

Oral Presentation Abstracts

Contributions of Carl Morris in Sports Analytics

Jim Albert, Bowling Green State University, Bowling Green, OH

† E-mail: *albert@bgsu.edu*

Carl Morris was well-known for his pioneering research in Bayesian multiparameter inference and prediction. Morris was also known for his development of statistical thinking and methodology in sports. This talk provides an overview of Morris' contributions in sports. This includes Morris' experience in sports as a youth, summaries of some of Morris' best-known contributions using sports data, his influence working with students, and some of Morris' thinking about the interplay of statistics and sports.

Here Comes the STRAIN: Analyzing Defensive Pass Rush in American Football with Player Tracking Data

Quang Nguyen and Ronald Yurko, Carnegie Mellon University, PA; Gregory J. Matthews, Loyola University Chicago, IL

† E-mail: *nmquang@cmu.edu*

In American football, a pass rush is an attempt by the defensive team to disrupt the offense and prevent the quarterback (QB) from completing a pass. Existing metrics for assessing pass rush performance are either discrete-time quantities or based on subjective judgment. Using player tracking data, we propose STRAIN, a novel metric for evaluating pass rushers in the National Football League (NFL) at the continuous-time within-play level. Inspired by the concept of strain rate in materials science, STRAIN is a simple and interpretable means for measuring defensive pressure in football. It is a directly-observed statistic as a function of two features: the distance between the pass rusher and QB, and the rate at which this distance is being reduced. Our metric possesses great predictability of pressure and stability over time. We also fit a multilevel model for STRAIN to understand the defensive pressure contribution of every pass rusher at the play-level. We apply our approach to NFL data and present results for the first eight weeks of the 2021 regular season. In particular, we provide comparisons of STRAIN for different defensive positions and play outcomes, and rankings of the NFL's best pass rushers according to our metric.

When Data Meets Reality: Augmenting Sports Videos with Visualizations

Zhutian Chen, Tica Lin, Johanna Beyer, and Hanspeter Pfister, Harvard University

† E-mail: *ztchen@seas.harvard.edu*

Augmented sports videos have become increasingly popular among sports enthusiasts worldwide. They combine embedded visualizations and video effects to present data overlaid on videos to effectively communicate sports insights. However, creating these videos can be challenging, requiring significant time and video editing skills. In this talk, I will present a set of human-AI systems that ease visualizing data into sports videos to enhance data experiences for sports games. These systems enable natural interactions, such as using natural language or gaze, to interact with in-game data, catering to experts, general audiences, and fans alike. I will share ongoing research endeavors aimed at creating a more immersive, interactive, and enjoyable data-driven game-watching experience for the future.

A Weighted Curve Clustering Approach for Analyzing Pass Rush Routes in American Football

Robert Bajons and Kurt Hornik, Institute for Statistics and Mathematics, Vienna University of Economics and Business

† E-mail: *robert.bajons@wu.ac.at*

We present a weighted K-means approach for clustering weighted curves, i.e. curves which may be assigned weights at each observation of the curve. The methodology is applied to routes of defending players in American football (obtained from NFL big data bowl on Kaggle), where the aim is to automatically detect effective pass rushing routes from specific players or teams. Each route of a defensive player can be assigned a weight, which at each time point represents the probability of pressuring the quarterback. The weights may be derived by a machine learning model of choice (gradient boosted trees are used in this work). Since pass rush routes vary in length due to the variability in the duration of each play, the weighted route data is first preprocessed using Bézier curves, such that each trajectory is of the same length. Results demonstrate that the methodology used is able to cluster pass rushing routes effectively and much better than a classical (unweighted) K-means approach. The resulting clusters are finally used for various team and player analyses of pass rush plays.

Analytics, have some humility: a statistical view of fourth down decision making

Ryan Brill and Abraham Wyner, University of Pennsylvania

† E-mail: *ryguy123@sas.upenn.edu*

Expected points (EP) and win probability (WP) are value functions fundamental to strategic in-game decision making and player evaluation in American football. The EP and WP functions which

are widely used today are statistical models fit from historical data. These models, however, are subject to serious statistical flaws: selection bias, overfitting, ignoring autocorrelation, and ignoring uncertainty quantification. We mitigate selection bias in EP models by including measures of team quality as covariates, and we mitigate overfitting using a catalytic prior. Further, we conduct a simulation study to understand the effect of autocorrelation on WP models. Although WP models fit from autocorrelated data are largely unbiased, we find that autocorrelation significantly inflates WP standard errors. Finally, we alter the existing fourth down decision procedure to incorporate uncertainty quantification in win probability models by bootstrapping the fourth down decision itself. We find that decision making changes substantially. In particular, we find that that far fewer fourth down decision are as obvious as analysts claim.

NFL Ghosts: Evaluating pass defense with high-dimensional CDEs

Ron Yurko, Carnegie Mellon University, Pittsburgh, PA; Kostas Pelechrinis, University of Pittsburgh, Pittsburgh, PA

† E-mail: *ryurko@andrew.cmu.edu*

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 a framework for evaluating the ability of pass defenders to limit yards gained after the catch (YAC) by opposing receivers relative to expected defenders known as “ghosts”. This consists of two components: (1) an estimate for a receiver’s YAC distribution, and (2) an estimate for a “ghost” defender’s 2D position distribution. We model both distributions using flexible non-parametric, high-dimensional conditional density estimate techniques to account for rich features derived from player-tracking data. We discuss the challenges of modeling “ghost” defenders, exploring sensitivity to which moments of time and features are conditioned on. We are then able to compare how a defender’s observed positioning at the moment the ball is caught affects the estimated distribution of the receiver’s YAC relative to “ghost” defenders. By modeling the full YAC and “ghost” distributions, we are able to compute multiple novel ways of assessing player performance such as the difference in expected points or probability of certain events, e.g., obtaining a first down or scoring a touchdown. Although we focus on pass defense, we discuss how this approach can be extended to other aspects of American football for measuring a player’s leverage within a play and the difficulties in extending for continuous-time assessment.

Yellow fever: investigating referee consistency in the 'Big 5' men’s European football leagues

Pete Philipson, Newcastle University, UK.

† E-mail: *peter.philipson1@newcastle.ac.uk*

Scrutiny of referees in football is ever-increasing and likely at an all-time high. Is this fair? How consistent are modern day referees? In this work the number of yellow cards given to the home and

away teams over four seasons of data from 2018-2022 in each of the ‘Big 5’ European leagues in men’s football is modelled. This allows us to make an assessment of the affect of playing behind closed doors due to the pandemic as well on heterogeneity amongst both referees, primarily, and teams, secondarily. The underdispersed nature of the data from the small counts, and likely dependence of the cards issued within a game, leads us to a bivariate Conway-Maxwell-Poisson copula model to analyse the data. We also model home advantage, and whether this was diluted during Covid-19, and league effects, allowing an assessment of how consistent referees are across leagues and which league has the most heterogeneous men in the middle. The underlying model could be deployed in a range of other sports when the researcher is analysing small counts with low variability.

How much does Home Field Advantage matter in Soccer Games?

Guanyu Hu, University of Missouri, Columbia, MO

† E-mail: nealguanyu@gmail.com

In many sports, it is commonly believed that the home team has an advantage over the visiting team, known as the home field advantage. Yet its causal effect on team performance is largely unknown. In this paper, we propose a novel causal inference approach to study the causal effect of home field advantage in English Premier League. We develop a hierarchical causal model and show that both league level and team level causal effects are identifiable and can be conveniently estimated. We further develop an inference procedure for the proposed estimators and demonstrate its excellent numerical performance via simulation studies. We implement our method on the 2020-21 English Premier League data and assess the causal effect of home advantage on eleven summary statistics that measure the offensive and defensive performance and referee bias. We find that the home field advantage resides more heavily in offensive statistics than it does in defensive or referee statistics. We also find evidence that teams that had lower rankings retain a higher home field advantage.

Contextualized Generative Ghosting Model as Benchmark for Evaluating Player Movement in Football

Chaoyi Gu and Varuna De Silva, Loughborough University, UK; Mike Caine, Warwick University, UK; Ben Smith, Breakaway Data, UK

† E-mail: c.gu@lboro.ac.uk

Performance analysis models have been built on tracking and event data in football to evaluate players’ movements in the game. However, training ghosting players from both attacking and defending sides to simulate the game possession in given context and using the ghosting model as benchmark for real-life performance evaluation has been an open research question until now. Our research aims to build a contextualized data-driven ghosting model utilizing deep learning algorithms and assess the players’ movements based on the trained benchmarking ghosts. We

build a conditional Variational Recurrent Neural Network to predict the sequences of space-control heatmaps from each team in specific game context and develop a contextualized Convolutional Neural Network to transform the predicted sequences of heatmaps to players’ trajectories. The predicted players’ trajectories with space-control heatmaps are used to generate contextualized benchmarking ghosts for real-life player performance evaluation. The results suggest that this novel approach is interpretable, objective, and efficient in evaluating players’ movements in the game. The great performance of the model can be accredited to the rich amount of contextual information learnt by the ghosts while they are trained on the large datasets, and the ability of deep generative model and convolutional neural network in learning insightful patterns from high-dimensional data. Our research introduces a novel method to evaluate the football player’s in-game movements, which is believed to be a useful addition to state-of-the-art metrics for helping coach make data-oriented decisions.

