The Impact of Altitude Training on NCAA Division I Female Swimmers' Performance

Katherine L. Manzione
Department of Applied and Computational Mathematics and Statistics
University of Notre Dame
kmanzion@nd.edu

Bailey K. Fosdick
Department of Statistics
Colorado State University
Bailey.Fosdick@colostate.edu

Ryan Elmore
Department of Business Information and Analytics
University of Denver
Ryan.Elmore@du.edu

Connor Gibbs
Department of Statistics
Colorado State University
Connor.Gibbs@colostate.edu

September 23, 2021

Abstract

In this study, we investigate the effects of a female swimmer training at altitude prior to college and how that training ultimately affects performance in her college career. In particular, for female athletes training in college at altitudes above 5,000 feet, we test whether athletes that trained at altitudes above 5,000 feet prior to college had different time drops between collegiate and high school best times, compared to athletes whose pre-collegiate training was below 5,000 feet. We test the hypothesis that having trained in college at altitudes above 5,000 feet and training prior to college at altitudes above 5,000 feet compared to altitudes below 5,000 feet will result in different observed time drops between collegiate and high school best times. We considered swimmers from four NCAA Division I colleges located in Colorado at altitudes ranging from approximately 5,000 to 7,000 feet above sea level. We collected data on the best pre-collegiate and collegiate times of 167 sprint event (50 freestyle) swimmers,

164 middle-distance event (200 freestyle) swimmers, and 69 distance event (1650 freestyle) swimmers. The subjects were grouped by event specialty and whether or not they trained at altitude prior to their collegiate careers. Time improvements from best pre-collegiate times to best collegiate times between altitude groupings were computed for swimmers and event groups. We conclude that there is no significant drop in time between swimmers who swam at below 5,000 feet in altitude prior to their collegiate careers and those who swam at or above 5,000 feet in altitude.

1 Introduction

Altitude training is generally believed by athletes in many endurance-based sports to assist with higher levels of performance at lower altitudes [1, 2, 3, 4]. When traveling to lower altitudes after having trained heavily in a higher elevation, swimmers have traditionally felt that swimming is easier and that they can swim faster times at sea level [4]. Many coaches claim to have seen dramatic time improvements in their swimmers' performances after having trained at high altitudes and having competed at lower altitudes [3, 4]. Head Coach Christopher Woodard and Assistant Coach Lisa Ginder of the Colorado State University Women's Swimming and Diving team both suggest "time drops are attributable to positive psychological responses and psychological adaptations within the swimmers" [3, 4].

In particular, swimmers coming from low altitude locations normally see large changes in their times in the first year due to new high altitude training circumstances. In contrast, swimmers who remain at high altitudes into their collegiate training are believed to negligibly improve, having been already acclimated to higher altitudes [3].

Despite common belief, the degree to which training at high and racing at low altitudes is successful remains ambiguous [5]. Studies have instead argued for and proved with sound evidence the concept of "live high, train low," implying living at altitudes of approximately 2,000 to 3,000 meters and training at altitudes closer to sea-level. This suggests that training and living at altitudes above 5,000 feet may not have the most profound effect on time drops [6]. Various simulation studies with athletes training in simulated altitudes of 2,500 meters or 4,000-5,500 meters found that an athletes' submaximal economy, which has been estimated from relationships between and athlete's absolute oxygen intake and speed during exertion, was not improved [7]. However, observed time improvements were attributed to elevated training intensity rather than high altitude conditions. [8]. Furthermore, studies have concluded that the time drops of swimmers at altitudes at or above 5,000ft and the time drops of swimmers at altitudes below 5,000ft are similar when controlling for other factors such as recovery time, training intensity, and training style [8, 9]. These concepts and their effects have been explored and confirmed repeatedly in multiple studies across various endurance sports including cycling [2], running [6, 7, 10, 11], and swimming [7, 8, 9].

