**Predicting the Results of USA Swimming Olympic Trials**

With Wave 2 of USA Swimming Olympic Trials around the corner, levels of spectator speculation are reaching all-time highs. In an attempt to join the join the fun, and potentially have some long-awaited pickem’ success, I have attempted to predict race outcomes using the available data on usaswimming.org and a bit of programming.

**The Data:**

Let’s focus on predicting a few marquee events which will impact USA Relays in Tokyo: The Men’s and Women’s 100 and 200 Freestyle events.

To keep the analysis relatively recent and manageable, data used was from the beginning of the 2015/2016 season up until June 7, 2021. All athletes qualified for the 2021 USA Olympic Trials (Wave 1 and 2) were considered given that no start lists were available at the time of writing.

**The Method:**

For each swimmer, we broke swims down into seasons:

* 2015/16
* 2016/17
* 2017/18
* 2018/19
* 2020/21

The 2019/2020 season was not included due to Covid-19 interruptions.

Each swimmer had their performance plotted, an example of Caeleb Dressel’s 100 Free shown here:

![Chart, scatter chart

Description automatically generated]()

C. Dressel 100 Free Times

The gradient of each season’s linear progression was used to find an expected slope for the current season. In other words, we can look at how much faster or slower a swimmer typically gets throughout the season and apply that information to the current season. To continue the Dressel example, he typically gets steadily faster as a season progresses, and ends up between 2.5 and 3.5 seconds faster at the end of the season than the beginning.

Using linear regression with this predicted slope and the times already swum this season, we can project a 2021 Olympic Trial result.

This result was cross referenced with the trends of season-best performances since 2015. For example, if a swimmer has season best times of 50.0, 49.5, 49.0 in each of the last 3 seasons, we predict they will swim a 48.5 in the following season. Admittedly, this does not perfectly represent the nuances of dropping time in one’s swimming career.

This analysis attempts to consider the overall trend of a swimmer’s improvement, their tendency to drop time at taper, as well as the actual times they have been recently swimming. It does not consider external factors that are ultimately also at play, such as sickness or injury, experience, nerves, or covid related training difficulties. However, it does provide a useful and fun point of analysis, highlighting some exciting ‘up and comers’ that have been rising the ranks and could make a big impact at trials.

**The Results (top 30):**

**Men 100 Free**

|  |  |  |
| --- | --- | --- |
| 1 | Dressel, Caeleb | 47.06 |
| 2 | Apple, Zach | 47.48 |
| 3 | Howard, Robert | 47.69 |
| 4 | Rooney, Maxime | 47.83 |
| 5 | Held, Ryan | 47.98 |
| 6 | Smith, Kieran | 47.99 |
| 7 | Farris, Dean | 48.00 |
| 8 | Jackson, Tate | 48.03 |
| 9 | Pieroni, Blake | 48.10 |
| 10 | Theall, Mark | 48.32 |
| 11 | Seliskar, Andrew | 48.38 |
| 12 | Casas, Shaine | 48.41 |
| 13 | Molacek, Jacob | 48.50 |
| 14 | Adrian, Nathan | 48.55 |
| 15 | Maurer, Luke | 48.64 |
| 16 | Stewart, Coleman | 48.69 |
| 17 | Alexy, Jack | 48.77 |
| 18 | Chadwick, Michael | 48.81 |
| 19 | Krueger, Danny | 48.86 |
| 20 | Newkirk, Jeff | 48.94 |
| 21 | Pinfold, Brett | 48.97 |
| 22 | Conger, Jack | 49.10 |
| 23 | Kibler, Drew | 49.12 |
| 24 | Delakis, Paul | 49.20 |
| 25 | Wright, Colin | 49.22 |
| 26 | Lasco, Destin | 49.25 |
| 27 | Curry, Brooks | 49.26 |
| 28 | Loy, Drew | 49.29 |
| 29 | Shebat, John | 49.29 |
| 30 | Chaney, Adam | 49.32 |

**Women 100 Free**

|  |  |  |
| --- | --- | --- |
| 1 | Manuel, Simone | 52.74 |
| 2 | Weitzeil, Abbey | 53.12 |
| 3 | Deloof, Catie | 53.33 |
| 4 | Huske, Torri | 53.36 |
| 5 | McLaughlin, Katie | 53.38 |
| 6 | Curzan, Claire | 53.45 |
| 7 | Walsh, Gretchen | 53.68 |
| 8 | Brown, Erika | 53.74 |
| 9 | Comerford, Mallory | 53.83 |
| 10 | Kendall, Amanda | 54.11 |
| 11 | Tetzloff, Aly | 54.23 |
| 12 | Mack, Linnea | 54.25 |
| 13 | Geer, Margo | 54.25 |
| 14 | Hinds, Natalie | 54.37 |
| 15 | Tang, Amy | 54.48 |
| 16 | Smoliga, Olivia | 54.49 |
| 17 | Ledecky, Katie | 54.63 |
| 18 | Ivey, Isabel | 54.64 |
| 19 | Schmitt, Allison | 54.72 |
| 20 | DeLoof, Gabby | 54.73 |
| 21 | Fisch, Claire | 54.77 |
| 22 | Alons, Kylee | 54.80 |
| 23 | Walsh, Alex | 54.85 |
| 24 | Pelaez, Erika | 55.00 |
| 25 | Dahlia (Worrell), Kelsi | 55.11 |
| 26 | Garofalo, Isabella | 55.14 |
| 27 | Nelson, Beata | 55.15 |
| 28 | Parker, Maxine | 55.16 |
| 29 | Douglass, Kate | 55.19 |
| 30 | Perry, Ky-lee | 55.25 |

