

DRAFT of

Spatial Statistics to Evaluate Player Contribution in Ultimate

Authors: Jeremy C. Weiss, Sean Childers

Abstract

Overview

Existing statistical ultimate (Frisbee) analyses rely on summary statistics, such as completion percentage and scoring rate, to assess player strengths and weaknesses. However, these statistics are limited in their ability to evaluate a player's contribution to winning points; that is, two players of different value may look identical statistically. To better determine player contribution, we develop a spatially-aware measure. We leverage sequential, location-based data to build scoring probability maps that aggregate possession outcomes with a function of location. From these maps we define an Expected Contribution (EC) measure that captures a player's ability to increase their team's spatially-defined probability of scoring, which can be separated into unique contribution scores from throwing, receiving, and defense. Our measure weighs both positive and negative actions—completions and blocks, as well as turnovers and yards yielded—based on the change of the location-based scoring probabilities. We validate our model on real data from high-level ultimate, showing that our measure both aligns with ultimate intuition while also identifying undervalued “dark-horse” players who contribute statistically to wins without garnering attention.

Methods

We collected data at elite tournaments using the using the UltiApps iPad application that tracks thrower and receiver locations and point outcomes. Using the location-based data, we produce team-specific and aggregate scoring probability graphs using logistic regression, LOESS, and k -nearest neighbors models. Overlaying each play on the graph determines an attributable contribution, i.e. the change in scoring probability, which is assigned to each involved player. Averaging over all points yields a player's Expected Contribution.

Results

Our measure reliably distinguishes the best throwers, receivers, and defenders in both men's and women's divisions. The Expected Contribution measure combines these factors to produce an assessment of players that accounts for their strengths and weaknesses in each phase of the game. As a secondary outcome, the scoring probability graphs can be also be used to compare the effectiveness of strategies, such as punts versus possession and risk-seeking (high variance throws) versus risk-averse.

1. Introduction: upgrading ultimate statistics

Alongside recent acquisitions of ESPN contracts and two professional leagues vying for market dominance, the sport of ultimate is beginning to incorporate advanced statistics and inducing a “sabermetric” revolution. The increasing value of obtaining a competitive edge has led to teams turning towards more sophisticated data collection and analysis. The joint collaboration of Ultiworld journalists and start-up UltiApps has led to the publication of a series of statistical articles about the utility of advanced statistics in strategic decision-making and game planning and has gained broad readership within the ultimate community. As part of the impetus to introduce advanced statistics, we at Ultiworld have gleaned approaches for analysis from other sports to improve our own measures. In this work, we present and discuss player efficiency measures for ultimate, akin to, *e.g.*, the PER in basketball or WPA in baseball and football. We call these measures expected contributions (ECs), as they assess the relative contribution a player makes to improving the team's chances of winning games and provide a statistical method for ranking players.

In contrast to related sports, the rules of ultimate include one that enables casual yet spatially-aware data

collection: namely that the player in possession of the disc must remain stationary (*i.e.*, maintain a pivot). This allows observers to track an ultimate point using a sequence of player locations, entered real-time into the UltiApps mobile application. We show we can use this simple, low-cost solution and provide insightful analyses that can help guide ultimate decision making and player evaluation.

To get expected contributions (ECs), we first use the data to estimate a spatially-dependent probability of scoring. With such estimates, we can use the spatial movement of the disc involving throwers, receivers, and defenders and attribute the change in probability of scoring to each of those players. Tabulating over throws of the 2013 club season and adjusting for playing time, we derive a total EC measure and break down ECs specific for throwing, receiving and defending.

The EC measure allows us to identify the big contributors as throwers, receivers, and defenders and distinguish them from a larger set of players who previously looked good statistically. For example, the best throwers are ones who consistently complete difficult passes upfield, a throw that goes uncaptured in completion percentage and only has some correlation with the assist statistic. It allows us to compare players who prefer safer throws with high completion rates and players who attempt more difficult throws with more potential upside. Regressing player EC with point outcome, we can assess how much more likely a team is to score given higher EC values.