This study was conduced to determine if high altitude training, from a quantitative perspective rather than a biological standpoint, is associated with time drops of NCAA Division I, Mid-Major conference, female swimmers. We attempt to answer the question of whether or not swim coaches at high altitude (approximately 5,000 feet and above) schools should consider altitude training, or the lack thereof, when evaluating high school swimmers. In summary: does high altitude training have a positive effect on a swimmer's performance as the general swimming population tends to believe?

The paper is outlined as follows. Section 2 describes the data involved in this study, including the variables and sources from which they came. Section 3 outlines our motivating research hypotheses, and we summarize the results in Section 4. Finally, we provide a discussion and concluding remarks in Section 5.

2 Data

In collegiate swimming, there are different forms of formal competition, including dual meets, invitationals, and championship meets. Dual meets typically consist of one team swimming against another. Invitationals and championship meets involve multiple teams, while championship meets usually require time qualifications. In this study, times were collected from results posted across these various collegiate meet settings.

All swimmers' times, both high school and college, and hometowns were attained from a comprehensive swimming times database swimcloud [12], and the elevation data were attained from a diverse search engine housing various structured databases [13]. All data were scraped using `R` package `rvest` [14]. Subsequent data import, manipulation, tidying, and visualization were performed using the `tidyverse` ecosystem in `R` [15]. The data and code used are available on Github at https://github.com/rtelmore/swimming.

The final data set includes female swimmers from Colorado State University, the United States Air Force Academy, the University of Denver, and the University of Northern Colorado. The data collection includes the past ten seasons of collegiate competition for each school (years 2010 - 2020). We have $n = 164$ observations for the sprint 50 freestyle, $n = 162$ for the middle-distance 200 freestyle, and $n = 69$ for the distance 1650 freestyle. Swimmers pre-collegiate altitude training was categorized by the elevation of their hometown, i.e. did they train at altitude or not prior to their collegiate career. The cutoff for "altitude" was set at or above 5,000 feet in elevation, because each of the four schools in our study were all at altitudes of at least 5,000 ft. Note that 13 swimmers competed in all three events and 86 competed in two events. For simplicity, we treat swimmers across events as independent.

All time differences were taken as best college time subtracted from best high school time. It should be noted that positive time differences indicate an improvement in collegiate performance from high school. For example, a college time of 25 seconds being subtracted from a high school time of 30 seconds signifies a five second improvement in time. Intuitively, it is preferable to see a

positive improvement in time. Summary statistics and sample sizes for the data are presented in Tables 1 and 2, respectively.

Table 1: Summary statistics are shown for our response variable of interest, the difference between high school and college times. The differences are broken down by the type of event (50, 200, and 1650 yard freestyle) and altitude categorization.

Event	Altitude Group	Mean	Standard Deviation	Sample Size
50 Free	$\geq$ 5000 ft	-0.592	0.670	46
	$<$ 5000 ft	-0.529	0.826	118
200 Free	$\geq$ 5000 ft	-2.444	3.526	44
	$<$ 5000 ft	-1.680	3.784	118
1650 Free	$\geq$ 5000 ft	-19.858	34.513	18
	$<$ 5000 ft	-10.224	40.817	51

Table 2: Sample sizes for each school are shown for each event (50, 200, and 1650 yard freestyle) and altitude categorization. The schools listed are Colorado State University (CSU), University of Denver (DU), University of Northern Colorado (UNC), and the United States Air Force Academy (USAFA).

Event	Altitude Group	CSU	DU	UNC	USAFA
50 Free	$\geq$ 5000 ft	19	13	14	2
	$<$ 5000 ft	21	28	30	41
200 Free	$\geq$ 5000 ft	16	13	12	4
	$<$ 5000 ft	24	28	28	38
1650 Free	$\geq$ 5000 ft	6	5	5	2
	$<$ 5000 ft	10	8	14	19

3 Hypotheses and Methods

In this section, we will outline our research hypotheses and discuss the methods used to address them individually.

