Optimizing Collegiate Cross-Country Resource Allocation

Austin Stephen

July 2020

Why Collegiate Cross-Country Matters to Your School

In division 1 FBS there are 112 schools that maintain a cross country program, and in 2016 the median autonomy division 1 program spent 1.6 million dollars on cross-country with the average head coach garnering a salary of $108,000 a year (Fulks and Emeritus. 2016). Yet, these programs only average an annual revenue of $120,000, meaning they hemorrhage over a million dollars a year. Universities are not ignorant to this cost; they allocate resources to these programs with a highly targeted purpose. A sports program's successes bolster an academic institutions reputation and can draw the public's attention to the school. In fact, a 2009 study found success on the national stage for a major sports team could boost applications to a school by 9 percent (Pope and Pope, 2009). While smaller programs like cross-country do not generate the awe-inspiring crowds of football or basketball, program success undoubtably influences a school's prestige and national relevance. From this structure of benefits, teams are highly compelled to maximize their performance at national meets to justify their investment in these programs.

In the following analysis, I examine how teams could target their finite financial resources on specific areas that yield the highest improvement in overall placement. I also present an R script that details which runners on a team are the most sensitive to improvements that would raise a team's standing. Running this script will enable coaches to implement my conclusions in the context of their own teams.

Cross-Country Scoring Paradigm

In a cross-country meet, each team enters 7 runners and the 5 fastest of those runners count toward the teams' score. For example, a team who's top 5 runners place first, second, tenth, eighteenth and twentieth would receive a score of 51 (1+2+10+18+20). Therefore, the faster a team's runners run, the lower they place in the race, lowering the teams score and improving the team's overall place. This scoring dynamic uses a runner's place, not their time, to gauge their performance. While certainly time and place are tightly correlated, the relationship is non-linear. **Figure 1** demonstrates this relationship between time and place by graphing time as a function of place with each dot representing a single runner. **Figure 2** plots a linear regression line through the function to aid in visualizing how it deviates from the point of inflection. Figure 1 and 2 demonstrate the relationship between place and time is sigmoid curve experiencing logistical growth at the beginning and end of the distribution.

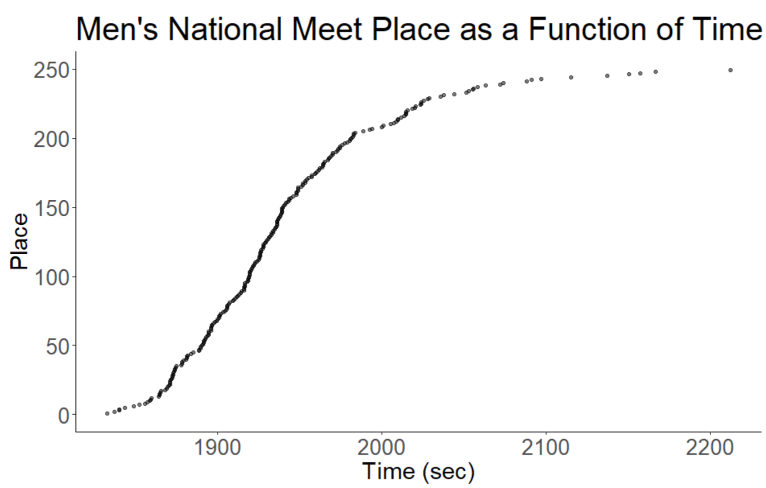

Fig. 1 Men's national meet 2019 results. X-axis the runners final time in seconds. Y-axis the runners overall place

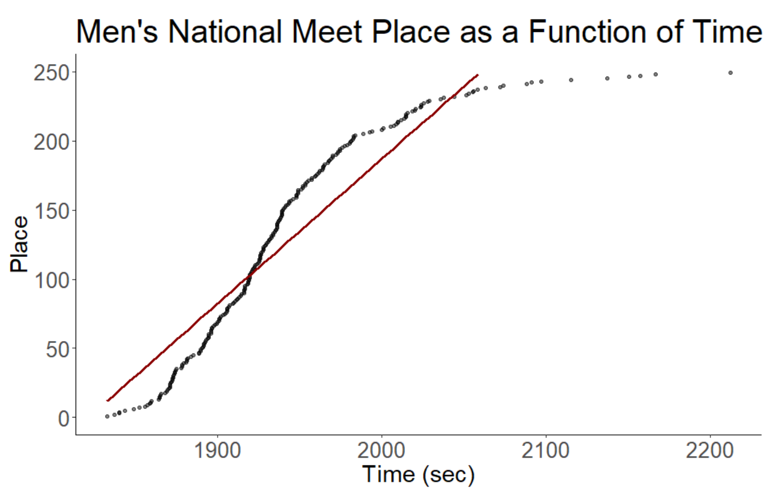

Fig. 2 Men's national meet 2019 results. X-axis the runners final time in seconds. Y-axis the runners overall place. Linear regression line (red) to highlight the curvature

This sigmoid curve is consistent across genders and types of championship races. **Figure 3** shows a highly similar curve for the Women's 2019 D1 national championship as well as the Men's and Women's 2019 Great lakes region championship.

