**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. 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 and as a result schools invest over a million dollars a year into cross-country alone.

A school’s return on investment in their cross-country program is determined almost exclusively by the team’s success, making it a natural candidate to examine for optimization. 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 in a single race. Of the teams 7 runners, the 5 fastest in that race get their overall place totaled yielding the team 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, lowering the teams score and improving the teams 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.

A screenshot of a cell phone

Description automatically generated

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

**A close up of a map

Description automatically generated**

**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.

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

**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. The incredible value of observing time improvements in the optimal runner/runners confers a distinct advantage on any team that can leverage this information.

A close up of a person

Description automatically generated

**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

A Model to Optimize Team Place Improvement

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. However, this overlooks an important component of a runner reducing their time, which is that all improvements are not created equal. The faster a runner gets the harder it becomes for them to improve their time because the improvement constitutes a larger percentage of the total time they spend running. For example, 30 a second improvement on a 30-minute 10k is 1.7% improvement, whereas a 30 second improvement on a 35-minute 10k represents a 1.4% improvement. Therefore, a model suggesting which athlete to improve must utilize relative time improvements instead of absolute.