Player evaluation based on game performance

Scott Powers

November 20, 2024

RICE UNIVERSITY
Sport Analytics

Outline

1. Player evaluation based on game performance
   - Descriptive metrics
   - Predictive metrics
2. Identifying undervalued players in the minor leagues w/ data
   - Targeted development plans
3. My Research
   - Basic descriptive metrics in volleyball
   - Advanced predictive metrics in baseball
   - Advanced descriptive metrics in soccer

Outline

1. Player evaluation based on game performance
    - Descriptive metrics
    - Predictive metrics
2. Identifying undervalued players in the minor leagues w/ data
    - Targeted development plans
3. My Research
    - Basic descriptive metrics in volleyball
    - Advanced predictive metrics in baseball
    - Advanced descriptive metrics in soccer

What is player evaluation?

"How good is this player?"

Scouting and roster construction

Trades and free agency

The draft

- If we add this player, how many more games will we win?

Big Picture

Player Evaluation

Training Measurements

Game Performance

Big Picture

Player
Evaluation

Training
Measurements

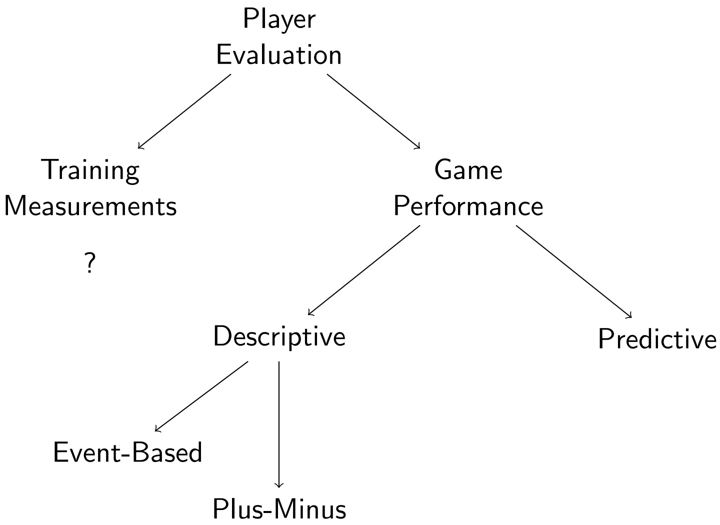
?

Game
Performance

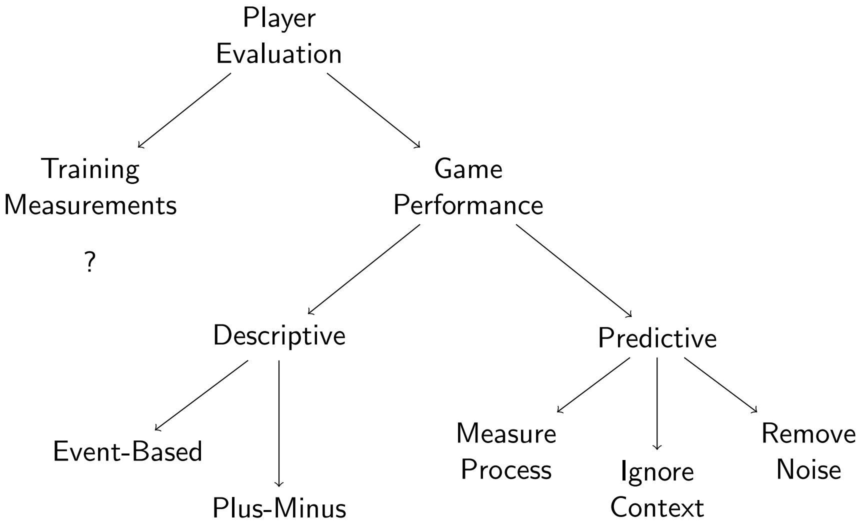
Big Picture

- Player Evaluation
  - Training Measurements
    - ?
  - Game Performance
    - Descriptive
    - Predictive