In Section 2, we briefly review fundamental ultimate concepts, describe the state of ultimate statistics, and propose our new p(Score) and EC measures. In Section 3, we show off our new measures on data collected from the 2013 club season. In Section 4, we describe applications of our model, including player ranking, comparing risk-averse and risk-seeking players, and guiding throw choice. In Section 5, we discuss limitations to the sequence-of-locations approach to data collection and offer directions for expanding upon our analysis or moving towards higher-cost data collection. We conclude in Section 6.

2. Methods: estimating the probability of scoring and attributing it to players

To review, ultimate is a two-team, seven-on-seven game played with a disc on a football field with the goal of possessing the disc in the opponent's endzone, *i.e.*, a score. The player with the disc must remain stationary (maintain a pivot) and release the disc within 10 seconds. If the disc touches the ground while not in possession or the first person to touch the disc after its release is the thrower, the play results in a turnover and a member of the other team picks up the disc with the intent of scoring in the opposite endzone. Each score is worth one point, and the game typically ends when a team reaches 15. Basic statistics are shown in Table 1 that offer an, albeit incomplete, assessment of player ability.

Table 1. Player statistics

Player	Points	Completions (%)	Throws	Goals (/p)	Assists (/p)	Blocks	Turnovers
Childers	65	88 (88)	100	20 (0.31)	4 (0.06)	10	8
Weiss	40	49 (98)	50	2 (0.05)	10 (0.25)	4	1
Eisenhood	55	20 (67)	30	10 (0.18)	1 (0.02)	20	15

Our work introduces a spatial model of the probability of scoring and uses it to produce the expected contribution (EC) measure. We present the framework first and subsequently discuss estimation. Let y denote the point outcome and x denote the recorded features of the play. Specifically y is a dichotomous outcome $\{0,1\}$ with 1 denoting the home team scoring the point and 0 otherwise, and x is a vector holding the start point of the throw

and categorical variables for home team, away team, and weather. Then we seek to model the conditional probability $p(y|x)$, which gives us a probability of scoring from anywhere on the field given possession of the disc. Given an estimator p' of the conditional probability distribution p , we can estimate the change in scoring probability of a play by $P(y|x_1)-p'(y|x_0)$, where x_1 and x_0 denote vectors holding the throw start and end locations, respectively, and $P(y|x_1)$ is 1 if the play resulted in a score, $p'(y|x_1)$ if the play resulted in a non-scoring completion, and $p'(y|x_1')$ if the play resulted in a turnover. The new vector x_1' corresponds to probability of scoring for the opponent, which requires rotating the field coordinates and swapping the home team and away team values.

With a change in scoring probability for each play, we can attribute the value to the involved players: i.e., the thrower, receiver, and defender. By tabulating over all throws, we acquire an aggregate value for throwing, receiving and defending, which we then sum to get a total contribution. Adjusting for playing time, we get the expected contribution (EC).

For estimation of the the conditional probability, we use (1) logistic regression and (2) the non-parametric LOESS on a reduced x vector containing only spatial coordinates. The logistic regression allows us to assess the relationship between probability of scoring and the covariates, while the LOESS provides a data-driven heat map to identify locally where teams have success on the field. As multiple plays are typically associated with a single outcome (multiple passes per point), the plays are not independent of one another and violate the i.i.d. assumption. Intuitively, this means that our analysis does not factor for the “flow” of a point. However, using multiple plays per point provides us a magnitude more data to better characterize contributions and probabilities of scoring. The dependence among plays could also be modeled, for example, by extending the x vector to contain information about previous plays.

Our data comes from the 2013 US club season. Team members or volunteers use the UltiApps mobile application to record the plays in real time by tracking the sequences of passes spatially and recording the involved players and the outcome of the point. This resulted in a data set containing 17,883 plays over 3099 possessions and 1,579 points in 68 games.

3. Results: top throwers, receivers, and defenders

Figure 1 shows the LOESS for probability of scoring given x -coordinate (yard line) for the point and for the possession. As expected field position is important: the probability of scoring increases the closer the disc is to the endzone. Likewise, we can see if there are pockets of the field with low or high probabilities of scoring. Ultimate wisdom suggests that the disc should remain away from the sidelines, where the thrower can get trapped. Figure 2 shows a k -nearest neighbor ($k=100$) map approximating the two-dimensional probability of scoring.