Estimating individual contributions to team success in women’s college volleyball

Scott Powers, Luke Stancil and Naomi Consiglio, Rice University, Houston, TX

[†] E-mail: *sp161@rice.edu*

The progression of a single point in volleyball starts with a serve and then alternates between teams, each team allowed up to three contacts with the ball (this sequence is called a “volley”). Using charted data from the 2022 NCAA Division I women’s volleyball season (4,147 matches, 600,000+ points, more than 5 million recorded contacts), we model the progression of a point as a Markov chain with the state space defined by the sequence of contacts in the current volley and additional context variables (e.g. the team’s current rotation). We estimate the probability of each team winning the point, which changes on each contact. We attribute changes in point probability to the player(s) responsible for each contact, facilitating measurement of performance on the point scale for different skills. Traditional volleyball statistics do not allow apples-to-apples comparisons across skills, and they do not measure the impact of the performances on team success. For adversarial contacts (serve/receive and attack/block/dig), we estimate a multinomial regression model for the outcome, with random effects for the players involved and fixed effects for context (e.g. origin location of the set leading to the attack); and we adjust performance for strength of schedule not only on the conference/team level but on the individual player level. We answer practical questions for volleyball coaches, e.g. What is the point value of using a substitution to have a defensive specialist in the back row?

A Bayesian two-stage framework for lineup independent assessment of individual rebounding in the NBA

Nicholas Kiriazis and Christian Genest, McGill University, Montreal, QC; Alexandre Leblanc, University of Manitoba, Winnipeg, MB

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

In basketball, existing methods to assess individual rebounding ability are problematic because they depend on all players present on the court rather than just the player of interest. Although there exist modelling approaches to correct for this dependence, they are generally unsuitable for events with binary outcomes. We propose a Bayesian, two-stage model for predicting conditional rebounding rates in the NBA. Although similar in flavour to the popular APM framework, it is different in that it does not assume that individual contributions are linearly additive on the response scale. Furthermore, we improve the regularization approach by using rebounding-specific heuristics. After defining the model, a simulation study is performed to verify its effectiveness, and the parameters are then estimated using data from the 2020–21 NBA season. Predictions are then made for rebounding in the 2021–22 season. It was found that there are some players who are excellent at stealing rebounds away from the opposing team without collecting them themselves, whereas others are simply collecting rebounds that their team has already secured. These subtleties are not captured by relying on traditional rebounding metrics.

Evaluating plate discipline in Major League Baseball with Bayesian Additive Regression Trees

Ryan Yee and Sameer Deshpande, University of Wisconsin–Madison, Madison, WI

[†] E-mail: *ryee2@wisc.edu*

We introduce a three-step framework to determine, on a per-pitch basis, whether batters in Major League Baseball should swing at a pitch. Unlike traditional plate discipline metrics, which implicitly assume that all batters should always swing (resp. take) pitches inside (resp. outside) the strike zone, our approach explicitly accounts not only for the players and umpires involved but also in-game contextual information like the number of outs, the count, baserunners, and score. Specifically, we first fit flexible Bayesian nonparametric models to estimate (i) the probability that the pitch is called a strike if the batter takes the pitch; (ii) the probability that the batter makes contact if he swings; and (iii) the number of runs the batting team is expected to score following each pitch outcome (e.g. swing and miss, take a called strike, etc.). We then combine these intermediate estimates to determine whether swinging increases the batting team’s run expectancy. Our approach enables natural uncertainty propagation so that we can not only determine the optimal swing/take decision but also quantify our confidence in that decision. We illustrate our framework using a case study of pitches faced by Mike Trout in 2019.

Aiming for Competitive Balance: Developing Fair Handicap Systems for Darts using a Markov Decision Process

Rachael Walker, Craig Fernandes, and Timothy C. Y. Chan, University of Toronto, Toronto, ON

[†] E-mail: *rachael.walker@mail.utoronto.ca*

Handicap systems are commonly used in sports to create competitive balance between players with different skill levels. The game of darts has two existing methods for handicapping. However, their effectiveness has never been rigorously evaluated and they have no guarantee of truly balancing competition. In this paper, we develop a novel Markov Decision Process (MDP) framework that is capable of modeling darts with different handicapping methods and players of varying skill levels. Powering our model with real data from professional dart players, we find that the existing handicap systems – which heuristically calculate a head-start for the weaker player – do not in fact create true competitive balance. We use our model to develop optimization-based versions of these head-start systems that resolve this imbalance. However, we find that even the optimized head-starts can be impractical or even infeasible to implement if the difference in skill level is high enough. This is because the main source of imbalance between players occurs at the end of the game and cannot be addressed by a head-start advantage. In response, we propose a novel handicap system that uses dynamic credits, which can be used at any point in the game and allow a player to deterministically select an outcome for a single dart throw. Moreover, we prove that this more flexible handicap not only creates competitive balance but is guaranteed to be practical and feasible to implement regardless of the difference in skill level.

Estimating Knee Movement Patterns of Recreational Runners Across Training Sessions Using Multilevel Functional Regression Models

Marcos Matabuena, Universidad de Santiago de Compostela; Marta Karas, Harvard University; Sherveen Riaazati, Nick Caplan, and Philip Hayes, Nortumbria University

[†] E-mail: *marcos.matabuena@usc.es*

Modern wearable monitors and laboratory equipment allow the recording of high-frequency data that can be used to quantify human movement. However, currently, data analysis approaches in these domains remain limited. This article proposes a new framework to analyze biomechanical patterns in sport training data recorded across multiple training sessions using multilevel functional models. We apply the methods to subsecond-level data of knee location trajectories collected in 19 recreational runners during a medium-intensity continuous run (MICR) and a high-intensity interval training (HIIT) session, with multiple steps recorded in each participant-session. We estimate functional intra-class correlation coefficient to evaluate the reliability of recorded measurements across multiple sessions of the same training type. Furthermore, we obtained a vectorial representation of the three hierarchical levels of the data and visualize them in a low-dimensional space. Finally, we quantified the differences between genders and between two training types using functional multilevel regression models that incorporate covariate information. We provide an overview of the

relevant methods and make both data and the R code for all analyses freely available online on GitHub. Thus, this work can serve as a helpful reference for practitioners and guide for a broader audience of researchers interested in modeling repeated functional measures at different resolution levels in the context of biomechanics and sports science applications.

Improving Medal Projections in Olympic Best Mark Sports

Suraj Bhuva and Dan Webb, United States Olympic & Paralympic Committee

[†] E-mail: *suraj.bhuva@usopc.org*

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 research develops 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 $m \times n$ 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 $m \times n$ predictions to then determine athlete medal probabilities, with the resulting calibration plots and Brier scores showing an improvement over existing methods.

Poster Presentation Abstracts

Estimating Shot Selection Efficiency in Basketball: A Revealed Efficiency Approach

Justin Ehrlich and Shane Sanders, Syracuse University, NY

† E-mail: *jaehrlie@syr.edu*

We create a shot selection measure called Shot Selection Efficiency that takes as input a player's or team's expected points and expected proportional volume from each observed shooting bin of the half court and computes the player's or team's shot-frequency weighted correlation between expected proportional volume and expected points from the field and free throw line. This represents a measure of a player's or team's shot selection efficiency. We then input Shot Selection Efficiency in an NBA team payroll/performance model to determine if shooting efficiency is fully "priced into" the NBA player market. We find that it is a source of wins not fully explained by player payroll, and therefore it is not fully priced into the player market. Future work will examine whether Shot Selection Efficiency is a measure of team coaching efficiency.

A Bayesian Approach to Modeling Golf Performance Over Time

Ami Gianchandani, Yale University, New Haven, CT; Alex Strasser, Carnegie Mellon University, Pittsburgh, PA

† E-mail: *ami.gianchandani1999@gmail.com*

Analysis of golf statistics is an underdeveloped and therefore underutilized tool for players and coaches looking to improve. Many players do not benefit from a one-size-fits-all approach to practice and need a more customized approach tailored to their game. Additionally, statistics has traditionally been reserved for tour professionals, but here we provide baseline measurements for golfers of all skill levels. This project aims to investigate a Bayesian approach using population-based priors to model a golfer's performance over time. Using strokes gained as a measure of performance, we can effectively diagnose where a player's strengths, weaknesses, and inconsistencies lie as well as see how they compare to their chosen skill-level population. This method provides a novel approach to analyzing golf statistics for individual players without sacrificing their individuality. The ideas presented in this paper can further be used to develop practice plans and drills specifically derived from a player's data.

Assessing Spatial Heterogeneity in Whiff Rate Using Geographically Weighted Regression

Riley Leonard, Cornell University, Ithaca, NY

[†] E-mail: *rileyleonard99@gmail.com*

Spatial statistical analysis offers an exciting opportunity for professional sports organizations to leverage a novel and under-utilized family of statistical approaches for a variety of complex data problems. Spatial statistics has particularly germane applications in the sport of baseball, where play-by-play data is nearly always accompanied by each observation’s vertical and horizontal locations in space—whether it be the coordinates of a pitch as it crosses home plate or the hit location of a batted ball. In these cases, the analysis of baseball data can benefit from the use of spatial statistical techniques.

One such spatial approach is geographically weighted regression (GWR). Whereas the fitted coefficient values of a typical parametric model may fail to capture spatial variations in the data, GWR models the local relationships between predictors and the response variable. In order to demonstrate the capacity of GWR to create accurate and insightful statistical inferences from spatial data, I create a model of whiff probability as it varies in coordinate space. For this analysis, whiffs are defined as events where a batter swings at a pitch and fails to make contact. The non-uniform distribution of whiffs in space suggests an underlying spatial component to the data generating process. As a result, a global parametric model produces confounded coefficient estimates and fails to account for pitch location. GWR produces a more reliable estimation of the relationships between different pitch quality parameters and whiff probability, and evinces pronounced spatial heterogeneity in whiff rate.

Quantifying Defensive Spin Moves in American Football: A Player Tracking Analysis

Samuel Jin, Gunn Sports Analytics Club, Palo Alto, CA

[†] E-mail: *samueljin99@gmail.com*

Defensive pass rushing and offensive pass blocking play pivotal roles in American football. Although previous public research has been done on football linemen in general, there has been limited exploration of individual moves and techniques. In this study, we leverage player tracking data from the 2023 NFL Big Data Bowl competition to develop a framework that identifies spin moves, one prominent type of rushing technique.

Using this identification system, we then train a gradient boosting model to evaluate spin move rushers and analyze the factors that affect the efficacy of a spin move.

The work also creates a spin move rushing metric to quantify the ability of each rusher to execute a spin move. The statistical properties of this metric are demonstrated, showing its correlation with

traditional pass rushing metrics and its stability over time, such as week-to-week correlation over the 2021 NFL season.

The findings of this study offer valuable insights for coaches, players, and fans, providing a comprehensive evaluation of one of the most significant pass rushing moves in football. Furthermore, the identification system can assist coaches in their film study process, allowing them to easily filter through spin moves and supplementing their evaluation of spin moves through quantitative measures.

Competing Risks Analysis of MLB Draft Data

Eric Gerber, Northeastern University, Boston, MA

[†] E-mail: *e.gerber@northeastern.edu*

Baseball is unique in the major US sports in that nearly every player drafted will spend years in the minor leagues (MiLB) before reaching the major league (MLB). Even then, less than half of players drafted will one day make it to MLB. Recently, steps have been taken to attempt to improve the MiLB experience. Since 2021, the MLB draft was reduced from 40 rounds to 20, MLB teams have been required to provide housing for players, and MiLB players have joined the MLBPA union. Yet there will still be hundreds of players drafted each year who will never reach MLB, and most MiLB players still earn poverty-level wages, creating significant risk in pursuing a career in baseball. This research applies competing risks analysis to investigate how different draft day factors influence the time it takes draftees to either reach MLB or retire before doing so.

A Comprehensive Survey of the Home Advantage in American football

Luke Benz, Harvard T.H. Chan School of Public Health, Boston, MA Thompson Bliss and Michael Lopez, National Football League, New York, NY

[†] E-mail: *lukebenz@gmail.com*

The existence and justification to the home advantage, the benefit a sports team receives when playing at home, has been studied across sport. Most research on this topic is restricted to individual leagues in short time frames, which limits a thorough understanding of the drivers of home advantage and prevents extrapolation.

Using nearly two decades of data from the National Football League (NFL), the National Collegiate Athletic Association (NCAA), and high schools from across the United States, we provide a framework for estimating home advantage in American football. We use STAN to fit a suite of Bayesian linear regression based paired-comparison models, with various plausible temporal trends for home advantage: a constant home advantage, a linear home advantage trend over time, and a trend in which home advantage is allowed to vary freely over time. Models are subsequently compared based on expected log pointwise predictive density (ELPD) estimated via the leave-one-out

cross-validation (LOO) approach of Gelman and colleagues. In the name of reproducible research, all code and data for this project is open-source.

To our knowledge, our work is the first to rigorously study home advantage in American football across all levels of the game. Our findings suggest home advantage has been declining in the professional and collegiate football over the past twenty years but remains heterogeneous across states in the amateur ranks of high school football. Taken together, these findings help us generate new hypotheses into the drivers of home advantage for American football.