**Men 200 Free**

|  |  |  |
| --- | --- | --- |
| 1 | Smith, Kieran | 1:44.76 |
| 2 | Foster, Carson | 1:46.34 |
| 3 | Dressel, Caeleb | 1:46.48 |
| 4 | Julian, Trenton | 1:46.63 |
| 5 | Seliskar, Andrew | 1:46.65 |
| 6 | Apple, Zach | 1:46.79 |
| 7 | House, Grant | 1:46.83 |
| 8 | Callan, Patrick | 1:46.98 |
| 9 | Magahey, Jake | 1:47.06 |
| 10 | Pieroni, Blake | 1:47.09 |
| 11 | Kibler, Drew | 1:47.10 |
| 12 | Haas, Townley | 1:47.31 |
| 13 | Urlando, Luca | 1:47.33 |
| 14 | Farris, Dean | 1:47.44 |
| 15 | Rooney, Maxime | 1:47.53 |
| 16 | Newkirk, Jeff | 1:47.62 |
| 17 | Davis, Wyatt | 1:47.98 |
| 18 | Bybee, Cody | 1:47.98 |
| 19 | Freeman, Trey | 1:48.00 |
| 20 | Sztolcman, Christian | 1:48.15 |
| 21 | Walker, Jack | 1:48.19 |
| 22 | Hill, Julian | 1:48.34 |
| 23 | Brown, Zach | 1:48.37 |
| 24 | Theall, Mark | 1:48.48 |
| 25 | Carrozza, Coby | 1:48.68 |
| 26 | Harting, Zach | 1:48.76 |
| 27 | Grothe, Zane | 1:48.81 |
| 28 | Stone, Lane | 1:48.84 |
| 29 | Delakis, Paul | 1:48.86 |
| 30 | Mitchell, Jake | 1:48.87 |

**Women 200 Free**

|  |  |  |
| --- | --- | --- |
| 1 | Ledecky, Katie | 1:55.69 |
| 2 | McLaughlin, Katie | 1:55.85 |
| 3 | Madden, Paige | 1:56.41 |
| 4 | Manuel, Simone | 1:57.06 |
| 5 | Schmitt, Allison | 1:57.21 |
| 6 | Forde, Brooke | 1:57.52 |
| 7 | Smith, Leah | 1:57.76 |
| 8 | Smith, Regan | 1:57.91 |
| 9 | Raab, Meaghan | 1:57.91 |
| 10 | DeLoof, Gabby | 1:57.94 |
| 11 | Flickinger, Hali | 1:57.98 |
| 12 | Cox, Madisyn | 1:58.08 |
| 13 | Gemmell, Erin | 1:58.37 |
| 14 | Margalis, Melanie | 1:58.41 |
| 15 | Weyant, Emma | 1:58.66 |
| 16 | Kozan, Justina | 1:58.71 |
| 17 | Tuggle, Claire | 1:58.83 |
| 18 | Laning, Erica | 1:58.89 |
| 19 | Comerford, Mallory | 1:59.20 |
| 20 | Strouse, Ashley | 1:59.37 |
| 21 | Runge, Cierra | 1:59.43 |
| 22 | Drabot, Katie | 1:59.49 |
| 23 | Smoliga, Olivia | 1:59.55 |
| 24 | Pitzer, Lauren | 1:59.55 |
| 25 | Nordmann, Lillie | 1:59.56 |
| 26 | Deloof, Catie | 1:59.59 |
| 27 | Nordin, Emma | 1:59.61 |
| 28 | Stege, Rachel | 1:59.61 |
| 29 | Hicks, Chloe | 1:59.64 |
| 30 | Pash, Kelly | 1:59.68 |

Disclaimer: This is by no means a perfect example of performance prediction. There are hundreds of ways to potentially predict the outcome of a swimming race, some using more historical data or mathematical abstraction than others. I tried to choose a model which produced accurate results and married the available data with the ability to share interpretable methods and my own (limited) skill level. If you would like to see my code or have any feedback and/or suggestions about my methodology, I would love to hear it. Let me know by email at [schaferjacqui@gmail.com](mailto:schaferjacqui@gmail.com).