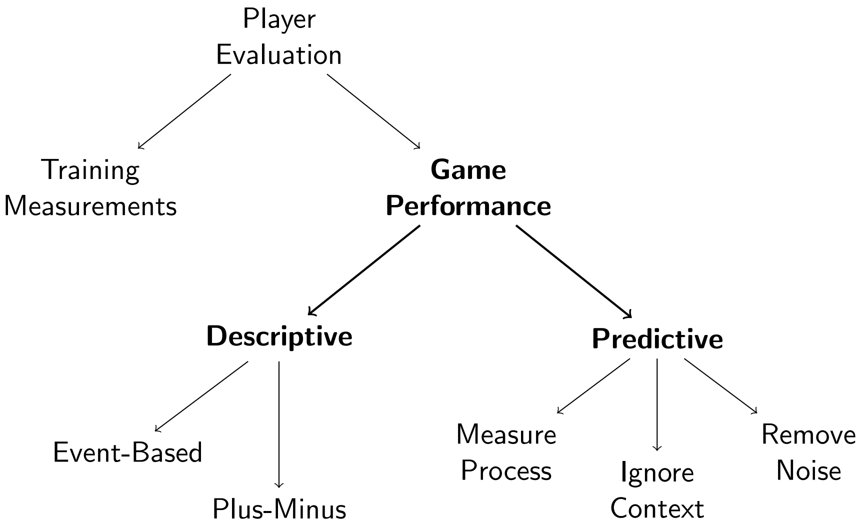
Big Picture


Player Evaluation
Training Measurements
?
Game Performance
Descriptive
Predictive
Event-Based
Plus-Minus

Big Picture


Player Evaluation
Training Measurements
?
Game Performance
Descriptive
Predictive
Event-Based
Plus-Minus
Measure Process
Ignore Context
Remove Noise

Big Picture


Player Evaluation
Training Measurements
Game Performance
Descriptive
Predictive
Event-Based
Plus-Minus
Measure Process
Ignore Context
Remove Noise

Player evaluation based on game performance

Two guiding principles:

1. (Descriptive) Measure impact on team wins
    - How do player actions cause us to win (or lose) games?

2. (Predictive) Separate signal from noise
    - How will the player perform in the **future?**

Player evaluation based on game performance

Two guiding principles:

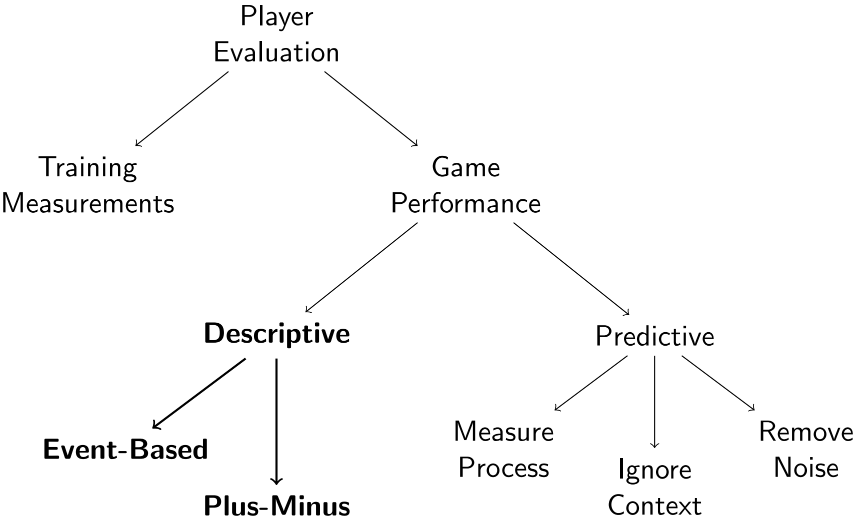
1. (Descriptive) Measure impact on team wins
   - How do player actions cause us to win (or lose) games?
2. (Predictive) Separate signal from noise
   - How will the player perform in the **future?**

Example (from *Moneyball*):

Batting Average $= \frac{H}{AB}$
- Less descriptive
- Less predictive

On-Base Percentage $= \frac{H+BB+HBP}{AB+BB+HBP+SF}$
- More descriptive
- More predictive

Big Picture


Player
Evaluation
Training
Measurements
Game
Performance
Descriptive
Predictive
Event-Based
Plus-Minus
Measure
Process
Ignore
Context
Remove
Noise

Descriptive Approach #1: Event-Based

Approach: Estimate win probability (or score expectancy) using event data. As score expectancy changes, assign credit to players responsible for those actions.