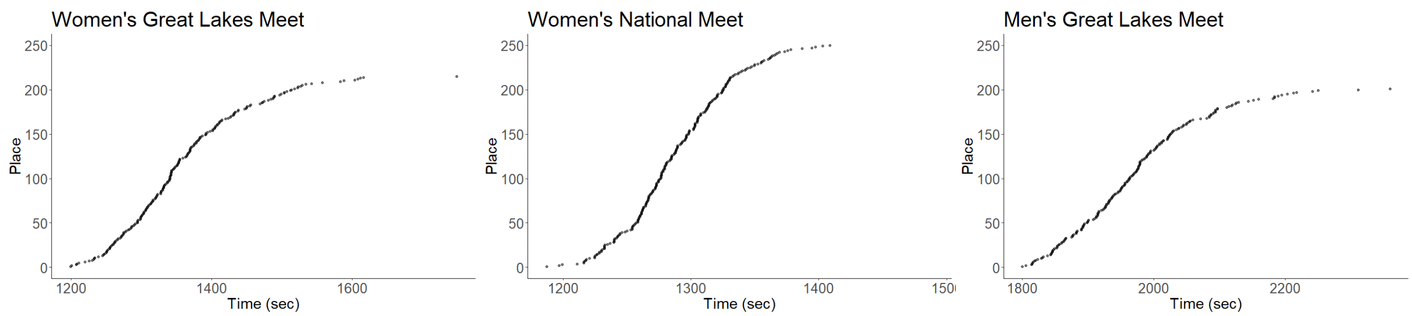

Fig. 3 Men' and Women's Great Lakes regional, and Women's national meet. X-axis the runners final time in seconds. Y-axis the runners overall place

There is one major implication of this relationship between place and time. The same improvements in time do not equally influence a team's place. The drastic difference between which runner improves can be seen in **Figure 4** which highlights in orange the runners for Virginia Tech. Should Virginia Tech's first runner improve by 30 seconds, their team score decreases by 3 points, however, if their fourth runner saw this same 30 second improvement, Virginia Tech reduces their team score by 40 points. Observing time improvements in the optimal runner/runners can result is major shifts in placement.

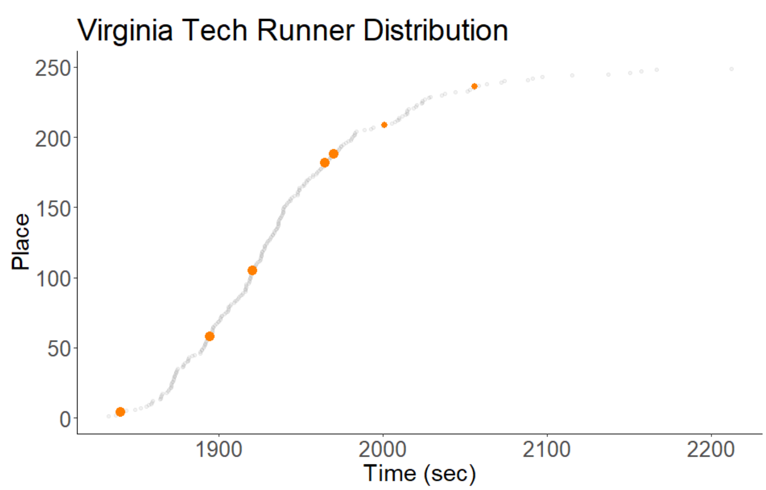

Fig. 4 Men's national meet 2019 with the first 5 Virginia Tech runners enlarged and all 7 runners highlighted in orange. X-axis the runners final time in seconds. Y-axis the runners overall place

Optimize Resource Allocation in Practice

The sigmoid curve modeling the relationship between runner place and time indicates the runner closest to the point of inflection would most dramatically impact team place. I created an R script that will allow a coach to tell what order of importance a runner offers to improving their teams score. This script orders the runners on a team by how close they are to the median of the distribution for a given race. The function takes in 2 parameters, first the seeding of the runners for the race, and second, the team of interest. The script then returns a list of the team of interest's 7 runners in order of how important they are to the team's score. This would give coaches a strategic edge in helping runners prepare for a given race.

Further Work

The biggest limitation of the solution I offer is a runner's PR is often a misleading metric for race performance. A more robust model could use all of a runner's times over a season weighting them by recency and normalizing them for course difficulty and elevation. Then it could seed the race using this metric offering a very realistic picture of how the runners compare. This would offer more accurate data on where a team's runners lie in the given field and would enable much more certainty in the accuracy of the ranking offered by the prediction model. The reason this was not pursued is the difficulty in accessing course information and runner's historical data with web scrapers.

Conclusion

Even resource distribution over a school's cross-country programs resources ignores the scoring paradigm of competition. For a team to optimize their placement in competitions they should focus resource allocation on the runners that are most influential to the team score. The R-script I offer can be used to tell a team which runners those are and help them achieve this goal.

References

Fulks and Emeritus. 2016. "Revenue and Expenses". NCAA Research. https://www.ncaa.org/sites/default/files/2017RES_D1-RevExp_Entire_2017_Final_20180123.pdf

Pope, Devin G., and Jaren C. Pope. "The Impact of College Sports Success on the Quantity and Quality of Student Applications." Southern Economic Journal, vol. 75, no. 3, 2009, pp. 750–780. JSTOR, www.jstor.org/stable/27751414. Accessed 25 Nov. 2020.

https://pttiming.com/results-current.php. Accessed July 2020.