

Oral Presentation Abstracts

rMetrics: A Statistically Motivated Framework for Player Evaluation Using Residualized Scores

Robert Bajons[†], Vienna University of Economics and Business; Lucas Kook, Vienna University of Economics and Business

[†] E-mail: *robert.bajons@wu.ac.at*

To efficiently assess and identify undervalued players, it is crucial to measure a player’s skills accurately. In dynamic games such as soccer, basketball, or ice hockey, a popular quantitative approach to evaluating player performance involves comparing an actual outcome to an expected outcome estimated by a statistical (or machine learning) model. Examples of this approach are goals (saved) above expectation (soccer and ice hockey), or shooter impact (basketball). Typically, analysts rely on flexible machine learning models capable of handling complex structures present in sports analytics when estimating expected outcomes. While machine learning models are superior to traditional statistical models in terms of predictive ability, using them for inferential tasks such as player evaluation is often inappropriate due to potential biases and a lack of uncertainty quantification.

In this work, we present a unifying framework for metrics of the above type. We first show that, when using parametric models to estimate the expected outcome, these metrics are directly related to score test statistics. Hence, valid statistical inference is possible under parametric assumptions. Motivated by this finding, we propose a natural extension to this framework using residualized versions of the original metrics. In this way, valid uncertainty quantification can be achieved when using machine learning models. Furthermore, we relate the proposed procedure to player-specific effect estimates in interpretable semiparametric models. This allows for a different view on and a deeper understanding of these popular player evaluation metrics. We present various use cases of our framework.

Reward Systems in Sports: Who’s the Fairest of Them All?

Benjamin S. Baumer[†], Smith College, Northampton, MA; Sarah Susnea, Smith College, Northampton, MA

[†] E-mail: *bbaumer@smith.edu*

One active thread of sports analytics research proposes mechanisms for awarding points, dollars, or other rewards that are more “fair”. However, we don’t always articulate a shared understanding of what constitutes fairness. In recent years, analysis of various statistical criteria for non-discrimination in the machine learning community has yielded some fascinating incompatibility results. Namely, Kleinberg, et al. (2016) prove that in all but the most trivial of cases, an algorithm cannot be unbiased across all reasonable definitions (of fairness). We apply this lens to sports, and explore how various existing reward systems embody different notions of fairness.

The Impact of Skating Speed and Style with Tracking Data

Meghan Chayka[†], Joe Gratz, Jeff Goeree

[†] E-mail: *meghan@stathletes.com*

Skating speed and style are fundamental drivers of hockey performance, shaping both defensive resilience and offensive effectiveness. This study leverages advanced player and puck tracking data to investigate how skating attributes influence game outcomes across defensive and offensive phases. Our dataset comprises 289 American Hockey League (AHL) games from the 2024–25 season (October 2024 – April 2025). Each game contains approximately 85,000–105,000 puck positions and 800,000–1,000,000 player positions, enabling fine-grained spatial and temporal analysis. Importantly, the tracking data is aligned with event-level outcomes (e.g., shots, goals, passes, takeaways), allowing us to assess the consequences of skating behaviors on scoring opportunities. First, we examine defensive skating styles, ranging from aggressive, high-pressure systems to passive containment and their impact on shot suppression and expected goals against (xGA) at both even strength (5v5) and during penalty kill (5v4). We assess whether aggressive defensive pressure reduces shot volume while concentrating risk into fewer but higher-quality chances and explore how these tradeoffs vary across game states. Second, we investigate how player skating characteristics, including acceleration, top-end speed, and edge control, influence offensive zone entry success. We test whether superior skaters generate more clean entries, odd-man rushes, and rush-based expected goals, and whether these advantages translate into faster zone establishment on the power play. This study highlights the indispensable role of tracking data in advancing hockey analytics and tactical understanding.

Fast Algorithm for Calculating Probability of Chess Winning Streaks

Guoqing Diao[†], George Washington University, Washington, DC

[†] E-mail: *gdiao@gwu.edu*

Motivated by the controversy in the chess community, where Hikaru Nakamura, a renowned grandmaster, has posted multiple impressive winning streaks over the years on the online platform chess.com, we derive the probabilities of various types of streaks in online chess and/or other sports. Specifically, given the winning/drawing/losing probabilities of individual games, we derive the probabilities of “pure” winning streaks, non-losing streaks, and “in-between” streaks involving at most one draw over the course of games played in a period. Numerical studies demonstrate the proposed algorithm is fast and accurate.

Expected Pass Value (xPV): A Holistic Framework for Evaluating Passing Situations in Soccer

Tobias Harringer[†], Vienna University of Economics and Business; Robert Bajons, Vienna University of Economics and Business

[†] E-mail: *Tobias.Harringer@wu.ac.at*

Passes are the most frequent events in soccer. Yet, to this day, analyzing passes statistically poses a substantial challenge. Existing metrics often lack context (e.g., successful passes) or are heavily outcome-dependent beyond the passer’s influence (e.g., assists, expected assists). In this work, we leverage spatio-temporal tracking data provided by PFF to gain deeper insights into passing. In particular, we develop a framework for valuing any real or hypothetical pass on the pitch based on two well-established components: (1) a pitch control model, and (2) a pitch value model. For (1), we derive a computationally efficient model to determine the likelihood of controlling a location on the pitch based on players’ velocity, movement direction, and angular mismatch to the location of interest. To estimate (2), we employ an XGBoost model based on a rich set of features that are derived from tracking data and capture situational context. Obtaining a value for every possible pass enables a holistic analysis of passing situations. More precisely, we can quantify the passing ability and decision-making of offensive players, and simultaneously the positional play of defenders. We apply our framework to obtain an expected pass value (xPV) for all passes from the 2022 World Cup. Results show that xPV outperforms existing metrics, as demonstrated by higher correlations with players’ future performance indicators and market value. We further analyze defenders by their ability to prevent xPV and quantify offensive decision-making by comparing the actual xPV to alternative passing options.