Blocking is a Drag: Physics-Based Approach to Modeling Offensive Linemen Effectiveness in American Football

Abhijit Brahme, University of California Santa Barbara, Santa Barbara, CA; Ishan Mehta, Houston Texans, Houston, TX

† E-mail: *abhijitbrahme@ucsb.edu*

The evaluation of linemen in American football had been considerably difficult prior to the existence of player tracking data. With the availability of high dimensional spatio-temporal data, we make two unique contributions. First, we adapt a Hidden Markov Model method from basketball (Franks et al. 2015) in order to determine an offensive line’s pass blocking responsibilities, as well as derive novel lineman metrics. Second, we present a generalizable method for modeling pass-rushing plays through the lens of fluid dynamics. We show that when we make use of these contributions in a hierarchical Bayesian linear model, we are able to identify top performing pass rushers and offensive linemen.

Evaluating the outcomes of passing plays from a quarterback’s perspective

Will Morgan, StatsBomb, UK

† E-mail: *will.morgan@statsbomb.com*

When evaluating quarterback decision making and performance, it’s important to consider the trade-offs as a play evolves; when pressured, judging the fine line between throwing the ball, scrambling or being sacked is a key skill.

To quantify this, we have developed a classification model to predict the outcome of a drop-back play, which we use to profile and evaluate quarterback performance across both the NFL and NCAA. The model incorporates game context information (yardline, down and distance) alongside features derived from our tracking data to quantify pressure on the quarterback. Inclusion of separation and route running features for all eligible receivers on a play are also key facets of the model feature set. The model’s target is whether the play ended with a completion, incomple-

interception, scramble or sack. Our initial modelling has utilized an eXtreme Gradient Boosting (XGBoost) approach, which we will compare with a graph neural network model that we are currently developing.

Early findings show that Patrick Mahomes achieves significantly more completions than would be expected when accounting for the probability of other outcomes, while also maintaining much lower sack rates than would be expected. This runs counter to tracking-based models of completion percentage that focus only on attempted passes, where he is typically closer to league-average in terms of Completion Percentage Over Expectation (CPOE). This demonstrates his signature ability to generate positive outcomes, while avoiding strongly negative plays.

Further insights from across the NFL and NCAA will be presented.

On Devon Allen’s Disqualification at the 2022 Track and Field World Championships

Owen Fiore and Jun Yan, University of Connecticut, Storrs, CT

† E-mail: *ofiore75@gmail.com*

Devon Allen was disqualified in the men’s 110 meter hurdle final of the 2022 World Track and Field Championships after registering a reaction time of 0.099 seconds, 0.001 seconds faster than what is allowed. Following the games, bloggers on the running website LetsRun concluded that the reaction time data from the 2022 World Championships seemed to be generally faster compared to the other datasets they considered, but they did not perform any formal statistical analyses. This paper questions the reaction time disqualification barrier, which is currently 0.1 seconds, to determine whether this is a reasonable threshold. We employ a generalized linear mixed model (GLMM) with a random venue effect in order to model the reaction time data. Additionally, we employ a signed-rank test for clustered data to compare reaction times for the same athletes at different competitions. This matter needs to be addressed because disqualification based on allowable reaction time will continue to be an issue in future world championships.

Leveraging PFF Data to Identify Off-the-Ball Runs in Soccer through NFL Route Running

Alexander Schram, Pro Football Focus, Cincinnati, OH; Amelia Probst, Pro Football Focus, Cincinnati, OH; Rory Breslin, Pro Football Focus, Cincinnati, OH

† E-mail: *alexander.schram@pff.com*

This study aims to build a bridge between football and soccer by identifying patterns across both sports. Leveraging PFF data, we analyze off-the-ball runs in soccer through the lens of NFL route running. Comparisons between both sports are not completely unheard of. As recently as June

2023, The Independent described Trent Alexander-Arnold’s performance for England against Malta as “a role that was more quarterback than that associated with the No 10 on his shirt”. An example that demonstrates how similarities between the sports arise when player objectives align.

In football, a route refers to a specific pattern or path executed by a receiver to create separation from defenders and become available for a pass. Similarly, in soccer, off-the-ball runs involve a teammate of the player in possession making movements to exploit open spaces and create passing options. In other words, there is a shared objective of finding open space and establishing passing opportunities in both sports.

We propose a supervised learning procedure that utilizes a combination of tracking data and event data. Our approach aims to identify the nearest neighbors of frames in the opposing sport. We extract key features from each frame, capturing player positions and velocities, creating a distinctive “fingerprint”. Incorporating labeled routes and off-the-ball runs as input data, the model is supervised to recognize relevant patterns and movements. This process aims to uncover shared characteristics in positioning, movement, and strategic decision-making, enhancing our understanding of both sports.

