★ March 10-12, 2023 ★ Phoenix, AZ ★

SABR

ANALYTICS CONFERENCE

Rawlings Est. 1887

Major League Baseball
Is Proud To Support
SABR

Welcome!

Welcome to the 12th annual SABR Analytics Conference, presented by Major League Baseball and Rawlings. This event will be held over three days from March 10-12, 2023, at the Beus Center for Law and Society, 111 E. Taylor St., on Arizona State University's Downtown Phoenix Campus in Phoenix, Arizona.

In 2023, we have invited another top group of speakers from throughout the baseball industry, including Dallas Braden, former major-league pitcher and TV analyst; Bobby Scales, former major-leaguer and Sports Info Solutions Vice President; Ari Kaplan, MLB front office executive; Colin Cahill of the San Francisco Giants; Roger Caplinger of the Milwaukee Brewers; Ethan Stewart of the Oakland A's; Jake Schuster of Gemini Sports Analytics; Nicholas Kapur of Zelus Analytics; Mike Petriello, Jason Bernard, Dana Bennett, Clay Nunnally, and Graham Goldbeck of Major League Baseball; Brian Hall of Tidal Quantitative Research; David Loving of Sage; Kyle Glaser of Baseball America; Meg Rowley and Ben Clemens of FanGraphs; Hannah Keyser of Yahoo! Sports; Rob Mains of Baseball Prospectus; Webster University professor Allison R. Levin; data scientist Stuart Neilson; and more. We'll also have many research presentations with a wide-ranging, diverse group of speakers throughout the conference.

Each year, the top minds of the baseball analytic community gather to discuss, debate and share insightful ways to analyze and examine the great game of baseball. The event is a natural for SABR. The Society for American Baseball Research has a long and storied history with baseball statistical analysis, evidenced by the link between our name and sabermetrics.

This year's conference will include a combination of guest speakers, panels, and cutting-edge research presentations — plus the unique Diamond Dollars Case Competition, in which high school and college students from across the country analyze and present a real baseball operations decision. Students will have opportunities for skill acquisition and development, including sessions on how to work with baseball-related data sets as well as databasing and SQL. Student registration includes a one-year SABR membership and a complimentary Rapsodo certification course.

General session programming will run on Friday from 1:00-4:20 p.m., Saturday from 9:00 a.m.-5:15 p.m., and Sunday from 9:00 a.m.-1:30 p.m.

Once again, we want to thank all of our attendees and our sponsors, which include Major League Baseball, Rawlings, Baseball-Reference.com, Sports Info Solutions, SMT, Rapsodo, Cleat Street Capital, Kinatrax, BaseballCloud, the University of the Pacific's Eberhardt School of Business, Arizona State University's Allan "Bud" Selig Sports Law and Business Program, and the following MLB teams: the Arizona Diamondbacks, Chicago Cubs, Cincinnati Reds, Cleveland Guardians, and Colorado Rockies.

You can visit SABR.org/analytics during and after the conference for complete coverage of the 2023 SABR Analytics Conference.

Scott Bush, SABR Chief Executive Officer

Events Schedule

Zoom webinar details will be sent by email to all student attendees.
All general sessions will be held in the Beus Center Great Hall, except where indicated.
Times listed are local times in Arizona (Mountain Standard Time).
** denotes student track presentation*

Friday, March 10

♦ **8:30-8:50 a.m.: Introduction to Diamond Dollars Case Competition**
Vince Gennaro will introduce the 2023 Diamond Dollars Case Competition in the Beus Center Great Hall, with presentations to follow.

♦ **9:00 a.m.-12:00 p.m.: Diamond Dollars Case Competition**
Presentations of the Diamond Dollars Case Competition in Beus Center rooms 240, 250, 442, 450.

♦ **9:30 a.m.-4:30 p.m.: Registration Open (Beus Center Plaza)**
You can pick up your SABR Analytics Conference registration packet in the Beus Center Plaza.

♦ **12:00-12:45 p.m.: Lunch (Beus Center Plaza)**
Analytics Conference registration includes lunch, which will be served in the Beus Center Plaza.

Session 1: Presented by Kinatrax

♦ **1:00-1:05 p.m.: Conference Welcome**
SABR CEO Scott Bush welcomes attendees in the Beus Center Great Hall.

♦ **1:10-2:10 p.m.: Integration of Biomechanics in Baseball Operations**
Panelists: Colin Cahill, San Francisco Giants; Roger Caplinger, Milwaukee Brewers. Moderator: Ethan Stewart, Oakland Athletics.

♦ **2:15-2:45 p.m.: RP1: Comparing Age Curves Across the MLB, KBO, and NPB**
John Asel and Jeremy Losak

♦ **2:45-3:15 p.m.: RP2: Effect of Previous Innings Played on Plate Appearances***
Alex Oppel

♦ **3:20-4:20 p.m.: 2023 Rule Changes Panel**
Panelists: Dallas Braden, NBC Bay Area; Kyle Glaser, Baseball America; Bobby Scales, Sports Info Solutions. Moderator: Rob Mains, Baseball Prospectus.

♦ **4:30-7:00 p.m.: Networking Reception with Baseball Industry Network, presented by Rapsodo (The Pub at Thunderbird)**
The networking reception is an opportunity for conference attendees to meet some of our panelists, speakers, writers and others working throughout the baseball industry. Light hors d'oeuvres served with a cash bar. Located next door in The Pub at Thunderbird, 401 N. 1st Street, 5th floor.

Events Schedule

Saturday, March 11

♦ **8:15-8:45 a.m.: Networking Coffee Talk, presented by Cleat Street Capital (Beus Center Mezzanine, 2nd floor)**
Grab a cup of coffee while we discuss career paths and network. Networking coffee talks are for industry executives and students looking to get into the game and will include industry job postings. Located in the Beus Center Mezzanine on the 2nd floor.

Session 2: Presented by SMT and the University of the Pacific's Eberhardt School of Business

♦ **9:00-9:30 a.m.: RP3: Predicting Hitting Power using Energy Flow Analysis in High School, Collegiate, and Professional Baseball Players**
Arnel Aguinaldo

♦ **9:30-10:00 a.m.: RP4: The Next Frontier of Biomechanics in Baseball: Deception, Pitch Tipping, and Mechanical Evaluation**
Alex Caravan and Kyle Wasserberger

♦ **10:00-10:30 a.m.: RP5: Impact of Stride Length on Pelvis Energy Flow: A Retrospective Analysis***
Taylor La Salle

♦ **10:30-11:00 a.m.: Diamond Dollars Case Competition**
Encore presentation from the winning team of the Diamond Dollars Case Competition.

♦ **11:15 a.m.-12:15 p.m.: Players We Love for 2023**
Panelists: Ben Clemens, FanGraphs; Hannah Keyser, Yahoo! Sports. Moderator: Vince Gennaro, former SABR Board President.

♦ **12:15-1:00 p.m.: Lunch (Beus Center Plaza)**
Analytics Conference registration includes lunch, which will be served in the Beus Center Plaza.

Session 3: Presented by Baseball-Reference.com

♦ **1:15-1:30 p.m.: SABR Analytics Conference Research Awards presentation**
SABR CEO Scott Bush will recognize winners of the 2023 SABR Analytics Conference Research Awards.

♦ **1:30-2:30 p.m.: Projection Systems Panel**
Panelists: Nicholas Kapur, Zelus Analytics; Jake Schuster, Gemini Sports Analytics; Meg Rowley, FanGraphs. Moderator: Ari Kaplan, MLB front office executive.

♦ **2:30-3:00 p.m.: RP6: Evaluating Minor League Defense Using Detailed Contextual Data**
Sarah Thompson

♦ **3:00-3:30 p.m.: RP7: Digging Deeper into the Seam Shifted Wake Effect**
John Garrett

♦ **3:30-4:00 p.m.: RP8: Pitchers Hitting: A Requiem**
Rob Mains

Events Schedule

♦ **4:15-5:15 p.m.: The Future Impacts of Artificial Intelligence and Machine Learning on Baseball**
Panelists: Brian Hall, Tidal Quantitative Research; Allison R. Levin, Webster University; David Loving, Sage; Stuart Neilson, data scientist.

Sunday, March 12

♦ **8:15-8:45 a.m.: Networking Coffee Talk, presented by Rapsodo**
(Beus Center Mezzanine, 2nd floor)
Grab a cup of coffee while we discuss career paths and network. Networking coffee talks are for industry executives and students looking to get into the game and will include industry job postings. Located in the Beus Center Mezzanine on the 2nd floor.

Session 4: Presented by BaseballCloud

♦ **9:00-9:30 a.m.: New Statcast Data: Weather and High-Frame-Rate Bat and Biomechanics Tracking**
Panelists: Graham Goldbeck and Clay Nunnally, Major League Baseball.

♦ **9:30-10:00 a.m.: RP9: How to Effectively Communicate Baseball Research**
Rebekah Callari-Kaczmarczyk

♦ **10:00-10:30 a.m.: RP10: Using Seam Orientation Data from Optical Sensors for Pitch Modeling and Design**
Glenn Healey

♦ **10:30-11:00 a.m.: RP11: Quantifying In-Game Command: Observed Data Meets Intended Location**
Ryan Reinsel

♦ **11:00-11:15 a.m.: SABR Analytics Conference Lifetime Achievement Award presentation**

♦ **11:15-11:45 a.m.: Baseball Savant's Newest Developments**
Panelists: Mike Petriello, Jason Bernard, Dana Bennett, Major League Baseball.

♦ **11:45 a.m.-12:15 p.m.: RP12: Trends in MLB Pitch Tempo Since the Start of Pitch Tracking**
Andy Andres

♦ **12:15-12:45 p.m.: RP13: Noisy Judgments: A Probability Surface-Based Analysis of Umpiring Variability**
Emily-Anne Patt and James Stockton

♦ **12:45-1:15 p.m.: RP14: A Markov Chain Matchup Model Using Multi-variable Approach***
Jeff Jin and Chris Zexin Chen

♦ **1:15-1:30 p.m.: Conference Wrap-Up**
SABR Chief Executive Officer Scott Bush.

Virtual Student Sessions

In addition to the complete general session program, the following virtual sessions are available on Thursday, March 9 to anyone who registers as a student for the 2023 SABR Analytics Conference. A special Zoom link, separate from the general sessions, will be sent to all students. Times listed are local times in Arizona (Mountain Standard Time).

Thursday, March 9

♦ **11:00 a.m.-12:00 p.m.: Introduction to Baseball Databases**
Sean Lahman, SABR

♦ **12:15-1:15 p.m.: Working with Baseball-Reference Data**
Jonah Gardner and Zoe Surma, Sports Reference

♦ **1:30-2:00 p.m.: Working with Baseball Savant and R**
Shanna Shi, Associate Data Quality Analyst, Major League Baseball

♦ **2:00-3:00 p.m.: Confidence Intervals/Statistical Difference Tests**
David Hyland, Xavier University

♦ **3:00-4:00 p.m.: SLICED with SABR**
Nick Wan, Director of Analytics, Cincinnati Reds

♦ **5:00-6:00 p.m.: Student Networking Breakout Room**
We invite you to an informal breakout session with your peers. This is a great way to build connections for your network and discuss the conference with other attendees.

THE OFFICIAL BASEBALL GLOVE OF
MAJOR LEAGUE BASEBALL

AUTHENTIC COLLECTION

Speakers/Panelists

Conference Welcome
1:00-1:05 p.m. MST, Friday, March 10

♦ **Scott Bush** is SABR's Chief Executive Officer. During his tenure, SABR membership has reached record levels while the organization has expanded research domains and added analytics certification programming to its offerings. He joined the organization in 2018 after serving as the Senior Vice President for Business Development with the Goldklang Group. Since graduating from the University of Minnesota, Bush held positions with increasing responsibilities in both sports and media, including a five-year stint as Assistant General Manager for the St. Paul Saints, where he played a key role in establishing CHS Field in St. Paul, Minnesota.

Integration of Biomechanics in Baseball Operations
1:10-2:10 p.m. MST, Friday, March 10

♦ **Colin Cahill** is in his fourth season as the San Francisco Giants' Vice President of Player Performance and Wellness. In his role, he works across Baseball Operations to help drive the organization's efforts in medicine, sports science, and technology. Previously, he spent 20 years working at the intersection of healthcare and technology in the Bay Area and in Boston. He has a master's degree in Biological Sciences and an MBA from Stanford University.

♦ **Roger Caplinger** recently retired after spending 33 seasons with the Milwaukee Brewers. As Vice President, Medical Operations, Health & Safety for the club, he oversaw all medical operations, player health and psychological services, and health and safety initiatives. He was the Brewers' head athletic trainer from 2002 to 2011, and the training room in the Brewers' home clubhouse is named after him. Since 2009, he has served as the chairman for the Major League Baseball Electronic Medical Records Committee. He is a licensed athletic trainer in the states of Wisconsin and Arizona. He graduated from the Metropolitan State University of Denver in 1991.

♦ **Moderator: Ethan Stewart** is the Lead Sport Scientist and Biomechanist with the Oakland A's. Previously, he spent time with the Cincinnati Reds, Baltimore Orioles, and Tampa Bay Rays. He is a founding member and board member for the American Baseball Biomechanics Society. He completed both his B.S. and M.S. degrees from the University of Kentucky in Kinesiology and Biomechanics, and is currently a Ph.D. candidate at Mississippi State University studying Biomechanics/Neuromechanics.

Speakers/Panelists

2023 Rule Changes Panel

3:20-4:20 p.m. MST, Friday, March 10

♦ **Dallas Braden** is a former major-league pitcher and television analyst from Stockton, California. He spent five seasons with the Oakland A's from 2007 to 2011, and pitched the 19th perfect game in MLB history on May 9, 2010, on Mother's Day against the Tampa Bay Rays. He has worked for ESPN and NBC Sports California as a broadcaster since 2014. He co-hosts the "Baseball is Dead" podcast and a SiriusXM radio show, "Dialed in with Dallas Braden."

♦ **Kyle Glaser** covers Major League Baseball and top prospects for *Baseball America*, where he also assists with West Coast draft coverage. He previously covered the San Diego Padres, the California League and amateur baseball for the *Riverside Press-Enterprise* and *Orange County Register*, and freelanced for the Associated Press. He is a graduate of Arizona State University.

♦ **Bobby Scales** currently serves as a Vice President at Sports Info Solutions, where he is responsible for advancing the future product, research and strategic direction in the baseball division. After he was drafted out of the University of Michigan, he played professional baseball for 14 years and spent parts of two seasons with the Chicago Cubs in 2009 and 2019. He also played in the Padres, Red Sox, and Phillies organizations, and in Japan for two seasons. He worked in the player development departments for the Los Angeles Angels and Pittsburgh Pirates before joining Sports Info Solutions in 2022.

♦ **Moderator: Rob Mains** is a writer for Baseball Prospectus. His "Veteran Presence" column focuses on analytics, history, and finance topics. He also manages *BP en español*, the site's Spanish-language content. He is a former Wall Street equities analyst.

Players We Love for 2023

11:15 a.m.-12:15 p.m. MST, Saturday, March 11

♦ **Ben Clemens** is a writer at FanGraphs. His work has also appeared at Viva El Birdos. He lives in San Francisco with his wife and their wonderful but nervous dog, Ruby.

♦ **Hannah Keyser** is a writer and reporter with Yahoo! Sports. She has covered sports and food for Major League Baseball, Deadspin, and Vice. She lives in New York, after growing up outside of Philadelphia and getting an ancient history degree from the University of Pennsylvania that almost never comes up in her current line of work.

♦ **Moderator: Vince Gennaro** is the former President of SABR's Board of Directors, author of *Diamond Dollars: The Economics of Winning in Baseball*, and host of a weekly national radio show, *Behind the Numbers: Baseball SABR Style on SiriusXM*. He is also the Associate Dean and Clinical Associate Professor of NYU's Preston Robert Tisch Institute for Global Sport. He is a consultant to MLB teams and appears regularly on MLB Network. He is also the architect of the Diamond Dollars Case Competition series, which brings together students and MLB team and league executives and serves as a unique learning experience, as well as a networking opportunity for aspiring sports executives.

Speakers/Panelists

Projection Systems Panel
1:30-2:30 p.m. MST, Saturday, March 11

♦ **Nicholas Kapur** is a Senior Data Scientist at Zelus Analytics. Previously, he worked in the Los Angeles Dodgers and New York Yankees front offices as a quantitative analyst for five seasons. He holds a bachelor's degree in Statistics and Applied Mathematics and a master's degree in Statistics from North Carolina State University.