What Influences the Field Goal Attempts of Professional Players?

Guanyu Hu[†], The University of Texas Health Science, Houston, TX

[†] E-mail: *nealguanyu@gmail.com*

Basketball shot charts provide valuable information regarding local patterns of in-game performance to coaches, players, sports analysts, and statisticians. The spatial patterns of where shots were attempted and whether the shots were successful suggest options for offensive and defensive strategies as well as historical summaries of performance against particular teams and players. The data represent a marked spatio-temporal point process with locations representing locations of attempted shots and an associated mark representing the shot’s outcome (made/missed). Here, we develop a Bayesian log Gaussian Cox process model allowing joint analysis of the spatial pattern of locations and outcomes of shots across multiple games. We build a hierarchical model for the log intensity function using Gaussian processes, and allow spatially varying effects for various game-specific covariates. We aim to model the spatial relative risk under different covariate values. For inference via posterior simulation, we design a Markov chain Monte Carlo (MCMC) algorithm based on a kernel convolution approach. We illustrate the proposed method using extensive simulation studies.

A case study analyzing the shot data of NBA legends Stephen Curry, LeBron James, and Michael Jordan highlights the effectiveness of our approach in real-world scenarios and provides practical insights into optimizing shooting strategies by examining how different playing conditions, game locations, and opposing team strengths impact shooting efficiency.

Tackling Causality: Estimating Frame-Level Defensive Impact with Multi-Agent Transformers

Ben Jenkins[†], Florida Atlantic University, Boca Raton, FL

[†] E-mail: *benrossjenkins@gmail.com*

Evaluating defensive performance in American football remains fundamentally challenging due to the indirect, distributed nature of defensive impact. Traditional metrics like tackle counts fail to capture causal contributions, and standard correlational models cannot answer counterfactual questions about defensive value. We propose a novel causal inference framework that estimates frame-level Conditional Average Treatment Effects (CATE) of tackling using multi-agent Transformer models trained on NFL tracking data.

We formulate tackling as a binary treatment at the defender level with Expected Points Added (EPA) as the outcome. By conditioning on high-dimensional spatiotemporal player tracking data, our model isolates each defender’s causal impact throughout a play. Our architecture employs multi-head attention mechanisms to address interference between units, sequential modeling to handle time-varying confounding, and a doubly robust objective combining outcome prediction and propensity estimation. We incorporate MMD and adversarial balancing techniques to ensure representation equilibrium across treatment conditions.

Our frame-level CATE estimates reveal previously hidden defensive contributions: anticipatory positioning, tackle opportunity creation, counterfactual movement value, tackle quality evolution, and missed opportunities. These temporal patterns provide dynamic insights into when and how defenders influence play outcomes, beyond who ultimately records the tackle. This work represents the first integration of causal inference with transformer-based multi-agent modeling at the frame level in football, introducing a statistically rigorous foundation for defensive evaluation that aligns with the sequential, interactive nature of the sport.

Digital Health in Sport Science

Marcos Matabuena[†], Harvard University, Boston, MA

[†] E-mail: *mmatabuena@hsph.harvard.edu*

The monitoring of physical activity through wearable devices has enabled the collection of a vast amount of physiological, biomechanical, and environmental variables. These data provide quantitative information about an athlete’s condition, allowing, for example, the assessment of fatigue over time during both training sessions and competitions.

The efficient use of the collected information for each athlete has the potential to enhance individual performance over time, optimizing their performance during key competition periods, which ultimately determine their overall results. In general, the collected data are in the form of time series with different sampling frequencies, recorded over multiple days. These data often exhibit a functional and multilevel structure, adding complexity to their analysis and interpretation. However, this complexity also opens new opportunities for applying advanced data analysis techniques, providing competitive advantages at both the individual and team levels.

The objective of this talk is to illustrate how statistical models in functional spaces, capable of handling multilevel structures, can become key tools in professional sports. These models have been successfully applied in the past, for example, to predict maximum oxygen consumption ($\text{VO}_2 \text{ max}$), as well as in the medical field to evaluate the progression of patients with chronic diseases.

Finally, in the last part of the talk, the author will discuss the role of generative artificial intelligence models in sports practice, exploring their potential and challenges in performance optimization and training personalization

The Best of Both Worlds: Predicting Football Coverages with Supervised and Unsupervised Learning

Rouven Michels[†], Bielefeld University, Germany; Robert Bajons, Vienna University of Economics and Business, Austria; Jan-Ole Koslik, Bielefeld University, Germany

[†] E-mail: *r.michels@uni-bielefeld.de*

Choosing between man and zone coverage is one of the most critical strategic decisions a defensive coordinator must make before each offensive play in American football. While experienced offensive coordinators and quarterbacks often rely on visual cues to recognize these defensive schemes, the growing availability of player tracking data opens up new possibilities for uncovering such tactics. In this project, we first use an XGBoost to predict whether a defense is in man or zone coverage based on the positions of all players once both teams are set. In a second step, we incorporate additional information from pre-snap player movements. Specifically, we apply a hidden Markov model (HMM) to capture defenders' movement trajectories, using hidden states to represent the offensive players they might be covering. By including summary statistics from the decoded state sequences - such as average entropy - as features in our pre-snap motion model, we significantly enhance the model's predictive power regarding coverage schemes. However, our approach goes beyond mere classification. It can also be used to explore broader tactical questions, such as how pre-snap motion helps offenses identify defensive coverages. To this end, we compute both pre- and post-motion coverage probabilities (with the latter derived from the HMM) for each team, allowing us to assess how effectively these models anticipate the actual defensive schemes.