Methods: Markov chain (for low-dimensional game state) or machine learning (for high-dimensional game state)

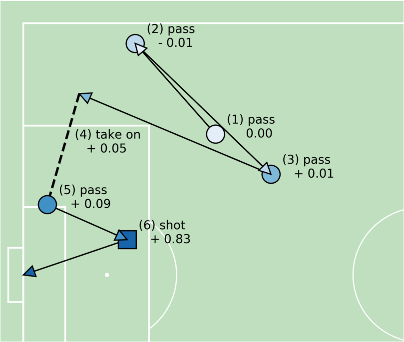
Baseball Example: Base-out run expectancy

Runners	0 Outs	1 Out	2 Outs
Empty	0.461	0.243	0.095
1 _ _	0.831	0.489	0.214
_ 2 _	1.068	0.644	0.305
1 2 _	1.373	0.908	0.343
_ _ 3	1.426	0.865	0.413
1 _ 3	1.798	1.140	0.471
_ 2 3	1.920	1.352	0.570
1 2 3	2.282	1.520	0.736

https://thebaseballscholar.com/2017/08/14/sabermetrics-101-re24/

- Goal expectancy estimated via Markov chain model
- Bases empty, 0 out, single $= 0.831 - 0.461 = +0.370$ runs
- Bases loaded, 2 out, strikeout $= 0 - 0.736 = -0.736$ runs

Soccer Example: VAEP (Decroos et al. 2019)



Decroos et al. 2019

- Goal expectancy estimated via machine learning

See also: Expected Threat (Singh 2018)

Descriptive Approach #2: Plus-Minus

Approach: Infer player contributions based on outcomes when they are playing vs. when they are off. Use regression to control for quality of teammates and quality of opposition.

Methods: Ridge regression, hierarchical Bayesian modeling, or linear mixed-effects regression

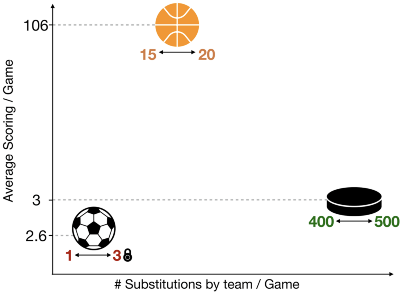
Basketball Example: Regularized Adjusted Plus-Minus

$$\begin{bmatrix} 40 \\ -25 \\ 60 \\ 0 \\ -30 \\ -12.5 \\ -12.5 \\ -30 \\ 50 \\ 30 \\ -50 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & -1 & -1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & -1 & -1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 & 1 & -1 & 0 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & -1 & -1 & -1 \\ 0 & 1 & 1 & 1 & 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & -1 & -1 & 0 & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 & 0 & -1 & -1 \\ 1 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 & -1 & 0 & 0 & -1 & -1 \\ 1 & 0 & 1 & 1 & 0 & -1 & -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & -1 & -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8 \\ \beta_9 \\ \beta_{10} \end{bmatrix}$$

(Jacobs 2017)

- Works well (many substitutions and scoring events)

Soccer Example: **Augmented** Adjusted Plus-Minus

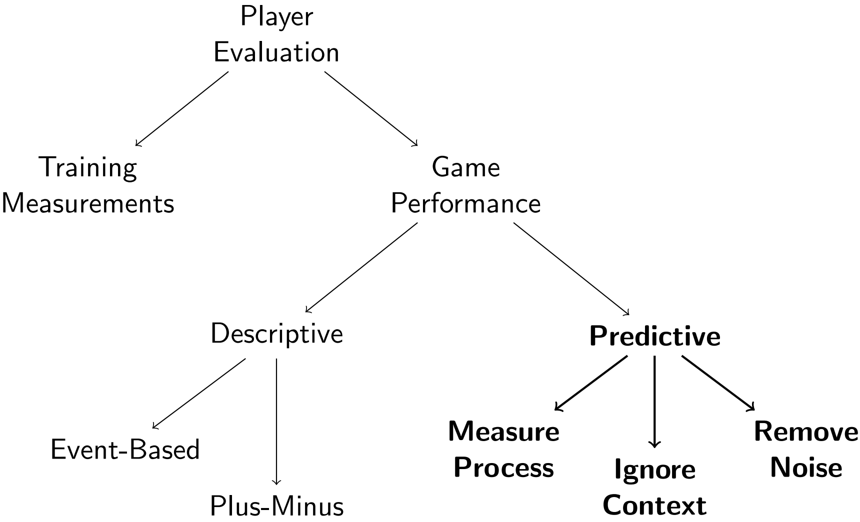


(Matano et al. 2023)