- **Hypothesis 1 (H1)**: Is there a difference in a swimmer's improvement based on whether or not they trained at altitude in high school?
- **Hypothesis 2 (H2)**: Does a swimmer's improvement based on whether or not they trained at altitude in high school vary across colleges?

The first hypothesis investigates the difference between swimmer best high school and best collegiate times based on whether or not they trained at altitude during high school. As mentioned previously, the two altitude groups are determined by high school training at or above 5,000 feet or below 5,000 feet. We test for differences between altitude categories using a Welch's two-sample t-test for differences in means. The test is applied for each of the three events: 50, 200, and 1650 yard freestyle.

Figure 1 illustrates the density of time difference between the altitude categories across the three events. This figure suggests that training above 5,000 feet prior to college might correspond to larger performance improvements in college; however, this will formally be tested in the next section.

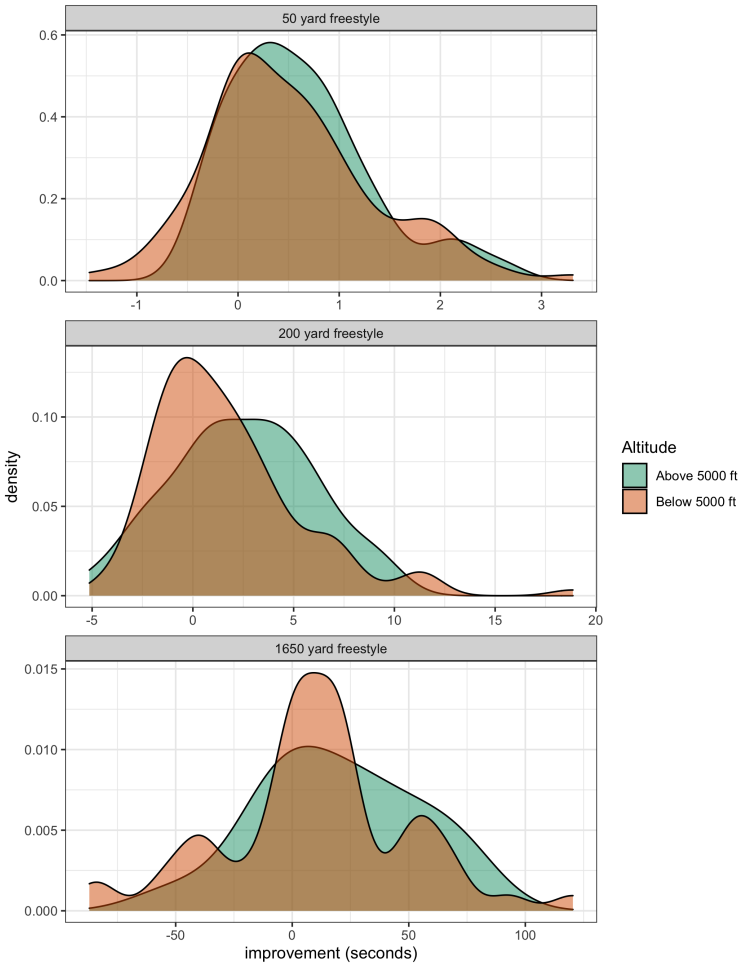

Figure 1: Densities for time improvements by event are shown for the two altitude categories (red is above 5,000 feet and blue is below 5,000 feet).

The second hypothesis investigates if there is evidence that any observed time improvement varies by school. These comparisons were made by using a two-way, non-additive analysis of variance (ANOVA) model for each event (50, 200, and 1650 yard freestyle) with factors for school attended and pre-college altitude training, as well as the interaction between the two factors. We hypothesized that altitude grouping would not have an effect due to previously determined findings of other studies. Boxplots for the cross-categorizations are depicted in Figure 2. In Figure 2, we observe a lack of variation in time difference between groups and between school attended, in general, for the 50 and 200 freestyle events. However, for the 1650 freestyle event we do observe variation in time difference between school attended.

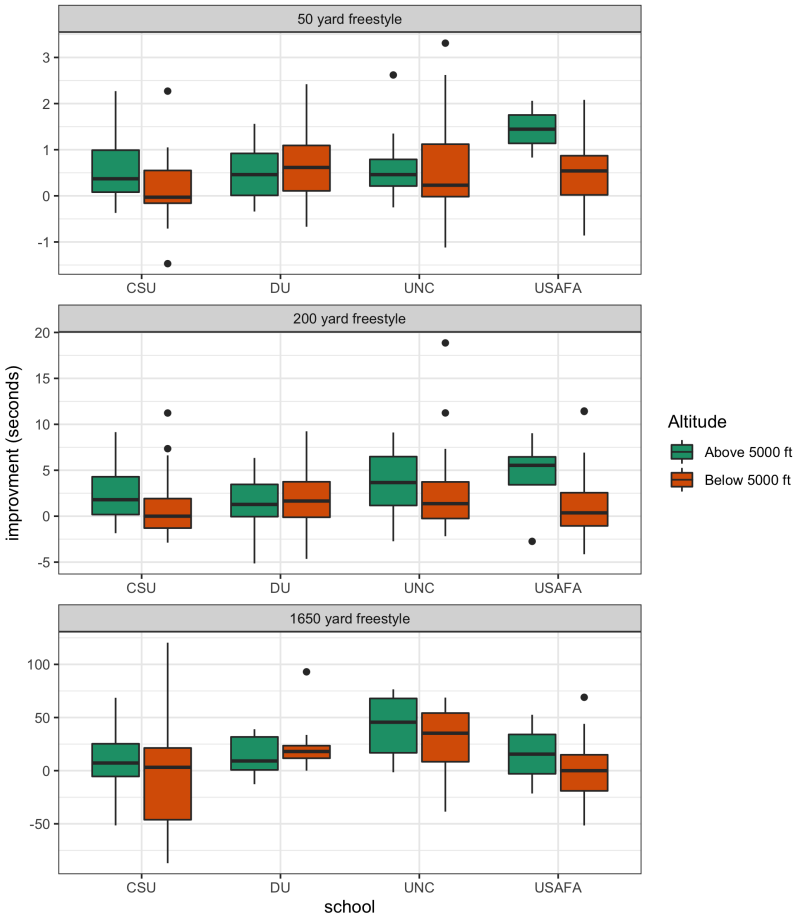

Figure 2: Boxplots of the observed time differences between altitude categorizations (red is above 5,000 feet and blue is below 5,000 feet) subset by the swimmers' college. Sample sizes for each event by school attended are given in Table 2.

4 Results

4.1 Primary Results

We first address the results related to H1. The results of the Welch's two-sample t-test for the difference in best high school and best college swimming times for all events are presented in Table 2. The tests indicate there is insufficient evidence to claim significant differences between the best high school and best collegiate times across the two altitude groups for each of the three events. In other words, there does not appear to be any significant improvement for those swimmers who train at altitude (or not) prior to attending college.

Table 3: Welch's two-Sample t-test for difference in time improvement by high school altitude group

Event	95% Confidence Interval	t	df	p-value
50 Free	(-0.311, 0.184)	0.508	100.55	0.613
200 Free	(-2.028, 0.500)	-1.202	82.306	0.233
1650 Free	(-29.817, 10.548)	-0.969	35.025	0.339

The two-way ANOVA results related to H2 are presented in Table 3 for each event (50, 200, and 1650 yard freestyle). We first note that altitude grouping as well as the interaction between the altitude grouping and school attended did not have significant ($p > 0.05$) effects on the difference in swimmers' time improvements for the 50 freestyle and 200 freestyle events. Recall that Figure 2 illustrates the lack of variation in time difference between groups and between school attended, in general, for the 50 and 200 freestyle events.

The 1650 yard freestyle event, however, does show (see Table 3) a significant difference in improvement across schools ($p < 0.05$). This aligns with that seen in Figure 2 where we see variation by school for the 1650 freestyle event. We note that although school attended has a significant effect on time difference, this does not pertain to the original hypothesis of altitude grouping being important.

4.2 Secondary Results

In this subsection, we examine an alternative approach to analysing the two primary hypotheses of interest. We outline it here as a secondary result since this approach was not specified *a priori*; however, we feel its inclusion is sufficiently interesting. Here we employ a regression-based design as opposed to the ANOVA, or the difference in means-based approach. Rather than categorizing the swimmers as having trained at altitude or not prior to their collegiate careers, we simply compute the difference in altitudes between their high school college. This would allow for an ANCOVA analysis and for the quantification of any observable differences in improvements per unit change in altitude between high

Table 4: The results from fitting two-way, non-additive ANOVA models across the three events under study (50, 200, and 1650 yard freestyle) are summarized here.

Event	Variable	df	F	p
50 Free	Altitude	1	0.529	0.468
	School	3	1.168	0.324
	Altitude*School	3	1.374	0.253
200 Free	Altitude	1	1.488	0.224
	School	3	1.351	0.260
	Altitude*School	3	1.259	0.290
1650 Free	Altitude	1	0.842	0.363
	School	3	2.842	**0.045**
	Altitude*School	3	0.350	0.789

school and college. Informally, this is depicted in Figure 3, and it shows a possible significant change in improvement per foot increase in elevation for the 1650 yard freestyle.

In order to formally test this affect, we fit a non-additive regression model with improvement (y) as a function of event (x_1) and change in altitude (x_2). As alluded to above, the slope of the regression line associated with the 1650 yard freestyle is indeed significantly different from zero (estimated slope = 0.0038, p-value < 0.0001). A similar effect is neither observed for the 50 yard freestyle (p-value = 0.966) nor the 200 yard freestyle (p-value = 0.627).

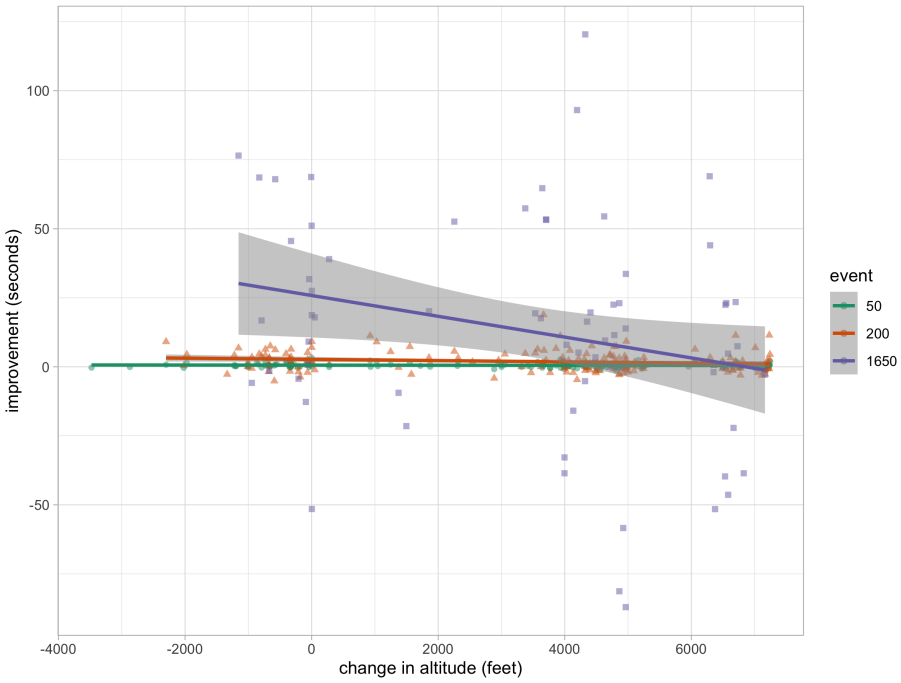

Figure 3: The change in elevation between high school and college (x) versus a swimmer's improvement (y). The individual points and lines show the effects for each event: 50 yard freestyle (green), 200 yard freestyle (orange), and 1650 yard freestyle (purple).