NFL CoD: A Bayesian circular mixed-effects model for explaining variability in directional movement in American football

Quang Nguyen[†], Carnegie Mellon University, Pittsburgh, PA; Ronald Yurko, Carnegie Mellon University, Pittsburgh, PA

[†] E-mail: quang@stat.cmu.edu

Change of direction is a key element of player movement in American football, yet there remains a lack of objective approaches for in-game performance evaluation of this athletic trait. Using tracking data, we propose a Bayesian mixed-effects model with heterogeneous variances for assessing a player’s ability to make variable directional adjustments while moving on the field. We model the turn angle (i.e., angle between successive displacement vectors) for NFL ball carriers on both passing and rushing plays, focusing on receivers after the catch and running backs after the hand-off. In particular, we consider a von Mises distribution for the frame-level turn angle and explicitly model both the mean and concentration parameters with relevant spatiotemporal and contextual covariates. Of primary interest, we include player random effects that allow the turn angle concentration to vary by ball carrier nested within position groups. This offers practical insight into player evaluation, as it reveals the shiftest ball carriers with great variability in turning behavior. We illustrate our approach with results from the first nine weeks of the 2022 NFL regular season and explore player-specific and positional differences in turn angle variability.

Personnel-Adjustment for Home Run Park Effects in Major League Baseball

Jason A Osborne[†], North Carolina State University, Raleigh, NC; Richard A Levine, San Diego State University, San Diego, CA

[†] E-mail: jaosborn@ncsu.edu

In Major League Baseball, every ballpark is unique, each with its own dimensions and climate. Quantifications of home run friendliness for these ballparks abound, but are often based on a limited number of seasons and typically do not include uncertainty assessment. Further, there don’t seem to be any that incorporate personnel effects from individual players. We consider a generalized linear mixed effects model for home run counts taking as our observational unit the combination of game and batter-pitcher handedness matchup. The Poisson distribution provides a good fit to counts observed in the 2010 to 2024 seasons, involving more than 2.5 million plate appearances. Our model allows for four theoretical mean home run frequency functions for each ballpark. We model personnel effects by constructing “elsewhere” measures of batter ability to hit home runs and pitcher tendency to give them up, using data from parks other than the one in which the response is observed. The fitted model facilitates estimation of the mean number of home runs hit in a game-matchup, adjusted to a team with average batters and pitchers appearing with average batter-pitcher handedness frequency. Standard errors are easily obtained for these adjusted means or for ratios of log-intensities for pairwise comparison of ballparks. In our analysis, we find

that the estimated home run frequencies adjusted to average personnel are substantially different from observed home run frequencies, leading to considerably different ballpark rankings than often appear in the media.

A Paradox of Blown Leads: Rethinking Win Probability in Team Sports

Jonathan Pipping[†], University of Pennsylvania, Philadelphia, PA; Jiaoyang Huang, University of Pennsylvania, Philadelphia, PA; Abraham Wyner; University of Pennsylvania, Philadelphia, PA

[†] E-mail: *jpipping@wharton.upenn.edu*

In roughly half of evenly matched games, the team that ends up losing is, at some stage, a heavy favorite to win. This surprising and seemingly paradoxical pattern of eventual losers appearing dominant mid-game challenges conventional interpretations of win-probability graphics, which have become ubiquitous across sports broadcasts and media. While in-game win probability estimates are easily interpretable as the probability that a team wins from a particular game state, their full trajectories across a game are anchored by the observed outcome; that is, every winner’s curve must finish at 100%, and every loser’s at 0%. Since the win probability trajectories of winners and losers fundamentally differ, any post-game analysis of their attributes (such as their maximum and minimum) is inherently biased and requires that we first condition on the eventual outcome of the game.

Accordingly, we investigate blown leads by analyzing the distribution of a single statistic: the maximum win probability attained by the losing team. Simulations reveal that in roughly half of evenly matched contests, the loser attained a win probability of at least 66.7%, a finding we confirm with NFL and NBA play-by-play data. Leveraging Donsker’s invariance principle, we recast calibrated win-probability paths as Brownian motions and derive closed-form expressions for the full distribution of these maxima as a function of relative team strength. These theoretical curves align closely with simulations and empirical data, providing a principled baseline for quantifying anomalous collapses and comebacks.

Do Behavioral Considerations Cloud Soccer Penalty-Kick Location-Optimization? Game Theory, GAM, and Lasso Analysis

Ava Uribe, Syracuse University, Syracuse, NY; Shane Sanders[†], Syracuse University, Syracuse, NY; Justin Ehrlich, Syracuse University, Syracuse, NY; James Reade, University of Reading, Reading, England; Carl Singleton, University of Stirling, Stirling, Scotland

[†] E-mail: *sdsander@syr.edu*

For soccer penalty-kicks, classical/statistical game-theoretic models predict a non-negative relationship between goal area partition-conditional shot-volume and conversion-rate for a goal-optimizing

penalty-kick taker. However, we find a strong negative relationship for our study data of 536 penalty-kicks from 2020 UEFA Champions and Europa Leagues. Estimated indifference-sets and underlying inferential statistics of linear, polynomial, and ML-regularized Lasso regression models indicate penalty-takers significantly value both conversion-rate and on-target rate when locating PK-shots. They behave as hybrid decision-makers who deviate from optimal PK-locating strategies in a manner consistent with the behavioral-valuation of optics of highly-skilled play by limiting likelihood of missing the goal entirely as might a novice-player. Players are revealed to value optics of performance rather than strictly performance-optimization. Optimization utility is estimated as 3.12 times as important on margin as behavioral utility from estimated indifference sets. Models are highly-explanatory: optimization and behavioral factors explain approximately 85.2% of shot-location for top professional players. Results also suggest PK-takers are risk-averse, as their partition-dependent shot-volume increases at a decreasing rate in conversion-rate. As the purpose of a PK-attempt is to maximize expected goal (likelihood) directly, presence of risk-aversion here represents another behavioral factor.