Table 2 shows the offensive EC scores for selected players alongside basic statistics. A more complete table is provided in the Appendix. The players selected epitomize the offensive thrower (Rebholz), the offensive receiver (Kittredge), and two defensive throwers (Lance and Malecek). The EC breakdown effectively captures their respective specialties. Furthermore it neatly distinguishes Lance and Malecek (1) in throw EC, which is hinted at in the assists per point, and (2) defensive EC, which is supported by their blocks per defensive possession (0.11 versus 0.04) and yards given up (173 versus 388). Note Lance’s throw EC is higher despite his lower completion percentage; EC awarded his throws more on average and determined the additional risk was justified.

Table 2. Player statistics and EC.

Player	Completions (%)	Goals (/p)	Assists (/p)	Blocks	Total EC	Throw/Catch/D EC
--------	-----------------	------------	--------------	--------	----------	------------------

Rebholz	167 (97.8)	3 (0.03)	14 (0.12)	1	0.22	0.19/0.12/-0.09
Kittredge	112 (94.1)	20 (0.20)	11 (0.11)	7	0.14	-0.01/0.17/-0.02
Lance	45 (90.0)	1 (0.02)	10 (0.20)	5	0.14	0.08/0.02/0.04
...						
Malecek	54 (96.4)	2 (0.03)	8 (0.10)	4	0.03	0.03/0.01/-0.01

4. Applications:

The table provided in the Appendix provides us a direct way of identifying overall and role-specific talent. To our knowledge, EC is the first composite ultimate metric, which clearly identifies known talent as well as identify underrated players. Also, it allows us to compare risk-averse and risk-seeking players, who typically bias themselves towards high completion percentages or high assists. Compared to the risk-averse, risk-seekers will have a higher variance in contributions, but not necessarily a higher mean.

To develop an idea of the value of a unit of EC, we regressed point outcomes (y) against average total EC controlling for whether the point was an offensive or defensive one. An additional EC of 0.1 increases the log odds of scoring the point by 0.25 (e.g. if your team was 1:1 to win the point, 0.1 EC would make it $e^{0.25}:1 = 1.28:1$, or 56%).

In addition to the EC measure, the conditional probability models can help guide throw choice and defensive strategy. Let us use synthetic probability of scoring graphs for illustration. First, suppose in addition we have a probability of completion graph for a particular player at a particular location on the field. Then we can make a first order approximation of expected value of the throw as function of throw location as shown in Figure 3 (the positive expectation is $p(x_1 \& \text{complete})p(y|x_1)$, the negative expectation is $p(x_1 \& \text{incomplete})p(y|x_1')$). In this case, the best throw is a short throw to the middle, and the long throw has a slightly lower expected value. The difference in expected values accumulates over many plays, suggesting that making optimal throw choices has large effects on the outcome.

On the defensive side, suppose we can employ strategies that alter the probability of completion graphs, ignoring how they influence the probability of scoring graphs for now. We can use the graphs to suggest which defense is most likely to succeed. For example, given a prevent defense that gives up high probability short throws and given pressure defense that covers short throws but gives up positioning on long throws, we can again compare expected outcomes under optimal offensive play. Figure 4 shows that the minimax outcome of 0.05 suggests that prevent defense is better defensive option and is best countered by the short, high-percentage forward throw.

Finally, a similar analysis to Figure 1 performed on women's data leads to the probability of scoring graph in Figure 5. While the graph is likely substantially different under different conditions and opponents, the graph hints strongly at a particular strategy: the Hail Mary (or punting). A 40-yard punt to the opponent's end-zone line changes the probability of scoring from 0.70 to 0.65, requiring only a completion percentage of 4.8% to make the throw worthwhile. The Hail Mary strategy, known in ultimate as Huck and D, is seen in all divisions of ultimate and particularly in poor weather conditions, but is less fun for players and unappealing to spectators. Figure 5 lends support to proponents in the ultimate community of shrinking the women's playing field and of postponing games in poor but non-life threatening conditions.

5. Discussion: constraints and limitations, mining for new strategies

We have presented the first composite ultimate measure, EC, an efficiency measure quantifying the contribution

of players to winning points. The measure is dependent on accurate modeling of the probability of scoring, which could be improved with more sophisticated data collection. For example, our system only tracks the location of the disc, not the other thirteen players on the field. Defensive strategy and offensive positioning clearly affect the probability but are not incorporated. Sample size and the independence of samples are additional concerns. Additionally, the total EC measure also aggregates throwing, receiving and defending, but in some sense the offensive measures are double-counted, leading to larger contributions on offense. One solution would be to divide the offensive contributions in half, but often the difficulty of pass completion lies more in the execution of one player, so equal attribution may be a mistake.