- Problem: Soccer has few scoring events AND few subs
- Solution: Set your prior belief using additional information (e.g. video game player ratings, Matano et al. 2023)

See also: Box Plus-Minus (Myers 2020)

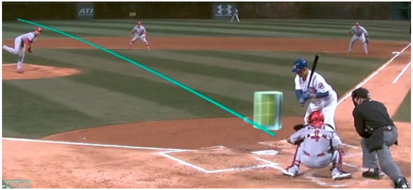
Big Picture


Player Evaluation
Training Measurements
Game Performance
Descriptive
Predictive
Event-Based
Plus-Minus
Measure Process
Ignore Context
Remove Noise

Predictive Strategy #1: Measure process, not results

Strategy: Measure more granular data. Create predictions of outcomes using the more granular data. These predicted outcomes (usually) have higher signal-to-noise ratio than actual outcomes.

Baseball Example: Pitch outcome modeling



Soccer Example: xG vs. goals

Predictive Strategy #2: Use context-neutral metrics

Strategy: (Sometimes) It helps to ignore the context in which an event occurred. This happens when the value of the event is sensitive to context, but the performance of the event is not.

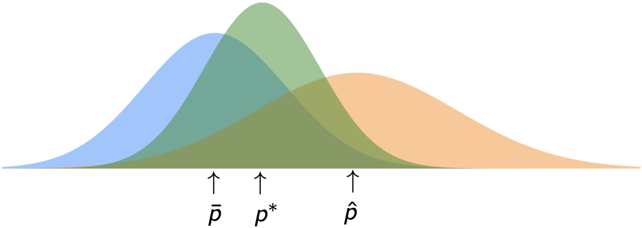
Baseball Example: Linear weights

2015 Linear Weights (Relative To Outs)	
Event	Run Value
BB	0.55
HBP	0.57
1B	0.70
2B	1.00
3B	1.27
HR	1.65

https://library.fangraphs.com/principles/linear-weights/

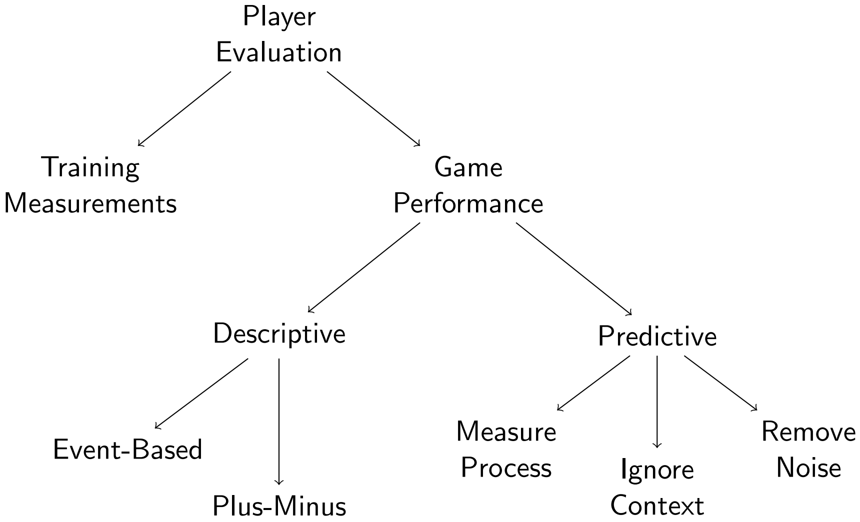
Predictive Strategy #3: Extract signal from noise

Strategy: Performance = Skill + Luck. Regression to the mean estimates a player's skill based on their performance. The amount of mean regression depends on the signal-to-noise ratio of the stat.



Soccer Example: The average take-on success rate is $\bar{p} = 45\%$. If a player is successful in 30 of 46 attempts ($\hat{p} = 65\%$), we expect his future success rate to be $p^* = 52\%$.