College Football Volatility: A Bayesian State-Space Model of the Transfer Portal and NIL Impact

Ronald Yurko[†], Carnegie Mellon University, Pittsburgh, PA; Luke Benz, Harvard T.H. Chan School of Public Health, Boston, MA

[†] E-mail: *ryurko@andrew.cmu.edu*

The landscape of American college football has changed dramatically in recent years with conference realignment, increased usage of the transfer portal, and the seismic legal ruling regarding athletes’ name, image, and likeness (NIL). In this work, we use a Bayesian state-space model to capture the impact of transfers and NIL on the volatility of team strengths over time. Specifically, we extend the classic autoregressive process for team strength by modeling the between-season innovation variance as a function of incoming transfers and NIL rule changes. This approach enables us to predict greater variance for team strengths in an upcoming season based on roster changes. Our results from the playoff era (2014-2024) reveal variation between schools that have embraced the volatility of transfers (e.g., Deion Sanders at Colorado) with those who have been reluctant to change (e.g., Iowa and Clemson). We explore variability in the effects between different positions of incoming transfers, and compare the predictive performance of our novel approach with simpler state-space models.

Ball path curvature and in-game free throw shooting proficiency in the National Basketball Association

Ruoqian (Judy) Zhu[†], Massachusetts Institute of Technology, Cambridge, MA; Dave Love, NBA Shooting Coach; Scott Powers, Rice University, Houston, TX

[†] E-mail: *judy05@mit.edu*

Basketball shooting coaches agree that smoother shooting motions are better, but there is less agreement about what “smooth” means quantitatively or what part of the shooting motion needs to be smooth. Using ball tracking data from the 2023-2024 National Basketball Association regular season, we explore the relationship between ball path curvature and free throw shooting performance. We fit Bezier curves to the ball tracking data in the sagittal plane and test different methods of calculating path curvature. We find that both max curvature and terminal curvature are negatively associated with shooting performance, but terminal curvature explains much more of the between-player variance in free throw shooting performance. This suggests that shooting coaches would be better off focusing on the smoothness at the end of the shot rather than at the beginning of the forward motion of the ball.

Poster Presentation Abstracts

Clustering Aging Curves to Classify Athlete Development and Predict Career Trajectories

David Awosoga, University of Waterloo, Waterloo, ON; Yushi Liu[†], University of Waterloo, Waterloo, ON; Samuel Wong, University of Waterloo, Waterloo, ON

[†] E-mail: *y239liu@uwaterloo.ca*

Athletes reach elite performance at highly variable rates. Some progress rapidly, while others develop more gradually over time. In this project, we model and cluster the rate of progression of Olympic track and field athletes using their aging curves. Unlike team sports, individual sports such as athletics allow for precise comparison through standardized scoring, as individual results are more easily measured in objective terms. To compare athletes across different event disciplines such as jumps and throws, we convert raw performances to a unified scale using the scoring system designed by World Athletics, the world governing body for athletics.

We apply functional data analysis to smooth each athlete’s aging curve and extract dominant patterns using Functional Principal Component Analysis (FPCA). Because competition frequency varies by athlete and event, we address irregular and sparse observations using FPCA methods suited for unevenly spaced data. We extend the work from Cavan et.al (2025) to account for selection bias, since more successful athletes tend to have longer careers.

Clustering the resulting FPCA scores allows us to group athletes based on how fast they improve with age, identifying profiles such as early risers, steady performers, and late bloomers. These clusters provide both descriptive insights into developmental patterns and predictive tools for forecasting when rising athletes might reach elite standards based on early-career data.

This focus on the rate of progression enables a more adaptable and predictive framework allowing coaches, performance analysts, and high performance directors to analyze athletic development.

A League-Normalized Plus-Minus Approach in Soccer for Talent Identification

Gabriel Reis Gama Barbosa[†], Gemini Sports; Amod Prasad Sahasrabudhe, Gemini Sports

[†] E-mail: *gabriel@geminisports.ai*

Plus-minus metrics have historically faced skepticism in soccer analytics due to challenges like low scoring, limited substitutions, and lineup collinearity. Using event data only, this study shows that, with careful methodological design, an xG-based Estimated Plus-Minus (EPM) can be both reliable

and insightful, even capturing defensive contributions typically overlooked by metrics like Atomic VAEP.

The methodology begins with xG-differential as the response variable in an adjusted Regularized Adjusted Plus/Minus (RAPM) model. Games are divided into finer-grained “stints” (periods between substitutions, goals, red card) to minimize collinearities and add context. A Statistical Plus-Minus (SPM) model is trained by using position-specific aggregated metrics to predict previously obtained RAPM metrics. Finally SPM is used as prior to stabilize EPM estimates.

The league translation leverages player transfer data. In a RAPM-style model, league adjustment coefficients are estimated using regression targets based on EPM differences between the originating and destination leagues. EPM scores from over 30 international leagues and multiple seasons are then normalized using these coefficients.

This metric is applied by clubs for player recruitment and talent identification. For researchers, the results demonstrate: 1. empirically stable EPM values across contexts; 2. defensively-informative estimations that complement existing metrics like Atomic VAEP; and 3. meaningful league translation adjustments that reduce bias in cross-league comparisons. Our approach offers a thoughtful alternative to the view that plus-minus metrics are unsuitable for soccer, providing a rigorous and standardized framework for objective player evaluation across diverse competitive environments.

Modeling NBA Production Curves using Bayesian non-parametric Concavity Priors

Abhijit Brahme[†], UCSB, Santa Barbara, CA

[†] E-mail: *abhijitbrahme@ucsb.edu*

In this work, we propose a Bayesian non-parametric model for jointly inferring how markers of skill co-evolve over an NBA player’s career. Our model explicitly accounts for multiple sources of variability in the metrics by accounting for player similarity, dependence between metrics and temporal dependence. We develop a novel approach which constrains the inferred aging curves to be concave, which facilitates interpretation and enables simple comparison of peak age performance across players. To capture similarity across players, we learn a latent embedding for each player which is associated with expected production curves across multiple metrics. We illustrate that these latent embeddings are useful in forecasting how young, high-value prospects (i.e high draft picks) will develop across multiple metrics, giving a more holistic projection of athletes’ career trajectories.

Unlocking NFL Value: A New Era of Player Evaluation with Partial Membership Clustering Models

Rayaan Damani[†], Rice University, Houston, TX

[†] E-mail: *rd71@rice.edu*

Modern NFL player roles are increasingly nuanced, yet traditional clustering methods often fail to reflect this complexity. We present an innovative application of partial membership models, leveraging 2020-2024 NFL running back and wide receiver performance data, to address this limitation. Our novel methodology allows players to simultaneously belong to multiple performance archetypes, a significant departure from conventional rigid categorizations. By integrating a multi-distribution framework (Negative Binomial, Poisson, Gamma, Beta, Normal) to precisely model varied metrics, we effectively identify core player types, quantify meaningful player similarities (e.g., 98.2% similarity between Tyler Allgeier and Jonathan Taylor), and pinpoint “hidden gems” exhibiting elite, multi-faceted skills such as Tyrone Tracy Jr. and Jerome Ford. Our analysis of 223 running backs and 192 wide receivers reveals a powerful opportunity for NFL teams: identifying statistically comparable players at considerably reduced costs, with potential annual salary cap savings exceeding \$ 15 million. This research offers critical data-driven insights to revolutionize player evaluation, roster construction, and talent acquisition within the evolving NFL landscape.

Evaluating NBA Player Consistency through GARCH Modeling

Joshua Davila[†], Texas A&M University, College Station, TX

[†] E-mail: *joshua.davila12@tamu.edu*

Player performance is arguably the most prominent indicator of success for a professional sports franchise. Whether a player has received all-NBA honors or plays a specific role off the bench, front offices seek to find players who can play their role at a consistent level. However, many of the standard measures of consistency don’t appropriately measure performance over a season. One solution is the use of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) modeling to evaluate player consistency. However, there have been limited uses of this approach in a sport context (Dolmeta et al., 2023). The dataset analyzed for this study features over 26 thousand NBA regular season games (collected using the “hoopR” package in R). Only players who average 10 minutes a game for at least 65 games were included in the modeling. All players were then sorted into a ranking to determine the relative consistency compared to the other players. The player ranked the least consistent for the 2023-24 NBA season was Tyrese Haliburton of the Indiana Pacers (volatility constant = 65.9201), while the Phoenix Suns’ Drew Eubanks was the most consistent (volatility constant = 0.00002) Using GARCH modeling allows teams to not only evaluate a player’s consistency but also forecast future performance. When used in tandem with models such as VAR (Vector Autoregression) and GAS (Generalized Autoregressive Score), it allows for front offices to assess a player’s total body of work while accounting for the time series nature of a season.

Analyzing Team-Sponsor Dynamics in Formula One Racing: A Social Network Analysis Approach

Grace Dickman[†], Toyota North America, Dallas, TX; Jonathan Jensen, Texas A&M University, College Station, TX; Joe Cobbs, Northern Kentucky University, Highland Heights, KY

[†] E-mail: *dickmangrace@gmail.com*

Sponsorship networks that are strategically designed through portfolio management (Kim et al., 2014) and roster design (Cornwell & Kwon, 2020) can realize benefits to brand equity on all sides of the network – inclusive of sponsors and sponsees (Groza et al., 2012) – and enhance business opportunities through structured networking platforms (Cobbs, 2011). Collectively, this informs that conceptualizing sport sponsorship from a network perspective is vital to advancing theory and practice. The sponsee-sponsor-audience relationship does not exist in isolation (Groza et al., 2012). The purpose of this study is to determine whether such networks are heterogeneous, resulting in insights for those who study and practice sponsorship-linked marketing.

The current study produced a complete network map of team sponsorship in Formula One motor racing. The network is undirected and weighted, where edges are weighted by the length of the sponsorship in years. For sponsors, Puma has the highest eigenvector centrality (0.338), meaning it is the most influential sponsor in the network, sharing strong connections with other influential teams and sponsors. The team with the largest degree is McLaren (39), however, Ferrari had the highest weighted degree (253). Ferrari has the highest betweenness centrality (8,150) which is the ability to facilitate connections with other sponsors. McLaren has the highest eigenvector centrality (1), meaning they are not just the most influential team but the most influential member of the entire network. While these are only preliminary results, results do support the heterogeneity of sponsor networks across similar teams in one sport.

No Toyota data or methodology was used in this research.

Wind vs 175 Grams of Plastic: Exploring Weather Effects in Professional Ultimate Frisbee

Braden Eberhard, University of Pennsylvania, Philadelphia, PA; Jacob Miller[†], Shown Space, Salt Lake City, UT

[†] E-mail: *shownspace.analytics@gmail.com*

The recent growth of analytics in professional ultimate frisbee has significantly advanced our understanding of team and player performance. However, current modeling approaches overlook one of the sport’s most influential contextual factors: weather. Wind can alter gameplay by increasing throw difficulty, lowering completion rates, and reducing scoring probability, which complicates fair evaluation across games. Anecdotally, players and coaches frequently cite weather, particularly wind, as a major factor in poor outcomes, often attributing missed throws and strategic adjustments to environmental conditions rather than executional errors.

To address this gap, we construct a novel dataset by integrating the Ultimate Frisbee Association’s (UFA) play-by-play data with historical weather records. Results reveal that higher wind speeds are associated with a shift in throw selection away from medium-distance attempts toward shorter throws, while the frequency of long throws remains largely unchanged. Wind speed also exhibits a measurable impact on throw completion probability, with the effect intensifying at longer distances. Incorporating wind-related covariates into generalized additive models (GAMs) for both scoring probability and individual throw completion significantly improves overall model fit, uncovers substantial nonlinear effects, and maintains or slightly improves classification metrics.

