Conference on Predictive Inference in Sports

June 4-5, 2025 Iowa State University







The conference is made possible by a gift from Iowa State University Professor Emeritus David Harville who established the C. R. Henderson Fund for Excellence in Predictive Inference and Its Applications.



C.B. Herderom

## Day 1 - June 4, 2025

Time	Event	Speaker	Title		
8:00	Breakfast				
8:45	Opening Remarks	Jamie Pollard			
9:00	Session I	Denise Bradford	Game-Changing Intelligence: Integrating Athlete Signals, Schedule Density, and Rookie Development within a Predictive Inference Framework		
9:30		Mark Broadie	Ranking Systems for Golf and Other Sports		
10:00		Nathan Sandholtz	An Experiment to Investigate the Spatial Component of Serving Strategy in Tennis		
10:30	Break				
11:00	Keynote	Luke Bornn	From Pixels to Points: Predicting Performance in Professional Sports		
12:00	Lunch				
1:15	Session II	Katerina Wu	Which League is Best? Using Paired Comparison Models to Estimate Hockey League Strength and Project Player Performance		
1:45		Michael Schuckers	Estimating Player Age Curves Using Regression and Imputation		
2:15		Namita Nandakumar	Twitter Is Real Life? Translating Student Research To Team Decision-Making		
3:00	Jack Trice Stadium Tour				
4:00	Poster Session and Appetizers				

### Day 2 - June 5, 2025

Time	Event	Speaker	Title	
8:00	Breakfast			
9:00	Session III	Suraj Bhuva	Medal Projections and Performance Benchmarks in Olympic Sports	
9:30		Monnie McGee	Layered Dirichlet Modeling to Assess the Changing Contributions of MLB Players as they Age	
10:00		Jim Albert	Prediction in Baseball – Understanding Streaky Hitting and the Increase in Home Run Hitting	
10:30	Break			
11:00	Keynote	Sara Ziegler	Stats for Normies: How to Explain Predictive Inference in Sports to an Innumerate Public (or Even Just Your Boss)	
12:00	Lunch			
1:15	Session IV	Sameer Deshpande	Expected Hypothetical Completion Probability	
1:45		Ron Yurko	NFL Ghosts: A Framework for Evaluating Defender Positioning with Conditional Density Estimation	
2:15		Quang Nguyen	Step-and-Turn Modeling of Individual Player Movement in American Football	
2:45	Closing			

Talks (listed alphabetically by presenting author (\*) last name)

# Prediction in Baseball – Understanding Streaky Hitting and the Increase in Home Run Hitting

Jim Albert\*

#### Abstract

An important aspect of baseball analytics is prediction of future player performance since teams are constantly involved in obtaining new players through the draft, trades and free agency. We illustrate several applications of Bayesian prediction. Although there is a popular belief in streaky ability, is there evidence that there is more streakiness in hitting than one could predict based on coin-flipping models? In the Statcast era (2015 through 2024) there have been substantial increases in home run hitting. We use prediction to explore how much of the home run increase is due to player behavior and how much is due to changes in the ball construction.

#### Medal Projections and Performance Benchmarks in Olympic Sports Suraj Bhuva<sup>\*</sup> and Dan Webb

#### Abstract

Certain Olympic sports suffer from a lack of data due to the small number of competitions leading up to an Olympic Games. A side effect of this is that supervised learning predictions have high variance due to the lack of input data. This talk will demonstrate an improved method for modeling performance outcomes in best mark sports, particularly in instances where there is a small amount of input data.

Existing prediction methods utilized a supervised learning model on a window of observations over a specified period leading into an Olympic Games. The output of this model was then used to parameterize a probability distribution from which Monte Carlo simulations were run to predict athlete placements.

Under the new approach, we define a rolling window for data leading to an Olympic Games and then implement an ensemble learning method whereby we progressively stack various supervised learning models from past Olympic Games. In order to predict the next Olympic Games, we aggregate m \* n predictions for an athlete *i* (where *m* is the number of Olympic Games in the dataset and *n* is the number of supervised learning models trained on each Games). We then run Monte Carlo simulations by sampling from athlete *i*'s distribution which is parameterized by the mean and variance of these m\*n predictions to then determine athlete medal probabilities, with the resulting calibration plots and Brier scores showing an improvement over existing methods.

In addition, Performance Funnels are a way to provide benchmarks for talent development. Adapting existing research on the topic, we show a preview of how we have created these benchmarks and provide some guidance on potential use cases for this framework.