♦ **Jake Schuster** is the founder and CEO at Gemini Sports Analytics, which provides sports-specific, end-to-end Automated Machine Learning tools to sports organizations around the world. Previously, he was a Senior Sports Scientist at Vald Performance. He has also worked in strength and conditioning with Florida State University, USA Field Hockey, and other sports programs. He holds a Ph.D. in sport science from the Auckland University of Technology, a master's degree in sport and exercise nutrition from Loughborough University, and a bachelor's degree in kinesiology and exercise science from McDaniel College.

♦ **Meg Rowley** is the managing editor of FanGraphs and co-host of the Effectively Wild podcast. Previously, she served as managing editor of The Hardball Times and wrote for Baseball Prospectus, where she won the 2017 SABR Analytics Conference Research Award for Contemporary Baseball Commentary. She holds a bachelor's degree in political science from Bryn Mawr College and a master's degree in political science from the University of Wisconsin.

♦ **Moderator: Ari Kaplan** has been a major-league front office executive for more than three decades. He created and led the Chicago Cubs and Los Angeles Dodgers analytics departments, and served as a special consultant to the GM with the Baltimore Orioles. He has also worked with many Fortune 500 companies as a leader in the IT industry, including Oracle, 3Com, Hallmark, and Merck, among others. He received Caltech's "Alumni of the Decade" distinction for his groundbreaking sabermetric work and he has co-authored five best-selling books on analytics, databases, and baseball.

The Future Impacts of Artificial Intelligence and Machine Learning on Baseball
4:15-5:15 p.m. MST, Saturday, March 11

♦ **Allison R. Levin** (MA, JD) is a consultant and professor at Webster University in St. Louis. Her research builds upon an eclectic background in political science, economics, women's studies, communications, and law. In recent years her focus has been on the effects of social media on pop culture, communications, and sports, primarily baseball. Starting with that framework, she has looked at Artificial Intelligence and Machine Learning from a communication and experiential aspect, as seen in her article in the 2022 *National Pastime* entitled "Baseball in 2040: The Digital Viewer." In addition, she built an interactive database of baseball players' walkup music which she presented at the 2019 SABR Convention. In addition to serving as co-chair of the AI/ML research committee, Allison is the President of the Bob Broeg St. Louis SABR Chapter and on the SABR Board of Directors.

♦ **Stuart Neilson** is a data scientist working on risk models in the banking industry. Previously (before data scientists were called data scientists), he built similar models in the economic consulting field. He holds a bachelor's degree from Victoria University in New Zealand and master's degrees from Tufts University and Harvard University.

Speakers/Panelists

♦ **David Loving** is a Machine Learning Engineer at Sage, where he builds AI tools to power the future of finance and accounting for SMBs. Previously, he led machine learning development at Camugo to deliver a fully-automated sports event broadcasting camera system. He is finishing up his graduate studies at Harvard University by working with Harvard's StellarDNN group to apply AI to black hole images from the Event Horizon Telescope.

♦ **Moderator: Brian Hall** is the founder of Tidal Quantitative Research, an applied research company dedicated to AI and machine learning. He also teaches undergraduate courses on analytics, artificial intelligence, and machine learning at NYU. Prior to joining the NYU faculty, Brian served as Vice President at Two Sigma Investments, a New York City-based hedge fund primarily known for its use of artificial intelligence. At Harvard University, Brian's graduate studies combined his passion for sports with advances in computational programming, machine learning, and neural networks.

New Statcast Data: Weather and High-Frame-Rate Bat and Biomechanics Tracking

9:00-9:30 a.m. MST, Sunday, March 12

♦ **Graham Goldbeck** is a Product Manager, Baseball Data Platform, for Major League Baseball. Before joining MLB in 2018, he spent seven years managing data and analytics for Sportvision. He holds a bachelor's degree in mathematics and economics from the University of California at San Diego.

♦ **Clay Nunnally** is a Baseball Scientist at Major League Baseball, where he has worked since 2018. Previously, he served as the Chief Operating Officer at Applied Physical Electronics. He holds a bachelor's degree in physics from Abilene Christian University, a master's degree in engineering physics from the University of Virginia, and a Ph.D. in electrical engineering from the University of Missouri.

Baseball Savant's Newest Developments

11:15-11:45 a.m. MST, Sunday, March 12

♦ **Mike Petriello** is a Stats Analyst for Major League Baseball and host of the Statcast podcast. He partnered with Jason Benetti and Eduardo Pérez on the popular "Nerdcast" TV broadcasts for ESPN. He can regularly be seen on MLB Network and across all MLB platforms. Prior to joining MLB, he wrote for FanGraphs and ESPN.

♦ **Dana Bennett** is a Data Visualization Engineer at Major League Baseball. He works on all things related to the Baseball Savant website. He holds a bachelor's degree in mathematics from Bates College. His previous work was with the Atlanta Braves and Chess.com.

♦ **Jason Bernard** is a Manager, Baseball Research & Development at Major League Baseball, where he has worked since 2015. He has also served as an adjunct professor at the NYU-Tisch Institute for Global Sport. He earned a bachelor's degree in sports management from St. John's University and a master's degree in sports business from New York University.

STATHEAD BASEBALL

POWERED BY BASEBALL REFERENCE

IS BASEBALL REFERENCE YOUR HAPPY PLACE? THEN STATHEAD IS FOR YOU.

WHAT IS STATHEAD?

It's a sports seach engine with quick and easy tools to go deeper into stats and feed your curiosity

SAVE $20 WITH COUPON CODE

SABR23

CURRENT STUDENTS RECEIVE A 50% DISCOUNT. NO CODE REQUIRED.

VALID ON SINGLE SPORT OR ALL SPORT ANNUAL SUBSCRIPTIONS

This coupon code is valid on new subscriptions only and cannot be combined with existing offers.
This offer expires June 1, 2023

SABR

Speakers/Panelists

Virtual Student Sessions

11:00 a.m.-4:00 p.m. MST, Thursday, March 9

- **Sean Lahman** is SABR's Data Projects Manager and the creator of the Lahman Baseball Database, an open-source collection of baseball statistics. He has edited or contributed to more than a dozen sports reference books, including *Total Baseball* and the *ESPN Baseball Encyclopedia*. He is an award-winning investigative journalist for the *Rochester Democrat and Chronicle* in New York.
- **Jonah Gardner** is the Product Marketing Manager at Sports Reference, where he's responsible for coordinating marketing campaigns for Stathead, Sports Reference's suite of advanced research tools. Before that, he worked as Social Media Manager for Sports Reference, providing live coverage and up to the minute statistical research for the biggest games of the season from Baseball Reference's social media accounts. Jonah lives in Philadelphia with his dog, Berman, and roots for the World Champion Atlanta Braves.
- **Zoe Surma** is a Software Engineer at Sports Reference. She is a graduate of the University of Notre Dame with a degree in Computer Engineering. She is a Cubs, Fighting Irish and Formula One fan.
- **Shanna Shi** is an Associate Data Quality Analyst at Major League Baseball, where she has worked since 2021. She previously worked with the Toronto Blue Jays as a Baseball Operations Fellow and as a research assistant at McMaster University. She holds a master's degree in Public Health, Epidemiology, from Emory University and a bachelor's degree in Psychology from Union College.
- **David Hyland** is a professor at Xavier University's Williams College of Business in Cincinnati, Ohio, where he has worked since 2004. He also developed a course on sabermetrics, "Teaching Statistics with Baseball Data," at Xavier. He serves on the Board of Directors for the Florence Y'alls club in the independent Frontier League. His research has been published in journals such as *Financial Management* and *Financial Analyst Journal*. He earned his Ph.D. from Ohio State University, a master's degree from Purdue University, and a bachelor's degree from Indiana University.
- **Nick Wan** is the Director of Analytics for the Cincinnati Reds, where he has worked since 2017. In this role, he oversees the day-to-day analytics operations for the Major League team, Player Development, Health & Performance, Professional Scouting, and Amateur Scouting. He previously spent three years as a Data Scientist with the Reds. He earned his Ph.D in Psychology from Utah State University and earned his B.S. in Psychology from Saint Mary's College of California. He is also the game show host for SLICED, a competitive data science competition which is streamed on his Twitch channel.