By introducing weather as a covariate in game-state models, this study addresses a critical gap in existing analytical approaches and offers a more context-aware framework for evaluating performance. This approach enables more accurate player contribution and efficiency metrics and supports better-informed strategic decision-making in professional ultimate.

Quantifying Which NBA Coaches Get the Most Out of Their Players

Shane Faberman[†], University of North Carolina, Chapel Hill, NC

[†] E-mail: *shanetyler2005@gmail.com*

In the NBA, coaches play a crucial role in game strategy, player development, and managing team dynamics. However, quantifying coaching impact is challenging, as it is difficult to isolate a coach’s influence from that of their players. Unlike the plethora of statistics available for players, coaching performance is assessed far more subjectively. This study introduces a new metric, Box Plus-Minus (BPM) Over Expected (BOE), that evaluates coaches based on how their players perform relative to expectations, aiming to identify those who consistently maximize their players’ potential. To calculate BOE, Expected BPM (EBPM) was first computed for each player-season. Let a player’s age n season represent the season when they were n years old. EBPM was derived by adjusting a player’s BPM from their age $n - 1$ season based on average aging trends for qualifying players and the deviation from the league-average qualifying player at age $n - 1$. BOE was then calculated as the difference between a player’s BPM and EBPM in their age n season. According to BOE, media rankings tend to overrate championship-winning coaches. This is likely due to underrating the effect of superstars such as LeBron James and Stephen Curry. Championships and regular-season win-loss records depend on many factors not under a coach’s control. BOE provides a more nuanced assessment of a coach’s system’s impact by shifting the focus toward player performance relative to expectation. It serves as a valuable tool for evaluating coaching effectiveness.

Going Beyond Aggregates: A Dynamic Rating System for NFL Pass Rushers and Blockers

Max Gebauer[†], University of Pennsylvania, Philadelphia, PA; Victoria Lee, University of Pennsylvania, Philadelphia, PA; Kenny Watts, University of Pennsylvania, Philadelphia, PA; Jonathan Pipping, University of Pennsylvania, Philadelphia, PA

[†] E-mail: *gebauerm@sas.upenn.edu*

In American Football, the outcome of a pass rush is extremely impactful for game outcomes. Good pass blocking allows the quarterback time to progress through their reads, while good pass rushing limits a quarterback’s time in the pocket, resulting in more favourable defensive outcomes (hurries, sacks, turnovers, etc.). Traditional evaluation methods rely heavily on these aggregate statistics, which do not account for variability in matchup difficulty and limit fair comparison within positions. To overcome this limitation, we introduce a dynamic rating system that leverages NFL tracking data to update player ratings after each blocker-rusher interaction.

Our model distinguishes multiple outcomes – pass blocker wins, losses, hurries, hits, and sacks – while incorporating adjustments for penalties and many-to-one matchups. Each outcome weight is determined by existing expected-points models. Preliminary results indicate that our model produces a more representative ranking of players compared to traditional stationary metrics, while also allowing for probabilistic matchup prediction and capturing changes in player ability over time.

Understanding the Stabilization Points of NBA RAPM Ratings

Hua Gong[†], Rice University, Houston, TX

[†] E-mail: *hg37@rice.edu*

Regularized Adjusted Plus-Minus (RAPM) is a widely used metric in sports for evaluating player impact, as it accounts for teammates and opponents. A critical aspect of RAPM is understanding stabilization points at which a statistic becomes reliable and less influenced by noise. This study investigates the stabilization points of RAPM ratings in the NBA. Using a split-half correlation framework, we explore how three key factors influence stabilization: the number of possessions played, the degree of collinearity in player lineups, and the ratio of noise variance to talent variance. Through both theoretical derivations and empirical simulations, we demonstrate that stabilization is delayed when players frequently appear with the same teammates, leading to high collinearity and reduced identifiability of individual contributions. We further show that a higher noise-to-signal ratio, arising from either high game-to-game variability or low underlying talent variance, elevates the stabilization threshold. These dynamics are examined across a range of outcome variables, including traditional plus-minus, offensive and defensive rebounding, and foul rates. While RAPM estimates for high-variance outcomes like plus-minus require thousands of possessions to stabilize, more stable metrics such as rebounding reach acceptable reliability levels with substantially fewer possessions. Our findings provide practical guidance for analysts and decision-makers on the sample

size requirements for interpreting RAPM estimates and suggest when caution is warranted due to insufficient data.

Quantifying the Impact of NBA Hustle Plays on Regularized Adjusted Plus-Minus

Dylan Hahami[†], St. John's University, Queens, NY

[†] E-mail: *dhahami5@gmail.com*

Hustle has been a core part of the game of basketball for decades, representing the grit and determination of a player or team through selfless plays that do not show up on the traditional box score. Since the 2015-16 season, the NBA has tracked these high-effort, low-glory actions as hustle stats, finally quantifying the intangible coaching observation into workable data. This study investigates the relationship between NBA hustle statistics and Regularized Adjusted Plus-Minus (RAPM), a robust measure of a player's overall impact on team performance. Using player data from 2017 to 2024, we incorporate publicly available hustle stats into regression and machine learning models to examine their marginal contributions to RAPM, controlling for conventional box score stats. We explore whether hustle metrics independently predict positive on-court impact or serve as proxies for other undervalued traits like defensive positioning or effort consistency. Results will aim to explore implications for scouting, lineup construction, and advanced analytics in professional basketball.