Big Picture


Player Evaluation
Training Measurements
Game Performance
Descriptive
Predictive
Event-Based
Plus-Minus
Measure Process
Ignore Context
Remove Noise

Outline

1. Player evaluation based on game performance
   - Descriptive metrics
   - Predictive metrics
2. Identifying undervalued players in the minor leagues w/ data
   - Targeted development plans
3. My Research
   - Basic descriptive metrics in volleyball
   - Advanced predictive metrics in baseball
   - Advanced descriptive metrics in soccer

Identifying undervalued players in the minor leagues

Framework:
Player has overall skill X and specific skills X_1, X_2, X_3.

$$X = f(X_1, X_2, X_3) = X_1 + X_2 + X_3$$

Example:

- X = overall batting skill
- X_1 = plate discipline / swing decision-making
- X_2 = contact / bat-to-ball skills
- X_3 = power

If we know our coaches are particular good at developing X_3, what implications does that have for player acquisition?

Identifying undervalued players in the minor leagues

Suppose our player development department can't increase X_1/X_2, but they can increase X_3 by 1 unit (run) over some time period.

Player A: $X_1 = 1, X_2 = -1, X_3 = 1 \quad \Rightarrow X = 1$
Player B: $X_1 = 1, X_2 = 1, \quad X_3 = -1 \Rightarrow X = 1$

After player development intervention...
Player A: $X_1 = 1, X_2 = -1, X_3 = 2 \Rightarrow X = 2$
Player B: $X_1 = 1, X_2 = 1, \quad X_3 = 0 \Rightarrow X = 2$

So it's a wash, unless...

- Player development can help Player B *more* than Player A, or
- $f(X_1, X_2, X_3)$ is a "cliff" function rather than a linear function

Outline

1. Player evaluation based on game performance
   - Descriptive metrics
   - Predictive metrics
2. Identifying undervalued players in the minor leagues w/ data
   - Targeted development plans
3. My Research
   - Basic descriptive metrics in volleyball
   - Advanced predictive metrics in baseball
   - Advanced descriptive metrics in soccer

Project #1: Basic descriptive metrics in volleyball

Joint work with Luke Stancil and Naomi Consiglio

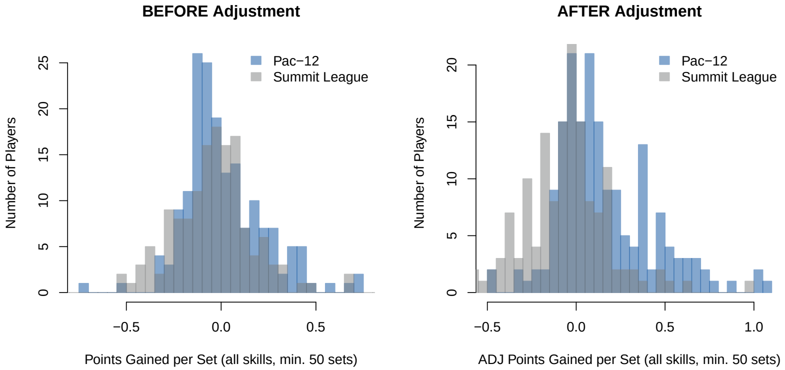
- 4,147 matches, 600K+ points, 5M+ contacts, ∼6,000 players
- We used a Markov chain to estimate point win probability

Player	Skill	Eval	State	P(Sideout)
Anna Deeber	Serve		(S, SV)	57%
Emma Halter	Reception	#	(R, R#)	63%
Saige K.-Torres	Set	#	(R, R#S#)	64%
Molly Phillips	Attack		(R, R#S#A)	64%
Raquel Lazaro	Dig	+	(S, D+)	49%
Elena Scott	Set	#	(S, D+S#)	47%
Claire Chaussee	Attack		(S, D+S#A)	47%
Kayla Caffey	Block	+	(R, B+)	56%
Phekran Kong	Dig	!	(S, D!)	51%
Raquel Lazaro	Set	#	(S, D!S#)	51%
Claire Chaussee	Attack		(S, D!S#A)	51%
Point Louisville				**0%**

Project #1: Basic descriptive metrics in volleyball

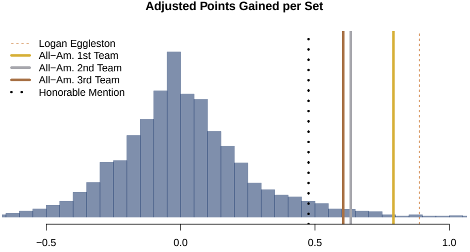
Joint work with Luke Stancil and Naomi Consiglio