The high cost of more sophisticated data collection is a strong barrier to conducting “all-14” analyses, particularly for a young sport such as ultimate. Instead, additional analyses based on sequential disc movement data could be investigated. Ultiworld has many new statistical analyses lined up, for example comparing the men’s and women’s game, combating the wind, and characterizing team progression and young talent over time.

Furthermore, we believe there is more to gain from this data specifically for efficiency ratings. For example, the term wins above replacement player is only loosely captured in our measure. A replaceability analysis could help answer questions like “how much does A having the disc instead of B in this situation improve the probability of scoring?”

Finally, EC-like measures could be developed in other sports as an attempt to quantify the value of non-scoring actions and attribute it to players. Integrating EC analysis and replaceability analysis, e.g. using propensity score methods, could lead to even more accurate and telling measures.

6. Conclusion

We introduce a new ultimate measure, expected contribution (EC), that evaluates how much a player contributes to the team probability of winning points. We characterize how much a unit EC affects the probability and show applications in player ranking and team strategy. EC is constructed based on conditional probability estimators, which can also be used to visualize spatial strengths and weaknesses. We suggest conditional probability estimators could be used in other sports to attribute value to non-scoring actions and improve player evaluation.

Figure 1. LOESS estimate of probability of scoring as a function of field position. Carets denote the playing field, and a score occurs between the right two carets (the endzone). Black denotes the home team and red denotes the away team. Solid lines are probability of scoring the point, dashed lines are the probability of scoring the possession.

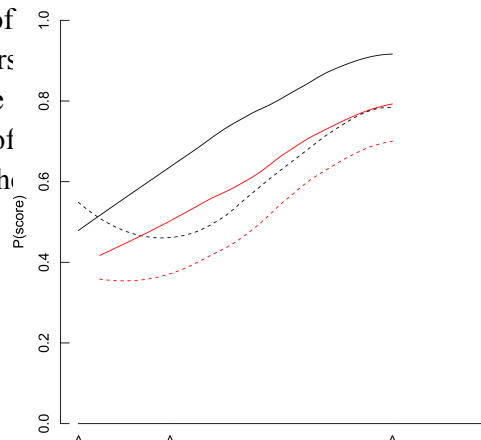


Figure 2. k-nearest neighbor estimate of probability of scoring as a function of field position with k=100. Central location of the disc does not seem important statistically.

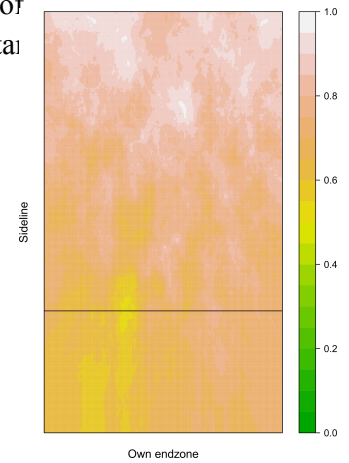


Figure 3. Probability of scoring (top), probability of completion (bottom left), and expected value of throw (bottom right). The probability of the each team scoring given they possess the disc at a particular location is given in the top two graphs. The expected value graph is determined by weighing the probability of scoring graphs by the probability of completion from the blue dot (bottom left at each location). The best throw location in this example is backwards toward the middle of the field (called a reset or dump).