Movement Dynamics in Elite Female Soccer Athletes: The Quantile Cube Approach

Jan Hannig, Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC; Kendall Thomas[†], Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC

[†] E-mail: *ketho@unc.edu*

This work presents an innovative adaptation of existing methodology to investigate external load in elite female soccer athletes using GPS-derived movement data from 23 matches. We developed a quantitative framework to examine velocity, acceleration, and movement angle across game halves, enabling transparent and meaningful performance insights. By constructing a quantile cube to quantify movement patterns, we segmented athletes' movements into distinct velocity, acceleration, and angle quantiles. Statistical analysis revealed significant differences in movement distributions between match halves for individual athletes. Principal Component Analysis (PCA) identified anomalous games with unique movement dynamics, particularly at the start and end of the season. Dirichlet-multinomial regression further explored how factors like athlete position, playing time, and game characteristics influenced movement profiles. This approach provides a structured method for analyzing movement dynamics, revealing external load variations over time and offering insights into performance optimization. The integration of these statistical techniques demonstrates the potential of data-driven strategies to enhance athlete monitoring in soccer.

Exploratory Network Analysis, Causal Interpretation, and a Transfer Destinations Predictor for College Football Transfer Portal

Jie He[†], University at Buffalo–SUNY, Buffalo, NY

[†] E-mail: *jhe53@buffalo.edu*

The combination of College Football Transfer Portal, NIL (Name, Image, and Likeness), and revenue sharing is creating the collegiate version of professional transfer market. Many student-athletes utilize the portal to achieve their goals and chase their dreams. Many CFB programs utilize the portal to improve their rosters and stay relevant in the race towards College Football Playoff. Because the portal is fairly new, people who do not closely follow collegiate sports know little about it. This study aims at using exploratory network analysis and machine learning methods to explore the most recent dataset (Season 2021 to 2025) from the network perspective. Network analysis has been used in broader contexts but has yet to be widely applied to NCAA CFB. A comprehensive network analysis on the CFB transfer network is an extension of the research direction. Temporal network analysis, community detection, and link prediction were employed to show the trends over time and the clear hierarchy in the portal among different levels of CFB. Meanwhile, prediction models specific to NCAA transfer destinations are insufficient. We created an Extreme Gradient Boosting model to predict CFB transfer destinations which was deployed to the public and supposed to be first-of-its-kind. Finally, causal interpretation of the quality of incoming transfers in CFB has not been investigated in previous research. We used Elastic Net Regression to identify the factors. Our findings can effectively benefit the public, researchers, and policymakers in sports management (e.g., player recruitment and retention).

Evaluating NHL Point Production Forecasting Methods for Russian Skaters

Erica James[†], Bryn Mawr College, Bryn Mawr, PA

[†] E-mail: *ecejames16@gmail.com*

One of the best ways to understand the skills of a prospect is through international tournaments among their age group and skill level as it allows for direct comparison against prospects in more familiar development systems. For recently drafted Russian prospects, this has not been possible for several years for a variety of reasons. This project attempts to simulate the first five years in the NHL of Russian skaters drafted in the first three rounds of the NHL Draft from 2020-2023 by comparing their weighted pre-NHL performance to that of Russian NHL players from 2004-2012 with similar development paths. All data was collected from QuantHockey and analysed in R using the forecast and tidyverse. This resulted in six forecasting models on point production based on position and a logistic regression for likelihood of reaching the NHL. I found that the logistic regression had a McFadden value approaching 1 while the most effective forecasting models for forwards and defenders were within 2 standard deviations from the dataset. This project shows that statistical models can give insight into outcomes from particular development systems in cases where

comparative viewing is limited. It does not however, negate the necessity of watching prospects as numbers alone are unable to identify all of the strengths and weaknesses of a particular prospect.

A Comps-Based Approach for Interpreting Tree-Based Predictions With an NFL Draft Application

Elisabeth Millington[†], Rice University, Houston, TX; Scott Powers, Rice University, Houston, TX

[†] E-mail: *emm16@rice.edu*

Random forests are a powerful yet opaque machine learning predictive model. This poses challenges for interpretability, particularly in environments where decisions can be very high-stakes, like professional sports. In this paper, we present a method for interpreting random forest predictions in a sports setting through the lens of player comps. Player comps are a common practice in player scouting, where prospects are evaluated based on their perceived similarity to past players. We can define interpretable similarity scores that quantify how much each training observation contributes to a given prediction by leveraging the connection between random forests and adaptive nearest neighbors. A player’s similarity scores can be viewed as quantifiable player comps, and the random forest’s prediction can be recovered by taking a weighted average of the outcome variable among those comps, with weights determined by the similarity scores. We apply this methodology to evaluate quarterback prospects in the 2025 National Football League Draft using ESPN’s Total Quarterback Rating (QBR) per season. We show that our approach captures meaningful structure in the data while providing interpretability, allowing us to identify comps for top prospects such as Cam Ward, Jaxson Dart, Shedeur Sanders, and Dillon Gabriel.

The Causal Role of Timeouts in Disrupting Scoring Runs in the NBA

Jiahe Niu[†], Boston University, Boston, MA; Jacob Park, Boston University, Boston, MA; Judith Lok, Boston University, Boston, MA

[†] E-mail: *jniu@bu.edu*

This study investigates the causal effect of calling a timeout during an opposing team’s seven-point scoring run in NBA basketball games. Using play-by-play data from the 2017-2018 and 2018-2019 NBA regular seasons, we apply propensity score modeling and inverse probability of treatment weighting (IPTW) to account for the following confounding variables: scoring team location (home/away), quarter, minutes remaining in quarter, score margin, and average player lineup quality measured by Box Plus-Minus (BPM) for both teams. Our outcome variable is the net point differential over the two minutes immediately following the 7-point run, capturing the relative scoring performance between the run team (team that scored 7-0) and the timeout-calling team. Our analysis reveals that calling a timeout during an opponent’s 7-0 run is associated with a small but statistically significant improvement in timeout-calling team’s subsequent performance, with an Average Treatment Effect (ATE) of 0.272 points (95% CI: 0.076 to 0.468, $p = 0.007$) and

an Average Treatment Effect on the Treated (ATT) of 0.310 points (95% CI: 0.115 to 0.505, $p = 0.002$) over the next two minutes of play. We also observe variation in timeout effectiveness across NBA teams, with some teams showing positive effects while others demonstrate negative effects. The results indicate that coaches should employ timeouts as an intervention to halt opponent runs, while recognizing that the primary benefit is momentum disruption rather than significant scoring improvement.