- We used a hierarchical linear mixed-effects regression to adjust each player's performance based on her quality of competition



Project #1: Basic descriptive metrics in volleyball

Joint work with Luke Stancil and Naomi Consiglio

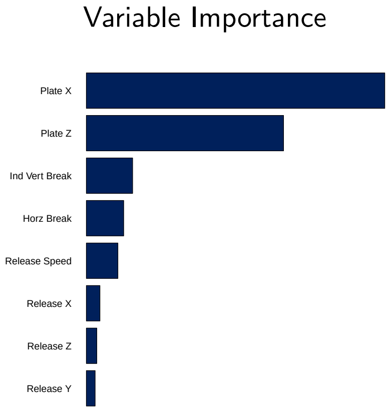


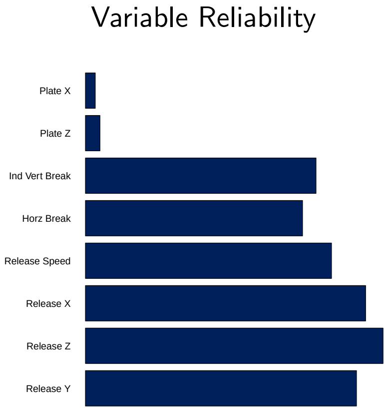
PLAYER	TEAM	CONF	POS	SETS PLAYED	POINTS GAINED PER SET (ADJ)	SERVE PG*/S	PASS PG*/S	SET PG*/S	ATTACK PG*/S	BLOCK PG*/S
Brooke Nuneviller	Oregon	Pac-12	OH	122	**+1.09**	+0.07	+0.41	+0.00	+0.56	+0.04
Mckenna Melville	Central Florida	AAC	OH	104	**+1.09**	-0.14	+0.23	-0.00	+0.79	+0.22
Claire Hoffman	Washington	Pac-12	OH	112	**+1.04**	+0.13	+0.23	-0.00	+0.65	+0.02
Julia Bergmann	Georgia Tech	ACC	OH	86	**+1.03**	+0.09	+0.25	-0.01	+0.64	+0.06
Kendall Kipp	Stanford	Pac-12	OPP	117	**+1.02**	+0.03	-0.02	-0.00	+0.72	+0.29
Amber Igiede	Hawaii	Big West	MB	102	**+0.98**	+0.07	+0.04	+0.01	+0.47	+0.38
Elizabeth Juhnke	South Dakota	Summit	OH	113	**+0.96**	+0.01	-0.01	-0.00	+0.69	+0.26
Madi Kubik	Nebraska	Big Ten	OH	109	**+0.94**	+0.05	+0.42	-0.01	+0.44	+0.05
Asjia Oneal	Texas	Big 12	MB	87	**+0.93**	+0.05	+0.04	+0.00	+0.35	+0.50
Logan Eggleston	Texas	Big 12	OH	91	**+0.89**	+0.09	+0.05	+0.01	+0.70	+0.05

Project #2: Advanced predictive metrics in baseball

Joint work with Vicente Iglesias

- Pitch outcome modeling is very useful, but the problem is that the most important variables are the least reliable!