The Causal Effect of Legalized Sports Betting on Fan Aggression

Hua Gong, Rice University, Houston, TX; Wenche Wang, University of Michigan, Ann Arbor, MI

† E-mail: hg37@rice.edu

Since the Supreme Court struck down the Professional and Amateur Sports Protection Act (PASPA) in 2018, a growing number of states across the U.S. have legalized sports betting. Meanwhile, there has been a concern over a potential link between sports betting and fan aggression. For instance, the NBA recently observed increased incidents of fans’ verbal and physical harassment towards players. Fan aggression and dysfunctional behavior can manifest into violence, abuse, and other types of misconducts and crimes.

Understanding the consequences of overturning PASPA has profound implications for policy makers. While sports wagering may bring significant benefits to the local economy, it has potential social impacts. The present study aims to investigate if the legalization of sports betting has led to more aggressive fan behavior and higher crime rates. Understanding the existence and extent of this impact can allow local governments to act proactively to mitigate the negative effects of legalized sports betting if they exist.

We use data from four sources: 1) the dynamic legal landscape of sports betting in the U.S. retrieved from the American Gaming Association, 2) incident-level crime data from the National Incident-Based Reporting System, 3) game-level information from major U.S. sports leagues, and 4) economic information sourced from the Bureau of Economic Analysis. We then employ a difference-in-differences design to causally estimate the effect of legalized sports betting on crime.

Predicting Prospect Projection in Hockey Using Point Shares

Claire Dudley, Northeastern University, Boston, MA and Boston Bruins, Boston, MA; Josh Pohlkamp-Hartt, Boston Bruins, Boston, MA

[†] E-mail: *dudley.c@northeastern.edu*

With players funneling into the NHL from a variety of different leagues, it can be difficult to predict how their stats will translate into the NHL. NHL_e is an example of an equivalency model that converts the value of a point in one league into NHL points. There are many different versions of NHL_e, and our goal is to incorporate a player's defensive value into the estimate of their NHL value. We use Kubatko's point shares method to account for a player's impact on team points, both offensively and defensively. Using Elite Prospects data from 2000-present, we calculate point shares for players in leagues with the a realistic potential to make the NHL in their career, for players aged 17 or older. After standardizing these values by league, we compare the value of a point share between respective leagues and translate point shares to their NHL equivalency. To translate to an NHL equivalent point share, we compare performance year-over-year for players that switch leagues and estimate league-pair translation factors. To handle leagues without a large sample of players with a direct path to the NHL we use the Wilson Method, first translating point share values to intermediary leagues like the AHL with good representation and then translating to the NHL. With this translation method, patterns in the career trajectories of successful players can inform as to how a current prospect may look in the NHL.

Regularized Ordinal Regression to Measure Impacts of Complementary Unit Performance on Scoring in College Football

Andrey Skripnikov, New College of Florida, Sarasota, FL

[†] E-mail: *askripnikov@ncf.edu*

American football is quite unique in the sense that the same team's offensive and defensive units typically consist of separate player sets that don't share the field simultaneously, which tempts one to evaluate them independently. Yet, some aspects of your team's defensive (offensive) performance tend to complement your offense (defense, respectively), e.g. turnovers forced by your defense could lead to easier scoring chances for your offense. We leverage detailed sequential drive-by-drive data from the 2014-2020 college football seasons in order to identify the features of complementary football that impact scoring the most, subsequently adjusting each team's offensive (defensive) performance to project it onto a league-average complementary unit. In particular, we conduct regularized ordinal regression modeling with an elastic penalty that allows variable selection with an added novelty of partial relaxation for the proportional odds assumption. Moreover, given the importance of contextualized team rankings in college football, we incorporate unpenalized components to guarantee the full adjustment of each team's scoring for the strength of their respective schedules and the home-field factor. For residual diagnostics of our ordinal regression models we apply the surrogate approach, creatively extending its use to non-proportional odds models.

Deep Possession: Identifying & Analyzing Ball Touches in Soccer

Robert Mackowiak, Maple Leaf Sports & Entertainment, Toronto, ON, Canada

[†] E-mail: *robert.mackowiak@mlse.com*

In soccer, possessions are often composed of many individual dribbles. However, soccer analytics has traditionally used an approximate approach for possession analyses: interpolating possession between discrete on-ball events. This narrow perspective fails to capture the intricate details and nuances of dribbling that significantly contribute to a player’s and a team’s possession strategies. To address this gap, we present a method to identify and analyze ball touches during possessions using player and ball tracking data.

By calculating horizontal ball acceleration magnitudes from tracked ball positions, we identify touches performed by players during ball possessions based on peaks in the acceleration signal. Using this data, we create spatial maps that provide comprehensive insights into the frequency and location of ball touches. We expand on the identification of ball touches to compare players’ abilities to complete in-game soccer movements with greater or fewer number of ball touches.

This exploration offers a more granular understanding of possession tendencies and better describes how players and teams interact with the ball during gameplay. This novel approach to possession analysis opens avenues for enhanced tactical evaluations, player performance assessments, and strategic decision-making in soccer.

Age-Conditioned Average Treatment Effects: Assessing Load-Management Strategies across Multiple Sports

Shinpei Nakamura-Sakai and Brian Macdonald, Yale University, New Haven CT

[†] E-mail: *s.nakamura.sakai@yale.edu*

Athletes’ performances improve, peak, and eventually decline. This curve is called the “age curve” and we expect this curve to have heterogeneity for the characteristics of the players. In this work, we focus on estimating the effect of rest between games on performance for each age. This helps make decisions about resting a player and so-called “load management”. We make three main contributions: First, we construct a Conditional Expectation Function (CEF) to compare the age curve for different covariates and treatments. Second, using a causal inference approach, we propose a methodology to construct an age-conditioned treatment effect (ACTE) for a given treatment. The ACTE can test causal hypotheses for each age on the treatment and outcomes of interest. Third, we apply this method to assess the effect of days between games on multiple performance metrics conditional on age.

Effect of Downhill Races on Boston Marathon Qualifying Time

Kellis Ward, Laura Albrecht, and Dorit Hammerling, Colorado School of Mines, Golden, CO;
Richard Smith, University of North Carolina, Chapel Hill, NC

[†] E-mail: *kellisward@mines.edu*

The Boston Marathon is the world’s oldest, and one of its largest, marathon races. Primary means of entry is achieving a qualifying time based on age and sex. Qualifying times may be achieved on any certified marathon course, however some courses have a net drop (calculated as the difference between the start and finish elevation) of as much as 5,000 feet. We focus our study on races with at least 3000 feet of net drop and evaluate if such races provide an advantage for qualification purposes. We stratify our analysis by splitting runners based on sex and age to analyze the effects of downhill courses separately for these groups. We find that net drop of at least 3000 feet leads to a higher proportion of qualifying times than non-downhill races. We suggest adjustments to qualifying times based on the net drop, i.e., making qualifying times for women 3-5 minutes faster and times for men 1-3 minutes faster if a downhill race is used for qualifying. We compare the performance of runners in the Boston Marathon, grouping them by those qualifying through a downhill versus a non-downhill race. Our results indicate that in the aggregate runners who qualify using a downhill race have a larger difference between their Boston Marathon qualifying and race time than runners who qualify on a flatter course. We thank and acknowledge the Boston Athletic Association, organizers of the Boston Marathon, for their input and suggestions related to this project.

Using the discrepancy between judges’ scores and athlete ability to determine order and difficulty bias

Monnie McGee, Southern Methodist University, Dallas, TX and Gabe Downey University of Kansas, Lawrence, KS

[†] E-mail: *mmcgee@smu.edu*

Sports such as diving, gymnastics, and ice skating rely on expert judges to score athletic performance during competition. To assure that scores are fair and comparable, judges typically follow a set of standards for various components of a routine and align their scores to these standards. However, eliminating subjectively is impossible even for the most conscientious and experienced judges. While nationalistic bias, particularly in ice skating and gymnastics, has received the most media attention, other forms of bias exist. These include order bias, where the order of the dives performed affects the scores, and difficulty bias, where athletes performing more difficult elements obtain higher scores than deserved. Measuring these types of biases requires measure of athlete ability, or what score an athlete “deserves”. In a previous study, we compared judges’ scores to measures of common diving elements obtained from video replay, which is not always available. Using records from *divemeets.com*, a data base of most diving competitions, we examined records 674 different meets by 5250 unique divers between 2017 and 2022. Each record contains information on the dive

performed, the degree of difficulty, the order of the dive, the age of the diver, and the judges' scores. Using the median score for each diver across all meets and dives as a measure of diver ability, we compare this measure to the dive performed, the difficulty of the dive, and the order of the dives to quantify difficulty and sequential biases and discuss scoring adjustments.

Attacking Tendencies by Setter Rotation in NCAA Women's Volleyball

Kevin Floyd, iFIT Inc., Logan, UT; Emily Wright, Volleyball Canada

† E-mail: *kfloyd14@hotmail.com*

In indoor volleyball, players rotate throughout all six positions on the court, three front court and three back court. Team offensive strategy may change depending on if the setter is in a front court or back court position. Using the nascent `ncaawvbR` package, we gather collegiate women's indoor volleyball play-by-play data from the 2021 and 2022 seasons from the NCAA database. For teams that rotate their setters without substituting, we analyze the frequency of sets to front and back court attackers based on setter position. We believe this, and other such questions answerable by rotation data gathered by `ncaawvbR`, will be useful for volleyball coaches and enthusiasts alike.