Does Furosemide Make Horses Run Faster?

Michele Sezgin[†], Carnegie Mellon University, Pittsburgh, PA; Ron Yurko, Carnegie Mellon University, Pittsburgh, PA; Joel Greenhouse, Carnegie Mellon University, Pittsburgh, PA

[†] E-mail: *msezgin@andrew.cmu.edu*

In 2023, the Horseracing Integrity and Safety Commission banned the use of furosemide within 48 hours of a race for two-year-old and stakes races. Though furosemide has been used in thoroughbred racing since the 1960s to treat exercise-induced pulmonary hemorrhage in horses, there has been speculation about the drug’s performance-enhancing effects. We study a perfectly-matched dataset of 28,617 two-year-old horses that have raced both with and without furosemide to estimate the causal impact of furosemide on race-day performance, as measured by probability of finishing in the money. When controlling for public betting odds, the fraction of total horses using furosemide in a given race, track, post position, and sex of horse, we find that using furosemide increases a two-year-old’s odds of finishing in the money by 55% on average. Future work will investigate the race-day performance impacts of furosemide in a larger population of horses matched along relevant characteristics.

Quantifying Pressing Effectiveness Through Expected Threat (xT) Transitions

Matthew Spivey[†], University of Pennsylvania, Philadelphia, PA; Adam Kuechler, University of Pennsylvania, Philadelphia, PA; Christopher Bugliosi, University of Pennsylvania, Philadelphia, PA; Jonathan Pipping; University of Pennsylvania, Philadelphia, PA

[†] E-mail: *spiveym@wharton.upenn.edu*

Effective pressing is universally recognized as an essential part of modern soccer. However, it remains a difficult action to evaluate with precision. Most existing metrics rely on simple event counts or changes in possession without capturing the broader impact of pressure on a team’s ability to advance the ball or create chances. In this project, we develop a continuous, spatial framework to assess pressing through changes in Expected Threat (xT) resulting from pressure events. We evaluate how pressure alters ball movement and value, whether by disrupting buildup, forcing a turnover, or limiting an opponent’s forward progress.

Our approach uses tracking data from the past three English Premier League seasons. We estimate expected goals (xG) using kernel density estimation, then proceed with conditional kernel density estimation (CKDE) to model how the ball moves from one location to another. These transition probabilities are combined with xG estimates to build a continuous xT model, allowing for more accurate evaluation of presses involving small movements or occurring farther from goal. Using this framework, we identify which players generate the most cumulative xT value through pressing and which are most efficient on a per-press basis. To better understand press quality, we also regress pressing outcomes on the skill of the player in possession.

By capturing the spatial and tactical nuances of pressure, our model provides an interpretable and practical tool for evaluating off-ball defensive contributions, offering insights for analysts, scouts, and coaching staffs.

A Comparison of Batter Performance and League Quality Across Domestic T-20 Cricket Leagues

Matthew Stuart[†], Loyola University Chicago, Chicago, IL; Gregory Matthews, Loyola University Chicago, Chicago, IL; Leigha DeRango, Loyola University Chicago, Chicago, IL; Hassan Rafique, Syracuse University, Syracuse, NY

[†] E-mail: *mstuart1@luc.edu*

This paper investigates batter performance across major domestic Twenty20 (T20) cricket leagues to understand how player effectiveness varies by competition context. Using a comprehensive dataset comprising individual match-level batting records from the Indian Premier League (IPL), Big Bash League (BBL), Caribbean Premier League (CPL), Pakistani Super League (PSL), and South Africa 20 (SAT), we fit a multinomial regression model for the number of runs scored per bowled ball accounting for batter and league conditions as well as in-match effects. A hierarchical modeling framework utilizing Bayesian methodology is employed to quantify both player-specific ability and league-specific correlations, allowing for quantifiable player comparisons across leagues and of the qualities of the leagues themselves. Our findings will offer insights into talent evaluation and the transferability of performance across T20 competitions.

Decoding Winning Patterns in Professional Basketball via Attention-Based Sequence Models

Joshua Zhang[†], Harvard University, Cambridge, MA; Amy Dong, Harvard University, Cambridge, MA

[†] E-mail: *JoshZhang@college.harvard.edu*

The NBA off-season is one of the most pivotal moments for franchises, as front offices decide whether to rebuild in pursuit of a championship. These decisions hinge on future team performance, yet prior

research lacks reliable team-based forecasting models to predict team win totals. We utilize publicly available player-level data spanning the 2016-2025 NBA seasons to develop two transformer-based neural network models to predict the overall win record per team for the regular and playoff seasons. Leveraging the architecture’s ability to capture long-range dependencies across player attributes and interactions, we represent each team as a sequence of up to 15 players, mirroring the size of an NBA roster. Each player is represented by a comprehensive set of features from the previous regular season, including statistical totals, per-game averages, per-36-minute metrics, draft position, age, physical measurements, player origin, shooting breakdowns, and other advanced analytics (e.g., BPM, VORP). Using a 80-20 (8-2 NBA Seasons) train-test split, our regular season model achieves a mean absolute error (MAE) of 7.14 and 6.49 on the training and test set, respectively. Our playoff season model achieves a MAE of 2.69 and 2.12 on the training and test set, respectively. These results are competitive given the NBA’s volatility, where small differences in win totals determine playoff qualification or seeding. Our model offers a practical tool to forecast roster impact, aiding decisions around trades, free agents, or the draft.