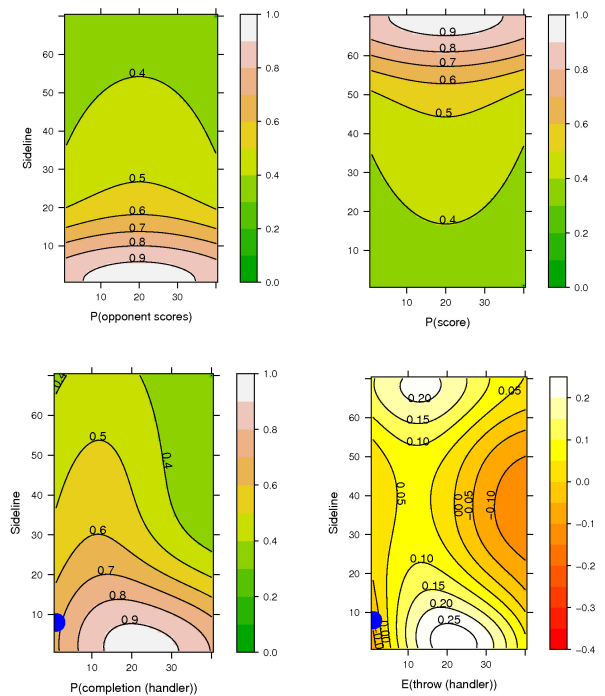


Figure 4. Expected value graphs under prevent defense (left) and pressure defense (right). The minimax outcome in the prevent defense graph, indicating that the defense should choose it, and the offense should counter with a short forward pass with expected value around 0.3.

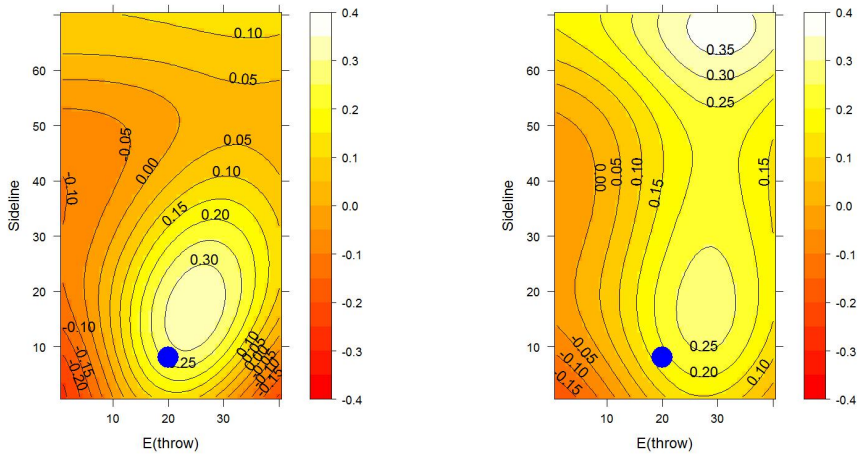
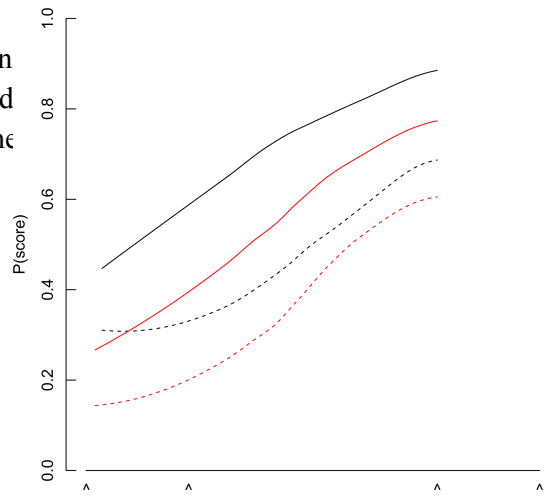


Figure 5. LOESS estimate of probability of scoring as a function of field position (women's). Black denotes the home team and red denotes the away team. Solid lines are probability of scoring the point, dashed lines are the probability of scoring the possession.



Acknowledgments

We gratefully acknowledge Ultiworld, UltiApps, and USA Ultimate for their continuing contribution and support.

References

- Jam it in. <Accessed 6 January 2014>
<http://ultiworld.com/2013/02/06/jam-it-in-new-ultiworld-scoring-charts-show-nexgens-endzone-dominance>
- Hollinger J. *Pro Basketball Forecast 2005-2006*. Potomac Books Inc, 2005.
- Weiss J and Childers S. Maps for Reasoning in Ultimate. *Machine Learning in Sports Analytics Workshop at the European Conference in Machine Learning*, Prague, CZ 2013. [\[link\]](#)
- Win Probability. <Accessed 6 January 2014>
<http://www.advancednflstats.com/2008/08/win-probability.html>

Appendix

See attached .xls/zip file.