The data you need to win

SIS provides cutting edge data to Teams, Sportsbooks, Media and Fans to help them win. We go deep and provide:

- Coverage of MLB, MiLB, KBO, NPB and Amateur
- Robust Defensive Metrics
- Stolen Base and Base Running Analytics
- Play-by-Play and Pitch Charting Alongside a Pitch Predictor Tool
- Synthetic Statcast for Leagues Below MLB
- Hit Locations, Trajectories and Timers
- Advanced Injury Information and Predictive Analysis
- Umpire Profiles
- 3rd Base Coach Profiles
- Universal ID System, Tying Athletes Across Levels of Baseball

sportsinfosolutions.com

SIS Baseball Podcast

@sportsinfo_SIS

linkedin.com/company/sports-info-solutions

Research Presentations

SABR and Sports Info Solutions are pleased to announce the research presentations that will be delivered at the 2023 SABR Analytics Conference. Times listed are local times in Arizona (Mountain Standard Time).

Friday, March 10

2:15-2:45 p.m. MST
RP1: Comparing Age Curves Across the MLB, KBO, and NPB
John Asel and Jeremy Losak

Identifying how players age is critical in baseball, especially when teams look to forecast player talent into the future. While aging patterns among Major League Baseball (MLB) players are well-documented, to this point there is no research examining aging patterns for players in different leagues. In this study, we compute aging curves for batters in the KBO (Korean Baseball Organization), NPB (Nippon Professional Baseball), and MLB. We then compare these aging curves to identify if players in different leagues exhibit different aging behaviors.

There are reasons to believe aging patterns may be different. First, different skill sets are prioritized in different leagues. For example, Leeds et al. (2012) showed that NPB rewards small ball performance factors more than MLB teams. Second, there may exist genetic-driven physical differences between players in each league (for example, the average adult male Japanese man is two inches shorter than the average adult male American man). Third, game load differences may lead to quicker physical decline. For example, MLB seasons consist of 162 regular season games, compared to 143 games in NPB and 144 in KBO.

Our data come from Lahman (MLB) and Baseball-Reference (KBO and NPB). We focus our analysis from 1982 to 2022, as 1982 is the first season in which we have data for all three leagues. In total, we have 27,682 player-season pairs (16,694 in MLB, 6,487 in NPB, and 4,501 in KBO), covering 5,157 unique players (2,918 MLB, 1,356 NPB, 889 in KBO), which we use to estimate aging patterns with generalized additive models. Results show mostly similar aging patterns between MLB and NPB. OPS peaks around 26.7 years old, K% peaks around 29.3 years old, BB% peaks during a player's early 30s, and HR% peaks between ages 27 and 29 with the peak being more pronounced for MLB hitters. MLB players also show slightly longer longevity. KBO players, on the other hand, exhibit a different aging pattern. Compared to MLB and NPB players, KBO players enter the league closer to their career averages, and then decline more quickly following their peak (for OPS, BB%, and HR%). KBO players, on average, are also younger relative to the other leagues. This may highlight cross-cultural differences in training or development regimens that lead to different aging patterns.

John Asel is a Syracuse University Sport Analytics and Economics student. He has interned for Driveline and the Tampa Bay Rays, and won the 2022 Doug Pappas Award at SABR 50. He will be joining the Baltimore Orioles analytics department upon graduation.

Jeremy Losak is an Assistant Professor of Sport Analytics in the Department of Sport Management at Syracuse University's David B. Falk College of Sport and Human Dynamics. His research focuses on the economics of sports, particularly baseball labor markets, attendance at sporting events, gambling markets, and college athletics. Previous SABR Analytics Conference research presentations include "Behavioral Biases in Daily Fantasy Baseball: The Case of the Hot Hand" (2022), "MLB Home Field Advantage Without Fans" (2021), and "What's Hanging? An Empirical Definition And Defining Attributes For The Hanging Pitch" (2020).

Research Presentations

2:45-3:15 p.m. MST
RP2: Effect of Previous Innings Played on Plate Appearances
Alex Oppel

While the effect of rest and workload on pitcher performance is well acknowledged, the evaluation of rest, or lack thereof, on hitter performance is limited. Consequently, this study quantifies the effect of position player workload on various plate appearance outcomes. Baseball Reference game log data from the 2010s are used to calculate each batter's innings played totals over the previous seven days. Binary logistic regressions using Retrosheet Event Files are built to model the probability of a hit, walk, strikeout, and home run occurring. Innings played totals over the past week are included in all four models to determine the effect of workload on batter performance.

Results indicate the effects of innings played over the previous seven days on walks, strikeouts, and home runs is insignificant; however, weak evidence exists to suggest a quadratic relationship for the effect of batter workload on hit probability. Holding all other variables constant, hit probability is maximized at 42 innings played, potentially suggesting batters could use an occasional day off to stay fresh.

Originally from Muscatine, Iowa, **Alex Oppel** is majoring in Sport Analytics at Syracuse University, where he serves as president of the Baseball Statistics and Sabermetrics Club, assists the softball team with game strategy and analytics, and works as a research assistant for Sport Analytics faculty. He previously completed an internship as a data analyst for the Northern Colorado Owlz and aims to pursue a position in baseball analytics after graduation.

Saturday, March 11

9:00-9:30 a.m. MST
RP3: Predicting Hitting Power using Energy Flow Analysis in High School, Collegiate, and Professional Baseball Players
Arnel Aguinaldo

Hitting performance is largely dependent on the batter's ability to produce a powerful swing with a kinetic chain through which mechanical energy flows between rotating body segments. Evidence in the current literature suggests that the energy generated, absorbed, and transferred by the net torques at the lead hip, back hip, and lumbosacral (L5S1) joints contribute to increased pitch velocity in baseball pitchers. However, it is unclear if energy flow patterns are similar during baseball hitting, particularly across different levels of baseball. Therefore, this analysis aimed to examine the relationship of energy generation, absorption, and transfer through the pelvis and bat velocity and to determine how they differ between high school, collegiate, and professional baseball hitters.

Kinematic and kinetic analyses were performed on high school (n=13), collegiate (n=13), and professional (n=11) baseball players using 3D motion capture and force platform technologies. Segment and joint torque powers were estimated using a conventional energy flow analysis that indicate the rates of energy generation, absorption, and transfer by joint torques. We employed a regularized regression analysis to determine the energy flow and ground reaction force predictors of bat speed, which was measured using an IMU-based bat swing sensor. High school, collegiate, and professional players exhibited bat speeds of 27.1 ± 2.7 m/s, 30.7 ± 1.9 m/s, and 32.3 ± 3.3 m/s, respectively, the difference between which was statistically significant ($p < .001$). While the rates of energy flow were similar across all levels of hitters, more

Research Presentations

experienced players exhibited higher levels in energy transfer at the lead hip during acceleration as well as in energy absorbed by the L5S1 joint during follow-through. Our model predicted bat speed with an RMSE of 1.1 m/s with roughly 60% of its variance explained by level, back leg impulse, back hip transfer, L5S1 generation and transfer during bat acceleration, and lead hip absorption and transfer during follow-through.

These findings reinforce the crucial role of rotations of the back hip and trunk during bat acceleration as well as the deceleration induced by the torques at the lead hip and L5S1 joints during follow-through. Hence, training strategies that enhance the muscular strength and mobility of the trunk and hips could potentially improve bat speed and overall power of the swing.

Arnel Aguinaldo is an Associate Professor of Kinesiology and Director of the Biomechanics and Pitching laboratories at Point Loma Nazarene University. He has earned an undergraduate degree in Bioengineering and a Ph.D. in Health and Human Performance and is a board-certified athletic trainer. Dr. Aguinaldo currently serves on the board of directors of the American Baseball Biomechanics Society and as a scientific advisor to Major League Baseball. His research aims to understand the biomechanical implications of injury risk and player performance in baseball pitching and hitting.

9:30-10:00 a.m. MST

RP4: The Next Frontier of Biomechanics in Baseball: Deception, Pitch Tipping, and Mechanical Evaluation

Alex Caravan and Kyle Wasserberger

Professional baseball is inundated with more human movement data than ever before. Of particular interest are the potential use cases to 1) help explain why some players over or underperform projections and 2) track player development over time. As early adopters of motion capture data in baseball, Driveline has gone through several iterations of biomechanics-driven player evaluation and development decision systems. Our presentation offers a brief overview of some of the ways Driveline incorporates biomechanics data for cross-sectional talent evaluation and longitudinal player development. We focus on previous work into pitch tipping, mechanical similarity scoring, and pitcher deception.

Alex Caravan is the Vice President of Business Operations at Driveline Baseball. Having recently transitioned from the company's R&D side, Alex's past and current responsibilities have involved building predictive performance-based models and research insight-based APIs that pipe into Driveline's training, software and enterprise offerings. He has been with Driveline Baseball since January 2018 and runs a weekly podcast that focuses on the ever-changing research frontier of player development in baseball. He graduated from the University of California at Berkeley in 2015.

Kyle Wasserberger is the Principal Sport Scientist and Coordinator of Health and Performance Research at Driveline Baseball, where his responsibilities include incorporating player and coach feedback to streamline and scale how Driveline uses biomechanics data and collaborating with researchers in academia and industry to advance baseball sport science research. He holds a bachelor's degree in exercise science from Calvin University as well as master's and doctorate degrees from Auburn University.

Research Presentations

10:00-10:30 a.m. MST
RP5: Impact of Stride Length on Pelvis Energy Flow: A Retrospective Analysis
Taylor La Salle

Proper timing and sequencing of the kinetic chain is key to efficient pitching, with contributions from lower proximal segments impacting upper distal segments. Current literature of energy flow suggests that the hip joints of the pelvis and the lumbosacral (L5S1) joints contribute to improved pitching performance among pitchers. Energy flow between lower proximal segments and upper distal segments can be characterized as energy generated, absorbed, and transferred. Previous research has shown that pitchers who incorporated a shorter stride exhibited insufficient total body momentum in the intended throwing direction despite maintaining ball velocity compared to their longer stride pitches. Therefore, the purpose of this retrospective analysis is to measure energy flow through the pelvis among pitchers who maintained ball velocity despite altering their stride length.

In a randomized cross-over design, pitchers were instructed to pitch on flat ground employing either a 25% increase from their normal stride length (OS) or 25% decrease from their normal stride length (US). Kinematic and kinetic analyses were performed on high school and collegiate pitchers (n=15) using 3D motion capture synchronously with ground reaction force (GRF) technology. Segment and joint torque powers were calculated using a conventional energy flow analysis that utilizes joint torques to determine the rates of energy generation, absorption, and transfer. Back hip and lead hip were isolated to determine rates of energy flow of the pelvis with special consideration for the L5S1 joint. Paired t-tests were applied to determine energy flow differences between stride conditions.

While rates of energy generation, absorption and transfer were similar between stride lengths, we did observe some differences in GRF and energy at the hip. On the drive leg, we observed a greater amount of force applied towards home plate during OS compared to US, suggesting a greater transfer of momentum from lower body to upper body over time during OS. The stride leg produced a greater amount of braking force in the OS compared to the US. During the stride phase, we observed greater energy absorption at the back hip in the OS compared to US pitches. We also observed a greater amount of energy transfer at the L5S1 joint during arm acceleration in the OS pitches compared to US. These differences in forces from drive leg to stride leg as well as the energy differences at the back hip and L5S1 joint showcase changes that occur for pitchers with greater stride lengths or for an individual who is trying to lengthen their stride during pitching.

Taylor La Salle is the Lab Coordinator for the Point Loma Pitching Lab at Point Loma Nazarene University. He has earned a bachelor's degree in Exercise Science at Utah Valley University and a master's degree in Exercise Physiology at the University of Nebraska at Omaha, followed by a master's degree in Kinesiology from Point Loma Nazarene University. With his background in exercise physiology and biomechanics combined with his passion for baseball, his research focuses on baseball player development.

Research Presentations

2:30-3:00 p.m. MST
RP6: Evaluating Minor League Defense Using Detailed Contextual Data
Sarah Thompson

As player tracking data and the ability to better evaluate players' defensive abilities has proliferated at the major league level, the need for similar data and insights at the minor league level has increased in importance. Sports Info Solutions (SIS) has been collecting data on the minor leagues for a decade now, which has allowed for the increased understanding of defensive performance at that level via metrics like Defensive Runs Saved. Over that time we have continued to expand the data collection to include things like fielder positioning data, which has in turn allowed us to refine the metrics for defensive evaluation. This presentation will feature SIS's latest research efforts to provide insight into minor league defense through the implementation of fielder positioning data, as well as how certain contextual elements of defensive play can help inform the evaluation of a player's performance.

Sarah Thompson is a Research Analyst at Sports Info Solutions. She previously worked in the healthcare software industry and received her bachelor's degree in Mathematics from Ursinus College and her master's degree in Statistics from Villanova University.

3:00-3:30 p.m. MST
RP7: Digging Deeper into the Seam Shifted Wake Effect
John Garrett

Since 2019, Particle Image Velocimetry research has been conducted in the Utah State University Experimental Fluid Dynamics Laboratory to better understand the impact of seams on baseball aerodynamics. Early research clearly showed seams can alter the wake formation on the baseball, and at times the position of seams can shift the wake to occur earlier on one side of the ball than the other. The resulting asymmetric pressure distribution due to this Seam Shifted Wake (SSW) creates a force acting on the ball that could affect pitch movement. This was shown in lab experiments at USU and further proven by MLB's Hawk-Eye system where pitch movement is evident in both the Magnus and non-Magnus directions. The SSW effect altering pitch movement has also been proven by 3D spin axis and break differences reported by previous iterations of Rapsodo and Trackman products.

Since SSW has the ability to create unexpected movement for pitchers, it is important to understand the underlying physics of why and when it happens. As with any research, the more we uncover, the more questions are revealed. This presentation will discuss research into the impacts that seam orientation, seam height, velocity, air density, and spin rate have on the physics of the SSW effect.

John Garrett is an Analytics Engineer at Rapsodo. He received his B.S. and M.S. degrees in Mechanical Engineering from Utah State University, graduating in 2022. He conducted research on baseball aerodynamics and the Seam Shifted Wake effect in the USU Experimental Fluid Dynamics Laboratory for three years as an Undergraduate and Graduate Research Assistant to Dr. Barton Smith. Before completing his M.S. degree, he served as a Baseball Data Intern with Major League Baseball.

Research Presentations

3:30-4:00 p.m. MST
RP8: Pitchers Hitting: A Requiem
Rob Mains

The 2022 season, the first full season with the designated hitter in both leagues, marked the end of pitchers batting. Several reasons were given for adopting the rule throughout MLB: Reduce pitcher injuries, lengthen benches by removing the need for a battalion of pinch-hitters, create uniformity between the leagues as interleague play increases. But, obviously, the main motivation was to increase offense, thereby scoring, increasing the sport's popularity and the revenues associated with that.

On the latter score, the change would appear to be a bit of a misfire. There were 4.34 runs scored per National League game with the universal DH in 2022 compared to 4.46 in 2021. That can be considered a victory, however, when contrasted to the much steeper decline, from 4.60 to 4.22, in the American League.

In this presentation, I will discuss how the elimination of pitcher hitting affected offense in the National League in 2022 and evaluate how it compared to the American League. In particular, I will examine how historical differences between the two leagues' strategies affecting both hitters and pitchers evolved during the 48 seasons during which they played under separate rules and note how the elimination of pitchers hitting could accelerate some trends while reversing others. I will examine scoring patterns, pitcher usage, substitutions, and the role of the designated hitter.

Rob Mains is a writer for Baseball Prospectus. His "Veteran Presence" column runs twice a week. He is a former equities analyst and was a finalist for the 2018 SABR Analytics Conference Research Award for Historical Analysis/Commentary. He is a SABR Analytics Certification course reviewer.

Sunday, March 12

9:30-10:00 a.m. MST
RP9: How to Effectively Communicate Baseball Research
Rebekah Callari-Kaczmarczyk

Communicating findings effectively is an essential element of disseminating good research. Effective research communication is understandable and convincing, but producing writing that fulfills these criteria can be challenging especially for a newer analyst. Research and technical writing is an extensive field of study with a multitude of readily accessible books, research publications and websites, but resources for the baseball analytics community specifically are lacking. In addition, many of the resources readily available to the public are based in opinion rather than research. In this presentation, I will discuss some of the main challenges of writing effectively about baseball analytics and will present a genre-based approach that can be used to make writing more understandable and convincing to fellow analysts and baseball decision-makers.

A genre is the communication of a discourse community (a group of people who communicate while working towards a common goal) and is typically defined by unique characteristics such as shared language patterns and vocabulary. Using these unique language characteristics marks an individual as part of the discourse community, making their discourse more understandable and convincing to other community members. The baseball analytics discourse community is composed of both researchers and readers from diverse backgrounds with varying levels and types of knowledge. For example, a student may have limited experience writing about their independent research and have received feedback from fellow researchers that their writing is disorganized or "not academic" enough while a statistician may have extensive writing experience

Research Presentations

in academia but struggle to communicate these findings in a way that is accessible to baseball leadership. Understanding the intended discourse community – what the readers know about the topic and what they expect – enables the writer to communicate in a convincing manner.

This presentation will provide research-based tools to help attendees clearly define their intended audience and produce the patterns or structures expected by that audience. I will review the typical structure of a research write-up and how the sections should vary depending on the intended discourse community. Although this will be most helpful to those new to writing about baseball analytics, this is relevant to anyone seeking to improve their written research communication.

Rebekah Callari-Kaczmarczyk is a Lead Software Engineer in Research and Development with the Philadelphia Phillies. Before transitioning to software engineering, she taught academic and technical communication to graduate students at Duke University. She has a certificate in Data Analytics and a master's degree in Applied Linguistics: TESOL from Georgia State University.

10:00-10:30 a.m. MST
RP10: Using Seam Orientation Data from Optical Sensors for Pitch Modeling and Design
Glenn Healey

Researchers at Utah State University have used Particle Image Velocimetry to show that a baseball's seams can advance boundary layer separation which causes a Seam Shifted Wake (SSW) force on the ball. High-resolution optical ball-tracking sensors enable measurement of the orientation of a ball's seams relative to the spin vector. This orientation can be characterized using various representations including spherical coordinates and Mollweide projections. We use a large set of seam orientation and pitch trajectory data acquired by the Yakkertech optical ball-tracking system to build a physical model for the dependence of the SSW force on seam orientation, pitch release parameters, and the atmospheric conditions. The result is a new aerodynamic model that includes all of the forces that are known to affect the trajectory of a baseball. We show that this model provides more accurate predictions for pitch trajectories than current models. The new model can be combined with intrinsic pitch values to enable the efficient design and optimization of pitches. We demonstrate the use of seam orientation to manipulate pitch shape using the BallR visualization environment. By quantifying the relationship between seam orientation, pitch trajectory, and pitch value, we also provide analysts with a new tool for evaluating, comparing, and forecasting the performance of pitchers.

Glenn Healey is a Professor of Electrical Engineering and Computer Science at the University of California, Irvine where he is director of the Computer Vision Laboratory. He received the B.S.E. degree in Computer Engineering from the University of Michigan and the M.S. degree in computer science, the M.S. degree in mathematics, and the Ph.D. degree in computer science from Stanford University. Dr. Healey's work focuses on combining physics, statistical signal processing, and machine learning methods for the development of algorithms that extract information from large sets of data. He has been elected a Fellow of IEEE and SPIE.

Research Presentations

10:30-11:00 a.m. MST
RP11: Quantifying In-Game Command: Observed Data Meets Intended Location
Ryan Reinsel

Teams and analysts have been trying to find a way to quantify command for decades. These processes have included measuring from the middle of home plate, from where the catcher setup, to using walk rate, edge percentage and manual tagging at both the college and professional level. Metrics have been created to measure how command works with called strikes in and out of the zone, but nothing has ever truly captured how well a pitcher can command his pitches during a game. With the advent of new data available, we can now know the intended location of each pitch. The combination of knowing the plate location height and side of a pitcher's intended location and the observed location allow for new metrics to be added for each pitch, such as pitch miss distance and pitch miss direction.

This presentation will explore how definitive location data opens up a new sector of command metrics that can be developed. From a game perspective we can look at Execution Rate, Sequential Command Grades, and overall Command Scores. We'll also take a look at the relationship with ballistic pitch and hit data regarding how well pitchers can not only throw a pitch where they wanted to, but does it have the spin and movement profile that is expected. For example, seeing how a pitcher's breaking ball movement and spin profile might change across the plate will provide explanation and new analysis for player development. From a game level, looking at command grades for a starting pitcher after having a low score compared to league average, yet taking a no-hitter into the 6th inning gives us new questions to ask about the connection between "Stuff" and actually hitting spots.

Defining in game command with precise data gives the industry an opportunity to grade, improve, teach, and scout pitchers in a unique way. From identifying a pitcher's command margin of error based off of their current arsenal, to leveraging seam-shifted wake models to better build and simulate an optimal sequence, we'll explore an untapped data avenue that is ready to be further researched.

Ryan Reinsel is Vice President of Innovation at BaseballCloud and the creative mind behind the baseball industry's newest data visualization software, BallR and PitchR. Equipped with a Fine Arts degree and a passion for technology, Ryan is a proven innovator at creating contextual applications to drive player development forward at the professional and collegiate level. Prior to BaseballCloud, Ryan held positions in multiple sectors of baseball including a role in Research and Player Development for the Chicago Cubs at their Arizona complex.

11:45 a.m.-12:15 p.m. MST
RP12: Trends in MLB Pitch Tempo Since the Start of Pitch Tracking
Andy Andres

This past year, MLB and Baseball Savant released pitch tempo data (the time between pitches) in advance of the new rules for the 2023 MLB Season. Statcast has recorded a time stamp for the start of the pitch when released, and Savant made the considered choice to include "takes" only when measuring pitch tempo, defined as "pitch release to pitch release." Year by year league, team and individual player trends since 2010, player service time and how it relates to recent measures of pitch tempo, and some early spring training data will be presented.

Andy Andres is on the faculty of Boston University where he has been teaching various natural sciences, mathematics, and data science courses for more than two decades. He developed and taught the

Research Presentations

first college course in Baseball Analysis and Sabermetrics at Tufts University where he was a Visiting Lecturer in Sabermetrics. He also initiated, designed, and taught the highly popular BU MOOC "Sabermetrics 101: Introduction to Baseball Analytics" to more than 70,000 registered students on the edx.org platform. His former students and mentees occupy front offices in the NBA, MLS, and MLB, where one was recently named General Manager. Andy is also a part-time MLB stringer and pitch clock operator at Fenway Park, a Data Analyst for BaseballHQ.com, and former Head Coach/Lead instructor for the MIT Science of Baseball Program. He has been a member of SABR since 2003.

12:15-12:45 p.m. MST
RP13: Noisy Judgments: A Probability Surface-Based Analysis of Umpiring Variability
Emily-Anne Patt and James Stockton

In any repeated evaluation of events, we expect variability in expert assessment. Within a corpus of data spanning multiple experts' evaluations across multiple events, this noise manifests as both interexpert and intraexpert variability in the assessed outcomes. Called balls and strikes in Major League Baseball provide a clear example of this noise in human judgment. To assess the prevalence, magnitude, and trends of this noise, we utilized 15 seasons of available PITCHf\x and Statcast data showing the locations of ~5.3 million individual pitches called either ball or strike. We generated a prior probability surface across the strike zone representing the average strike zone as called by the entire umpiring corps. This surface shows the actual behavior of these experts when making evaluations, providing a novel methodology for assessing trends and variability across MLB's umpire corps. We use this surface to evaluate changes in the actual strike zone over time as well as individual umpire performance. We present details of the probability surface construction method along with initial results stemming from its application. We preview future applications of the best fit strike zone surface model for in-depth player performance. These detailed analyses are made possible by the carefully designed fitting methodology for the strike zone probability surface. Accounting for sampling error and providing a well-defined functional form for the optimizer allows the model to arrive at reliable results from smaller subsets of data than previous methods. Pitcher-Batter, Umpire-Pitcher, and similar specific match-ups are available for study.

Emily-Anne Patt is the manager for quantitative intelligence and methodologies supporting security and resilience at Alphabet, Inc. based in Washington, D.C. Her background is in econometrics and financial economics, with prior experience at the US Department of State, Federal Reserve Board of Governors, US Department of Treasury, and as a data science consultant for US federal government clients.