#### From Pixels to Points: Predicting Performance in Professional Sports Luke Bornn\*

#### Abstract

In this talk I will explore how modern data sources are used to measure player performance and predict the resultant impact on team success. By blending advanced spatio-temporal models with techniques from causal inference, we are able to predict player skill far better than either individual tool allows. Using optical tracking data consisting of hundreds of millions of observations, I will demonstrate these ideas by characterizing defensive skill in NBA players and decision making in EPL soccer players.

Game-Changing Intelligence: Integrating Athlete Signals, Schedule Density, and Rookie Development within a Predictive Inference Framework. Denise Bradford\*

#### Abstract

Modern sports analytics necessitates an advanced approach that extends beyond traditional metrics, incorporating comprehensive athlete-monitoring data—including training load, neuromuscular readiness, and physiological markers—alongside fixture congestion and developmental trajectories. This integrated presentation aims to predict performance, mitigate injury risk, and optimize roster composition under conditions of uncertainty (Gelman et al., 2013).

#### Ranking Systems for Golf and Other Sports Mark Broadie\*

#### Abstract

This talk introduces a new ranking system, Strokes Gained to Points (SG2P), for ranking teams or players in a given sport. The system will be described mainly in the context of golf, but it applies to many others sports, including tennis, baseball, football, etc., and applies to individual and team sports. SG2P is the first ranking method that is fair, unbiased and includes a premium for winning (more precisely, a premium for above average performance). The SG2P ranking system is currently used to rank college teams and players in golf and is used to rank professional disc golfers in the Professional Disc Golf Association (PDGA).

#### **Expected Hypothetical Completion Probability**

Sameer Deshpande\* and Kathy Evans

#### Abstract

Using high-resolution player tracking data made available by the National Football League (NFL) for their Big Data Bowl competition, we introduce the Expected Hypothetical Completion Probability (EHCP), a objective framework for evaluating plays. At the heart of EHCP is the question "on a given passing play, did the quarterback throw the pass to the receiver who was most likely to catch it?" To answer this question, we first built a Bayesian non-parametric catch probability model that automatically accounts for complex interactions between inputs like the receiver's speed and distances to the ball and nearest defender. While building such a model is, in principle, straightforward, using it to reason about a hypothetical pass is challenging because many of the model inputs corresponding to a hypothetical are necessarily unobserved. To wit, it is impossible to observe how close an un-targeted receiver would be to his nearest defender had the pass been thrown to him instead of the receiver who was actually targeted. To overcome this fundamental difficulty, we propose imputing the unobservable inputs and averaging our model predictions across these imputations to derive EHCP. In this way, EHCP can track how the completion probability evolves for each receiver over the course of a play in a way that accounts for the uncertainty about missing inputs.

# Layered Dirichlet Modeling to Assess the Changing Contributions of MLB Players as they Age

Monnie McGee\*, Jacob A. Turner, and Bianca A. Luedeker

#### Abstract

The productive career of a professional athlete is limited compared to the normal human lifespan. Most professional athletes have retired by age 40. The early retirement age is due to a combination of age-related performance and life considerations. While younger players typically are stronger and faster than their older teammates, older teammates add value to a team due to their experience and perspective. Indeed, the highest-paid major league baseball players are those over the age of 35. These players contribute intangibly to a team through mentorship of younger players; however, their peak athletic performance has likely passed. Given this, it is of interest to learn how more mature players contribute to a team in measurable ways. We examine the distribution of plate appearance outcomes from three different age groups as compositional data, using Layered Dirichlet Modeling (LDM). We develop a hypothesis testing framework to compare the average proportions of outcomes for each component among 3 of more groups. LDM can not only determine evidence for differences among populations, but also pinpoint within which component the largest changes are likely to occur. This framework can determine where players can be of most use as they age. Twitter Is Real Life? Translating Student Research To Team Decision-Making Namita Nandakumar

#### Abstract

We will take a critical eye to my past public projects on NHL trends and inefficiencies and explore what seven years of research within the NFL/NHL have taught me about framing and solving sports analytics problems with a team-centric lens.

#### Step-and-Turn Modeling of Individual Player Movement in American Football Quang Nguyen<sup>\*</sup> and Ron Yurko

#### Abstract

In continuous team sports, player tracking data have driven significant advancements in the task of player evaluation. We present a framework for modeling the movement of an individual ball carrier within a play in the National Football League. Inspired by the animal movement literature, we characterize and model frame-level player movement based on two components: step length (distance between successive locations) and turn angle (angle between successive displacement vectors). In particular, we propose Bayesian mixed-effects models for both step length and turn angle that demonstrate careful consideration for their dependence and respective response distribution. We also account for dynamic tracking data features describing the relationships between players on the field, along with relevant player and team random effects. This approach offers practical insight into player evaluation, as it enables us to generate hypothetical player movement paths in any given play. Specifically, we draw posterior predictive step length and turn angle for a player at each frame, and update spatial covariates needed to simulate the next frame. This ultimately allows us to compare the observed player trajectory with the distribution of simulated paths via common valuation metrics in American football.

#### An Experiment to Investigate the Spatial Component of Serving Strategy in Tennis

Nate Sandholtz<sup>\*</sup>, Ron Hager, Stephanie Kovalchik, and Gil Fellingham

#### Abstract

We conducted an experiment with the Brigham Young University Men, s Tennis Team to investigate the spatial component of serving strategy in tennis Serve data including precise spatial coordinates of bounce locations were collected for eight collegiate athletes with known intended targets for each serve Using these data we estimate each players execution error defined as the distribution of bounce locations around the intended targets Because many serves are unobserved due to contact with the net we explicitly account for censoring when modeling these distributions The resulting estimates allow us to assess whether players stated targets align with their actual behavior as indicated by the centers of the estimated serve distributions We extend the analysis by estimating playerspecific optimal aiming locations for both first and second serves To account for the twostage nature of the problem players are permitted a second serve if they fail to hit the service region on their first serve we formulate the decision as a twoperiod Markov decision process MDP Solving this MDP requires estimates of the point win probability surface over the service area which we obtain using methods developed in prior work We compare the estimated optimal targets to the players stated targets and discuss cases of alignment and divergence.

#### Estimating Player Age Curves Using Regression And Imputation Michael Schuckers<sup>\*</sup>, Mike Lopez, and Brian Macdonald

#### Abstract

The impact of player age on performance has received attention across sport. Most research has focused on the performance of players at each age, ignoring the reality that age likewise influences which players receive opportunities to perform. Our manuscript makes two contributions. First, we highlight how selection bias is linked to both (i) which players receive the opportunity to perform in sport, and (ii) at which ages we observe these players perform. This approach is used to generate underlying distributions of how players move in and out of sport organizations. Second, motivated by methods for missing data, we propose novel estimation methods of age curves by using both observed and unobserved (imputed) data. We use simulations to compare several comparative approaches for estimating aging curves. Imputation-based methods, as well as models that account for individual player skill, tend to generate lower RMSE and age curve shapes that better match the truth. We implement our approach using data from the National Hockey League.

# Which League is Best? Using Paired Comparison Models to Estimate Hockey League Strength and Project Player Performance $Katerina Wu^*$

#### Abstract

Comparing hockey players from around the world and assessing how they may perform when transitioning to another hockey league is a complicated problem. Due to the different playing styles and opponent difficulty, there is not one consistent metric to make comparable evaluations of player performance for hockey leagues around the world. In this project, we introduced a new method for comparing and projecting player performance across leagues using an adjusted z-score metric. This metric controls for factors such as age, league, season, and position that affect a player's P/PG metric, and could be applied to any league of interest. After constructing a dataset of over a million pairwise comparisons of the performances of each player across multiple leagues using the adjusted z-scores calculated for each observation, we model the difference in player performance across seasons conditional on the leagues in which they played, testing several different regression methods including logistic regression, OLS regression, and Bradley-Terry Models. The coefficients from the models create a ranking of leagues determined by their strength. By using all comparisons from all the leagues to estimate league strength rather than using only league-to-NHL and/or NHL-to-league transitions, we incorporate a greater amount of information for a more accurate estimation of league strength.

#### NFL Ghosts: A Framework for Evaluating Defender Positioning with Conditional Density Estimation

Ronald Yurko\*, Quang Nguyen, and Konstantinos Pelechrinis

#### Abstract

Player attribution in American football remains an open problem due to the complex nature of twenty-two players interacting on the field, but the granularity of player tracking data provides ample opportunity for novel approaches. In this work, we introduce the first public framework to evaluate spatial and trajectory tracking data of players relative to a baseline distribution of "ghost" defenders. We demonstrate our framework in the context of modeling the nearest defender positioning at the moment of catch. In particular, we provide estimates of how much better or worse their observed positioning and trajectory compared to the expected play value of ghost defenders. Our framework leverages high-dimensional tracking data features through flexible random forests for conditional density estimation in two ways: (1) to model the distribution of receiver yards gained enabling the estimation of within-play expected value, and (2) to model the 2D spatial distribution of baseline ghost defenders. We present novel metrics for measuring player and team performance based on tracking data, and discuss challenges that remain in extending our framework to other aspects of American football.

Stats For Normies: How To Explain Predictive Inference In Sports To An Innumerate Public (Or Even Just Your Boss) Sara Zieqler\*

#### Abstract

Statisticians live in a world of probabilities and so understand what they mean. The general public does not. But for a new sports model or metric to be adopted, it must be understood by people who struggle with predictive inference. How can data analysts go about explaining their sports work to the rest of the world? I will discuss pitfalls in the understanding of such common predictive tools as win probability, and offer suggestions for communicating this work to people who are not as familiar with what it all means. **Posters** (listed alphabetically by presenting author (\*) last name)

## **2010** Association of Tennis Professionals Matches by Court Surface *Emilie Butler*\*

#### Abstract

This analysis examines whether the surface of a tennis court affects a player's chance of winning. Using data from the Association of Tennis Professionals' 2010 matches reveals no significant relation between the player's chance of winning on hard, clay, or grass courts. Logistic regression yielded a high standard error, but a Bayesian logistic regression led to more confidence in the results. These results show a player's true athletic ability given the court changes.

#### Analyzing the Effects of NBA Head Coaches

Andrew Cannon<sup>\*</sup>, Jared Fisher, Garritt Page, and Gilbert Fellingham

#### Abstract

Evaluating the value-added of coaches in the NBA is a challenging task as the coaches with the best win/loss records often have the best players. This prompts a question of attribution: if two coaches had the same roster, which one would win? This paper attempts to answer this question by introducing a method for quantifying coaching effect in the NBA. We propose a method for isolating the effect of a coach's in-game scheme on their team's probability of winning a game while controlling for other factors, namely the relative strength of the two competing teams. To control for team strength, player performance metrics are aggregated into "Team-Adjusted VORP Difference" or  $\Delta t$ VORP, meant to account for the difference in quality of on court product between both teams. We model each coach's win probability as a function of  $\Delta t$ VORP using Bayesian Additive Regression Trees. We find that monotonicity constraints improve the model fit based on out-of-sample prediction accuracy. In comparing coaches' win probability curves, we find some of the winningest coaches are close to average, while other coaches are found to be truly great contributors to their teams.

#### Predicting Transfer Portal Success for College Football Wide Receivers Lucas Libero\*

#### Abstract

With the recent changes to the athlete transfer rules there have never been more college football players transferring between schools. This has created a heated competition between college football programs to scout players from other programs and bring them in and hope they perform as well or better than they did before. This project aims to use Wins Above Average and PFF player grades to create a model that accurately predicts transfer portal player success for college football wide receivers and provides feature selection so that NCAA teams can more effectively conduct their scouting. PFF is a sports analytics company that started in professional football but has spread to college football as well. They collect play-by-play data for nearly every major football game in the United States. This project uses data obtained from PFF's Academia Program that works with students and professors at universities to connect applied sports analytics with academia. A variety of regression and machine learning approaches were applied to model player success after transferring using metrics from before they transferred. The models showed up to a 65% accuracy, with the most accurate predictions coming from using both WAA and offense grade as a response.

#### The Behavioral Impact of Anticipated and Actual Football Game Outcomes: Analyzing Unhealthy Food and Alcohol Consumption Patterns of Collegiate Fans S.T. Nestler<sup>\*</sup>, J.D. Boas, K.M. Pekar, J.R. Draves, and Y. Chang

#### Abstract

Football game outcomes can significantly impact fans' emotions and behaviors, influencing their consumption patterns of unhealthy food and alcohol. This study investigates how both predicted and actual game outcomes affect these behaviors among collegiate football fans. This research explores the psychological and emotional drivers behind consumption patterns, especially when game outcomes either align with or deviate from fan expectations.

The study focused on one collegiate football team from the Southeastern Conference (SEC) during the 2023 regular season. Transaction data from venues near the stadium were analyzed to assess food and alcohol consumption patterns. The team's actual game results and predicted outcomes, estimated based on rankings, past matchups, and performance, were modeled to test the hypotheses.

A series of 2 (predicted outcome: victory vs. loss) 2 (actual outcome: victory vs. loss) between-subjects ANOVAs were conducted in R. Results showed that predicted victories led to significantly higher unhealthy food consumption, gross sales, and alcohol consumption compared to predicted losses. Consumption peaked when a predicted victory resulted in an actual loss. Conversely, the lowest consumption levels were observed when both predicted and actual outcomes were losses.

Further implications of how fan expectations and actual game outcomes interact to drive unhealthy consumption, thereby informing strategic marketing campaigns, inventory planning, fan engagement efforts, and community health initiatives will be presented.