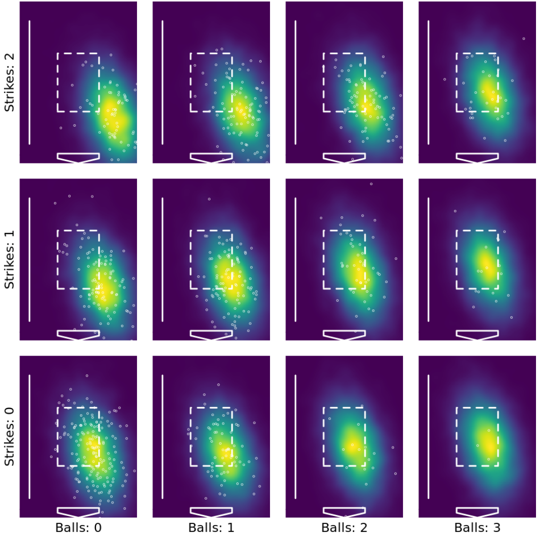


Project #2: Advanced predictive metrics in baseball

Joint work with Vicente Iglesias

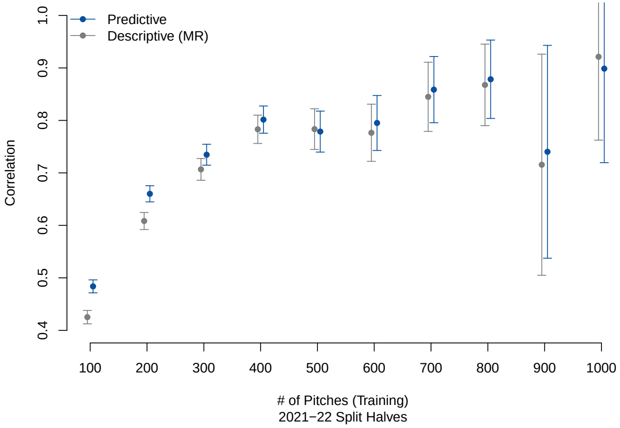
We used a Bayesian hierarchical model to estimate the probability distribution over the 9-dimensional pitch trajectory for each pitcher in each count.

We predicted future pitcher outcomes using this model.



Project #2: Advanced predictive metrics in baseball

Joint work with Vicente Iglesias



- Our **predictive** model outperforms mean-regressed pitch outcome model for pitchers with < 300 pitches.

Project #3: Advanced predictive metrics in soccer

Joint work with Andrew Kang

Motivation: Parma-led proposal for 2024 Opta Forum

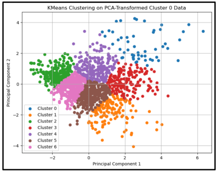
> *Using a combination of Opta Vision events, and spatial tracking data captured for all on-field players, propose a method for categorising players, based on role-specific performance traits, to group similar stylistic players to help enhance a recruitment profiling pipeline.*

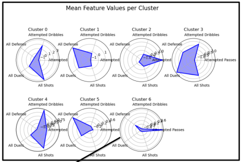
- BUT we had only 100 games of (anonymized) data
- We clustered players and fit cluster-specific box plus-minus
    - This identifies players well-suited for specific game models
- Andrew presented this work at Opta Forum!

Project #3: Advanced predictive metrics in soccer

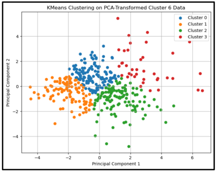
Joint work with Andrew Kang

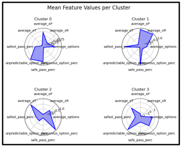
- Step 1: Cluster players based on *style*, not performance





Step 1: PCA and Initial Clusters



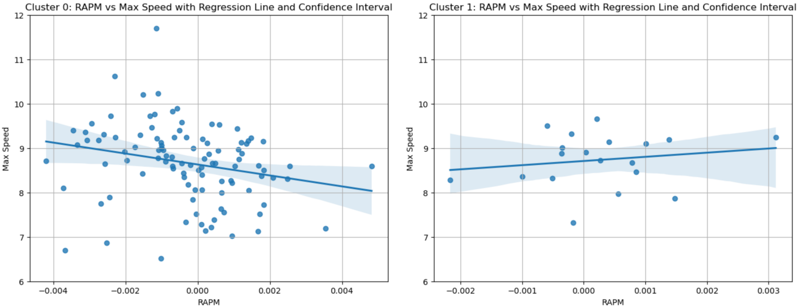


Step 2: PCA and Passers' Passing Sub-clusters

Project #3: Advanced predictive metrics in soccer

Joint work with Andrew Kang

- Step 2: Evaluate player performance relative to their style



A more refined equation that uses the cluster as an indicator variable is shown below.

$$\text{RAPM} = \beta_0 + \sum_{k=1}^{p} \beta_k b_k + \sum_{i=0}^{n-1} \sum_{k=1}^{p} \gamma_{ik} I(\text{Cluster} = i) b_k$$

Outline

1. Player evaluation based on game performance
    - Descriptive metrics
    - Predictive metrics
2. Identifying undervalued players in the minor leagues w/ data
    - Targeted development plans
3. My Research
    - Basic descriptive metrics in volleyball
    - Advanced predictive metrics in baseball
    - Advanced descriptive metrics in soccer

Thank You!

References

Decroos, T., Bransen, L., Van Haaren, J., & Davis, J. (2019). Actions speak louder than goals: Valuing player actions in soccer. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 1851-1861).

Jacobs, J. (2017). