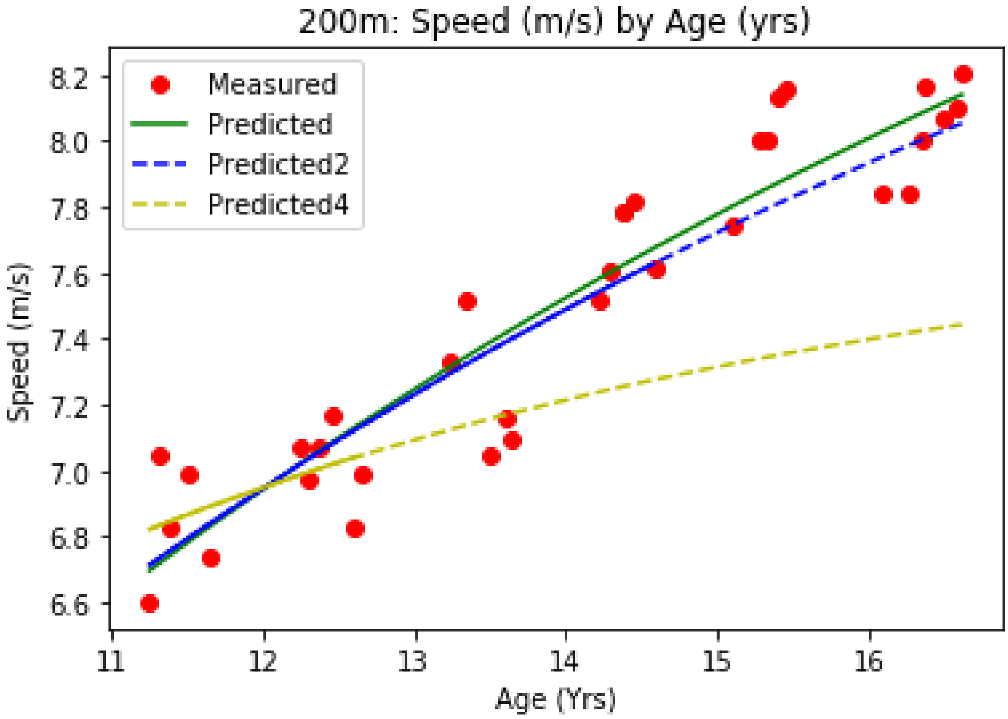
```
46468 Rachel Dickens
Optimal parameters
[ 8.57796751 -0.15555856]
R^2: 0.5683766513368346
[ 8.90908383 -0.13547624]
R^2_2yr:  0.5213274858449991
[18.31485967 -0.03715686]
R^2_4yr:  -5.800823188602775
```



```
22502 Jessica Tappin
Optimal parameters
[ 8.17320143 -0.1536522 ]
R^2: 0.5366308502727044
[ 7.82365574 -0.20675571]
R^2_2yr:  0.3136909337927497
[ 7.82475104 -0.20712954]
R^2_4yr:  0.3176941065581913
```



```
72867 Rachel Norris
Optimal parameters
[10.41547955 -0.09154503]
R^2: 0.874229760525121
[ 9.93836021 -0.10003888]
R^2_2yr:  0.8640782603612984
[ 7.81308568 -0.18344241]
R^2_4yr:  0.18685118387892719
```



```
75444 Abigayle Fitzpatrick
Optimal parameters
[ 8.480648   -0.13329522]
R^2: 0.738667180477154
[ 8.89600543 -0.1125365 ]
R^2_2yr:  0.6036039650372247
[ 8.42958506 -0.13608773]
R^2_4yr:  0.7337793542844246
```