James Stockton, PhD, is the lead data scientist at Altamira Technologies supporting the United States Air Force Chief Data and AI Office based in northern Virginia. His background is in astronomy and astrophysics and he has worked as a data scientist in private industry, academia, and supporting federal customers in the IC/DoD space for the past 10 years.

Research Presentations

12:45-1:15 p.m. MST
RP14: A Markov Chain Matchup Model Using Multi-variable Approach
Jeff Jin and Chris Zexin Chen

Markov Chain is a widely used analytical framework in baseball that can be used to estimate run scored and optimize lineups. In 2021, A novel approach was introduced at SABR by Connor Turner who used Markov Chain to simulate a hypothetical matchup between Mike Trout and Justin Verlander. Here, we would like to introduce a refined version of the Markov Chain matchup model (PA Markov Chain) as an expansion on Turner's work.

The PA Markov Chain uses a player's historical transitional probabilities from one count to another to model the expected event probabilities of his future PA. A player's tendency to move through a plate appearance is characterized by a probability matrix (T). We can then use T, through the Markovian Process to generate a stationary matrix (P), which represents the expected PA event probabilities of a certain player vs. the league. Turner proposed to take the weighted average between a pitcher's transition matrix T_p, and a batter's transition matrix T_b, as the new transition matrix T_{pb} as they were to play each other. We expanded on this method by first testing its validity using Muli-Outcome Brier Score, then incorporating additional variables (i.e. platoon, park factors, temperature, etc.) before fitting machine learning models to improve the accuracy. Our current model achieved noticeable improvement against Turner's baseline approach and is on par with the Log5 matchup model.

This work is still ongoing but the preliminary results are promising. We are currently in the process of incorporating additional variables and expect to make further improvements. This model will be beneficial to the analysis and prediction of matchup results. It can be further built into full-game Markov Chain simulation to perform team level analysis.

Jeff Jin is a recent graduate from the MS Sports Business program at NYU and the co-founder of NYU Sports Analytics and Technology Association. He is a self-taught sports analyst and data scientist whose recent work is featured in the presentation "Win With AI: Exposing Winning Strategies In Sports" at the 2023 MIT Sloan Sports Analytics Conference. He will be joining Cleat Street Capital as a Baseball Analyst.

Chris Zexin Chen is a current data science undergraduate at NYU and the co-founder of NYU Sports Analytics and Technology Association. He is a sports enthusiast with a strong interest in analytics and has experience in modeling basketball and baseball. He is part of the demo team for the presentation "Win With AI: Exposing Winning Strategies In Sports" at the 2023 MIT Sloan Sports Analytics Conference.

SMT DATA CHALLENGE 2023

In baseball, defensive analytics present challenges different from pitching and hitting. Not only is spatial information essential, but one fielder's location, movement, and decision-making can materially impact those of other players.

For this year's Data Challenge, we want you to **consider a game situation involving the interaction of two or more defensive teammates.** Moreover, since action is central to all sports, we want you to **frame your analysis as a story**. Show us how your work informs and contributes to others' knowledge of and appreciation for the game.

We encourage you to consider any situation that interests you, but meanwhile, here are a few sample questions that you may use or draw inspiration from:

- When making an outfield assist, does hitting the cutoff man really make a difference?
- Compare double plays using different combos of Players A, B, C, and D. Which are most/least effective? Why?
- How does the range of Infielder A affect surrounding players? Infielder B?

| ELIGIBILITY

- The Data Challenge is open to STUDENTS ONLY. You must be enrolled as a high school, undergraduate, or graduate student for the Fall 2023 semester. Participants are expected to register using a .edu email address or similar. If you have questions, please contact smt.datachallenge@smt.com.
- Participants must be at least 18 years of age.

| REGISTRATION (Deadline: May 31)

Teams of up to four will compete in High School/Undergraduate or Graduate Divisions. You can register on the CMSAC 2023 website or by going directly to the 2023 SMT Data Challenge Registration Form. Data will be made available on or after June 1.

| SUBMISSION (Deadline: August 31)

- A short, narrative paper on your study in PDF format (max: 2000 words)
- A GitHub repo link containing code files and .csv files with results

A panel comprised of people from academia, industry, journalism, and sports (including team executives, players, umpires, scouts etc.) will judge your submissions based on the following criteria:

- How rich and textured (not necessarily big or small) is the question being asked?
- How applicable is the analysis?
- How appropriate were the methods used?
- How specific and narrative was the topic of investigation?
- How well did you communicate your findings? This includes both written text and visualizations. How did the use of facts, data-supported narratives, anecdotes, visual aids, etc. buttress storytelling?

| FINALISTS (Announced: October 4) AND PRESENTATIONS (November 11)

Three finalists from each division will present their work at CMSAC 2023 on November 11. Presentations will undergo a second round of judging, and division winners will be announced at the end of the conference.

| WINNERS AND PRIZES

- Division winners: $1000
- Runners-up: $250

| QUESTIONS?

If you have questions, please contact smt.datachallenge@smt.com.

| IMPORTANT DATES

MAY 31: REGISTRATION DEADLINE
DATA AVAILABLE ON OR AFTER JUNE 1.
AUG 31: SUBMISSION DEADLINE
OCT 4: FINALISTS ANNOUNCED
NOV 11: FINALIST CMSAC 2023 PRESENTATIONS

OPTIONAL ANALYSIS DEMO/Q&A SESSIONS:

- *JUNE 14*
- *JULY 5*
- *JULY 26*
- *AUGUST 16*

Diamond Dollars Case Competition

The SABR Analytics Conference is pleased to host the unique **Diamond Dollars Case Competition**. High school and college students from schools across the country will compete against each other by preparing an analysis and presentation of a baseball operations decision — the type of decision a team's GM and his staff is faced with over the course of a season. The case was developed by Vince Gennaro, former president of SABR's Board of Directors, co-founder of the SABR Analytics Conference, author of *Diamond Dollars: The Economics of Winning in Baseball*, and consultant to MLB teams. The Case Competition is the first national competition to be based solely on baseball operations issues.

Four- to five-person student teams will be asked to evaluate a baseball operations case problem. Once the student team has prepared its case, they will have the opportunity to present their analysis and recommendations to a panel of judges, which will include MLB front office executives. They will have a dialog, receive feedback, and ultimately be evaluated based on the quality of their insights and analysis.

The competition will be divided by room, and winners will be recognized from each room. The competition will take place beginning at 8:30 a.m. MST on Friday, March 10. An encore presentation by the overall winning team will be delivered as part of the general session at 10:30 a.m. MST on Saturday, March 11.

Participating schools

- **Arizona State University** (Tempe, AZ)
- **Elon University** (Elon, NC)
- **Maggie L. Walker Governor's School** (Richmond, VA)
- **NYU-Tisch Institute for Global Sport** (New York, NY)
- **Rice University** (Houston, TX)
- **Syracuse University** (Syracuse, NY)
- **The Branson School** (Ross, CA)
- **Tufts University** (Medford, MA)
- **University of Arizona** (Tucson, AZ)
- **University of Florida** (Gainesville, FL)
- **Washington University in St. Louis** (St. Louis, MO)

Presentation schedule
Friday, March 10 ♦ Beus Center

♦ **Room 240** —8:55 a.m.: Syracuse University #1; 9:30 a.m.: NYU-Tisch Institute; 10:05 a.m.: The Branson School; 10:45 a.m.: University of Arizona #3; 11:20 a.m.: Maggie L. Walker Governor's School #5

♦ **Room 250** — 8:55 a.m.: Maggie L. Walker Governor's School #1; 9:30 a.m.: Syracuse University #2; 10:05 a.m.: University of Arizona #1; 10:45 a.m.: Maggie L. Walker Governor's School #1; 11:20 a.m.: Rice University

♦ **Room 442** — 8:55 a.m.: University of Florida; 9:30 a.m.: Maggie L. Walker Governor's School #2; 10:05 a.m.: Washington University of St. Louis; 10:45 a.m.: Tufts University; 11:20 a.m.: Syracuse University #3

♦ **Room 450** — 8:55 a.m.: Elon University; 9:30 a.m.: University of Arizona #2; 10:05 a.m.: Maggie L. Walker Governor's School #3; 10:45 a.m.: Syracuse University #4; 11:20 a.m.: Arizona State University

ARE YOU READY TO GET SABR CERTIFIED?

Whether you are a student or professional looking to set yourself apart as you break into the ultra-competitive baseball industry, or a broadcaster, coach, scout, or fan interested in gaining a better understanding of the present state and future of the game, the **SABR Analytics Certification** was created with you in mind.

With online, on-demand curricula offering professional development and learning opportunities, courses advance through progressively deeper dives into the world of baseball data and analytics, with virtual instruction provided by industry thought leaders and the opportunity to present your own analysis to be reviewed by esteemed professionals.

Learn more and register today:

SABR.ORG/ANALYTICS/CERTIFICATION

Research Awards

The **SABR Analytics Conference Research Awards** recognize baseball researchers who have completed the best work of original analysis or commentary during the preceding calendar year. Here are the 2023 finalists:

Contemporary Baseball Analysis

- Lindsey Adler, "Why the Yankees pitching staff is embracing a new slider and how the 'whirly' is different from a traditional slider," The Athletic, April 4, 2022.
- Michael Ajeto, "We're Talking About Seam-Shifted Wake Wrong," Baseball Prospectus, May 17, 2022.
- Davy Andrews, "Steven Kwan, Geraldo Perdomo, and the Victor Robles Problem," FanGraphs, December 19, 2022.
- Ben Clemens, "How Good Are Those Probabilities on the Apple TV+ Broadcasts?" FanGraphs, May 26, 2022.
- Ethan Rendon, Elijah Emery, Will Sugar and Tieran Alexander, "The Mystic Art of Pitch Tunneling," Prospects Live, August 23, 2022.

Contemporary Baseball Commentary

- Michael Baumann, "The Playoffs Aren't Too Big. The League Is Way Too Small," FanGraphs, October 4, 2022.
- Zach Crizer, "The Phillies are Dave Dombrowski's 5th World Series Team. Should More Teams Play His Version of Moneyball?" Yahoo! Sports, October 31, 2022.
- Patrick Dubuque, "One Step Behind," Baseball Prospectus, October 3, 2022.
- Hannah Keyser, "MLB lockout: There's no way to force team owners to put wins over dollars," Yahoo! Sports, February 16, 2022.
- Ginny Searle, "So Do Not Become Proud, But Fear," Baseball Prospectus, June 7, 2022.

Historical Analysis/Commentary

- Donna L. Halper, "Marvels or Menaces: How the Press Covered 'The Lady Baseballists,' 1865-1915," *SABR Baseball Research Journal*, Spring 2022.
- Jerry Nechal, "Were Pitchers More Likely to Throw at Black Batters? 1947–66," *SABR Baseball Research Journal*, Fall 2022.
- Marc Normandin, "1994 Explains What 'Labor Peace' Never Could," Baseball Prospectus, February 28, 2022.
- Mike Petriello, "How many games would Springfield's ringers win?" MLB.com, February 19, 2022.
- Leander Schaerlaeckens, "The Forgotten Men Behind The Ideas That Changed Baseball," Defector, November 14, 2022.

John Dewan Defensive Analytics Award

- Russell A. Carleton, "So You've Decided to Ban the Shift," Baseball Prospectus, March 9, 2022.
- Justin Choi, "The Blue Jays Make A Big Shift In Making Big Shifts," FanGraphs, May 11, 2022.
- Alex Eisert, "More Than a Putout Puzzle: Revisiting the Problem of Outfield Alignment," FanGraphs, December 5, 2022.

Voting for the winners was conducted online at SABR.org, BaseballProspectus.com, FanGraphs.com, and IBWAA.com, with results weighted equally at 25%.

Sponsors

SMT

BASEBALL REFERENCE

Rapsodo

KINATRAX

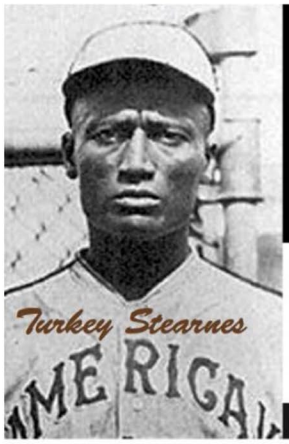
UNIVERSITY OF THE PACIFIC
Eberhardt School of Business

CLEAT STREET

BASEBALLCLOUD

ASU Sandra Day O'Connor College of Law
Arizona State University
Allan "Bud" Selig Sports Law and Business Program

Teams



JERRY MALLOY

NEGRO LEAGUE CONFERENCE

SABR Jerry Malloy Negro League Conference July 20-23, 2023 Detroit, Michigan

Hosted by SABR's Negro Leagues Committee since 1998, this event has been the only symposium dedicated exclusively to the examination and promotion of Black baseball history. All fans are welcome to attend!

★ Hear from former Negro Leagues players

★ Groundbreaking, peer-reviewed presentations on Black baseball

★ Tour of historic Hamtramck Stadium

★ Detroit Tigers game

★ Support student scholarships, library grants, and grave markers

Register now at SABR.org/malloy

The Society for American Baseball Research invites you to

SABR National Convention
July 5-9, 2023
Palmer House Hilton
Chicago, Illinois

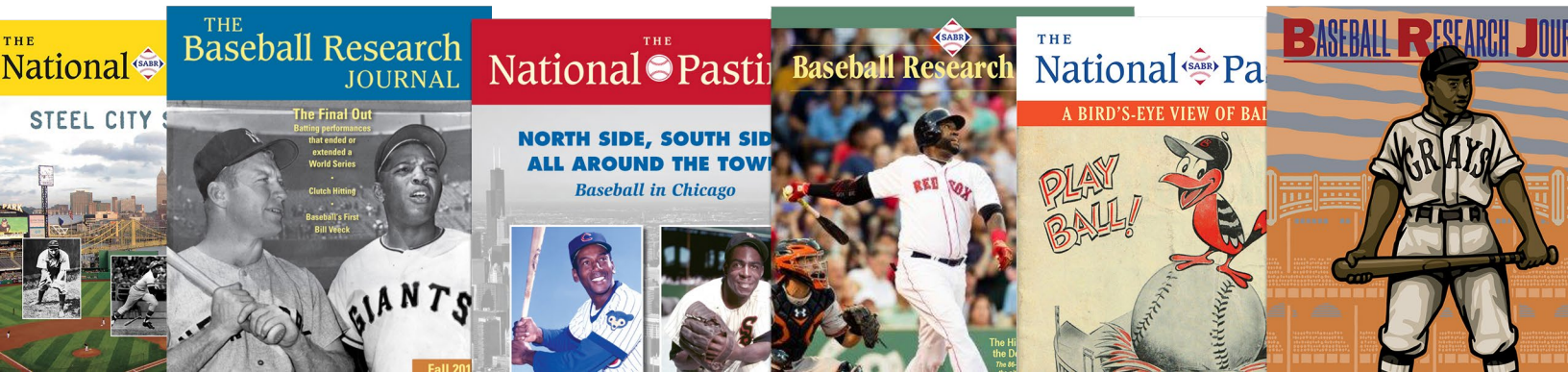
- Top speakers and panelists from around the baseball world (players, executives, historians, authors, media, and more)
- Groundbreaking, peer-reviewed presentations on baseball history, analytics, and the current state of the game
- Chicago White Sox game on Friday night
- Single-day registration available
- Centrally located in the heart of Chicago's Downtown Loop
- Walking distance from many tourist attractions

Contact Scott Bush at 602-496-1461 or sbush@sabr.org for sponsorship or vendor room opportunities.

SABR.org/convention

WHY SHOULD YOU JOIN SABR?

SABR member benefits include:

- Two editions (spring and fall) of the *Baseball Research Journal*
- *The National Pastime* convention journal
- New and classic e-books published by the SABR Digital Library, all FREE to members
- "This Week in SABR" e-newsletter every Friday
- Research committees and discussion groups
- Publish your research in peer-reviewed SABR journals
- Discounts on national conferences such as the SABR National Convention, the SABR Analytics Conference, and the Jerry Malloy Negro League Conference
- Join one of 80+ regional SABR chapters in the U.S., Canada, or around the world
- FREE access to *The Sporting News* archives online through Paper of Record
- Contribute to SABR Baseball BioProject or Games Project

LEARN MORE AT SABR.ORG/JOIN