5 Discussion

From this study, we see no significant differences in time between best high school and best collegiate times at the two altitude groups. We also see that altitude grouping and school attended did not have significant effects on the difference in swimmers' time improvements for 50 freestyle and 200 freestyle events but did for the 1650 freestyle event.

Limitations of this study include its observational nature, rather than being a randomized experiment. In order to run this study as a randomized experiment to investigate these hypotheses, we would have had to randomize the swimmers' hometowns, which is not plausible. We cannot control which swimmers are attending these schools. There may be differences between swimmers that attend these altitude schools and those who do not.

References

[1] P. U. Saunders, D. B. Pyne, and C. J. Gore. Endurance Training at Altitude. *High Altitude Medicine and Biology*, 10(2):135–148, 2009.

[2] Maria J. Brosnan, David T. Martin, Allan G. Hahn, Christopher J. Gore, and John A. Hawley. Impaired Interval Exercise Responses in Elite Female Cyclists at Moderate Simulated Altitude. *Journal of Applied Physiology*, 89(5):1891–1824, 2000.

[3] Christopher Woodard. Personal communication.

[4] Lisa Ginder. Personal communication.

[5] Damian Miles Bailey, Bruce Davies, Lee Romer, Lindy Castell, Eric Newsholme, and George Gandy. Implications of Moderate Altitude Training for Sea-level Endurance in Elite Distance Runners. *European Journal of Applied Physiology and Occupational Physiology*, 78:360–368, 1998.

[6] J. Stray-Gundersen, R. F. Chapman, and B. D. Levine. Living High-Training Low Altitude Training Improves Sea Level Performance in Male and Female Elite Runners. *Medicine and Science in Sports and Exercise*, 91(3):1113–1120, 2001.

[7] M. J. Truijens, F. A. Rodríguez, N. E. Townsend, J. Stray-Gundersen, C. J. Gore, and B. D. Levine. The Effect of Intermittent Hypobaric Hypoxic Exposure and Sea Level Training on Submaximal Economy in Well-Trained Swimmers and Runners. *Journal of Applied Physiology*, 104(2):328–337, 2008.

[8] M. J. Truijens, H. M. Toussaint, J. Dow, and B. D. Levine. Effect of High-Intensity Hypoxic Training on Sea-level Swimming Performances. *Journal of Applied Physiology*, 94:733–743, 2003.

[9] Ferran A. Rodriguez, Xavier Iglesias, Belen Feriche, Carmen Calderon-Soto, Diego Chaverri, Nadine B. Wachsmuth, Walter Schmidt, and Benjamin D. Levine. Altitude Training in Elite Swimmers for Sea Level Performance (Altitude Project). *Medicine and Science in Sports and Exercise*, pages 1965–1979, 2015.

[10] J. Stray-Gundersen and B. D. Levine. Implications of Moderate Altitude Training for Sea-level Endurance in Elite Distance Runners. *Scand J Med Sci Sports*, 18(1 Suppl):21–8, 2008.

[11] William C. Adams, Edmund M. Bernauer, D. B. Dill, and John B. Bomar Jr. Effects of Equivalent Sea-level and Altitude Training on Vo2max and Running Performance. *Journal of Applied Physiology*, 39(2):262–266, 1975.

[12] https://www.swimcloud.com/. online public database.

[13] https://www.basedig.com/us-government/list-of-cities-and-towns-in-the-usa 4934/. online public database.

[14] Hadley Wickham. *rvest: Easily Harvest (Scrape) Web Pages*, 2020. R package version 0.3.6.

[15] Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686, 2019.