Calculating Acceleration Vectors in American Football GPS Data

Steven Plaisance

2023-04-25

Abstract

This paper seeks to convert scalar acceleration values from athlete GPS data to a series of vectors describing different types of athlete movement. This increased specificity of measurement provides utility both from a physiological standpoint and a sport-specific production perspective.

1. Introduction

GPS data is an increasingly popular method of monitoring athlete workload and performance indicators in collegiate and professional sports organizations. Ongoing advances in sports medicine, sports science and sports performance are supported by increasingly sophisticated means of measuring and analyzing athlete workload and performance, leading to more advanced understanding of the relationship between workload and performance among sport coaches, strength coaches, athletic trainers and dietitians.

GPS data offers unique utility due its precision of measurement, width of descriptive possibilities and relatively low invasiveness. Compared with other leading forms of athlete monitoring, GPS sensors monitor athletes directly in their sport-specific training environment (ie, on the field during practice rather than in the weight room during lifts.) The presence of GPS sensors changes very little if anything in regards to the sport-specific training environment. Often times, sensors are placed within the shoulder pads prior to practice and athletes are not explicitly aware they are being recorded.

Within American Football, GPS data has facilitated a wide range of uses in both professional and collegiate organizations, though there is little crossover between the two. Professional organizations have focused on developing sport-specific performance and production metrics to evaluate free agents, draft prospects and potential trade targets. Meanwhile collegiate organizations remain focused on physiological performance indicators as a complement to their sports medicine and strength and conditioning departments.

One major limitation of raw GPS data: velocity and acceleration are provided in scalar values. But athletes move in a variety of ways, meaning any sophisticated assessment of physiological performance would account for the various types of movement an athlete may perform. In this paper, I will convert raw scalar acceleration into 8 different vector accelerations, each describing a different type of movement.

- Forward propulsion
- Forward braking
- Backward propulsion
- Backward braking
- Left propulsion
- Left braking
- Right propulsion
- Right braking

More specific measurements of athlete movement provide utility in both the physiological performance area and the sport-specific performance area. Relevant changes in athlete performance are likely to vary across different types of movement. Athletes may also be asked to perform different movement types at varying frequencies based on their position or role within a given playcall.

2. Literature Review

At this time, there is little public research available regarding American Football GPS data from a physiological perspective. Meanwhile, there is a host of football-specific research available thanks to the Big Data Bowl, an annual sports analytics contest hosted by the National Football League. The contest provides a set of GPS data from NFL games and offers contestants a broad prompt to answer creatively.

The winning submission of the 2021 Big Data Bowl (Peng et al.) used a variable framework to assess pass coverage performance based on the defensive playcall and the positioning of nearby receivers at the time of the pass. In essence, models such as this one minimize scouting time by automating tedious charting tasks, saving organizations precious time and money and creating a potential competitive advantage.

Another Big Data Bowl submission from 2019 used a clustering-based approach to identify different route types in GPS data and went on to model optimal route combinations based on defensive personnel and alignment (Chu et al.).

Both of these papers offer possible next steps as my research delves further into sport-specific measures. For example, the route identification could be used in tandem with acceleration vectors to assess different movement types based on the route run. Curl routes necessitate a strong forward propulsion followed by forward braking, while crossing routes require a smooth transition from forward propulsion to left/right propulsion depending on the direction of the route.

3. Data

The raw GPS data includes the following variables:

- athlete_id: A uid describing the athlete being recorded
- stream_type: Specifies whether the data was recorded via satellite positioning or local positioning
- x: field x coordinates (meters)
- y: field y coordinates (meters)
- ts: POSIX time in seconds since the start of the epoch
- cs: Observation time offset in centiseconds

- face: the magnetic facing of the unit
- v: velocity (meters per second)
- a: acceleration (meters per second per second)
- pq: positional quality (percentage)

The data is sampled at 10Hz, meaning there are 10 observations per athlete per second.

Due to data size limitations, the data provided in this paper is specific to one athlete from one practice.

4. Empirical Methods

A handful of new variables must be calulcated in order to arrive at acceleration vectors.

First, dx and dy represent the instantaneous change in positioning from the last observation in the x and y planes. These are used to calculate alpha (the direction the athlete is moving)

$$\alpha = atan2(dy, dx)$$

The key variable used to calculate acceleration vectors is designed as "theta" and describes the difference, in degrees, between the direction that the athlete is facing (face) and the direction that the athlete is moving (alpha)

$$\theta = \alpha - face$$

Next, we calculate four proportion variables, one for each direction of movement (F for forward, B for backward, L for left and R for right). These variables take a range from 0 to 1 and describe the proportion of scalar movement that is in the designated direction.

$$p_F = \begin{cases} 1 - (|\theta|/90),\ \theta < 90\ \&\ \theta > -90 \\ 0,\ Otherwise \end{cases}$$

$$p_B = \begin{cases} (|\theta| - 90)/90,\ \theta > 90\ OR\ \theta < -90 \\ 0,\ Otherwise \end{cases}$$

$$p_L = \begin{cases} 1,\ \theta = -90 \\ 0,\ \theta = -180 \\ (90 - (-\theta \mod 90))/90,\ \theta < -90 \\ -\theta/90,\ \theta < 0 \\ 0,\ Otherwise \end{cases}$$

$$p_R = \begin{cases} 1,\ \theta = 90 \\ 0,\ \theta = -180 \\ (90 - (\theta \mod 90))/90,\ \theta > 90 \\ \theta/90,\ \theta > 0 \\ 0,\ Otherwise \end{cases}$$

From there, the proportion variables are multiplied by the scalar distance variable to find vector distance in each direction. Vector velocities are calculated from changes in vector distance, and vector accelerations calculated from changes in vector velocity.

5. Research Findings

As expected, the athlete recorded the highest propulsion and braking acceleration values moving in the forward direction, weight both the left and right directions returning lower values compared to forward, and backward returning the lowest values. See table 1.1 for maximum acceleration values in each direction. I've also included the scalar acceleration value for comparison.

- Include distribution plot which shows more defined difference between each direction.
- Possibly include multiple practices and observe how each vector changed / didn't change over time.

6. Conclusion

This athlete in particular appears to move more explosively toward his left, both in terms of propulsion and braking. This could be an indicator of muscle imbalance or increased muscle fatigue in one side, or possibly a product of the athlete's physiology.

In terms of football-specific insights, an athlete that moves more explosively toward his left would be more effective running routes to his left rather than his right.

- Vector accelerations provide more specific insights from a physiological standpoint.
- They also open the door for more football-specific metrics that can be calculated down the road.

References

Peng, W., Richards, M., Walczak, S., & Werner, J. (2021). A Defensive Player Coverage Evaluation Framework. NFL Big Data Bowl 2021.

Chu, D., Wu, L., Reyers, M., & Thomson, J. (2019). Routes to Success. NFL Big Data Bowl 2019.

Tables and Figures

1.1

Maximum Values of Vector and Scalar Accelerations (m/s^2)

F Propulsion	L Propulsion	R Propulsion	B Propulsion	F Braking	L Braking	R Braking	B Braking	Scalar A
4.3	4	3.6	3.4	-4.6	-4	-3.3	-3.5	

- Don't know how to prevent this table from running off the egde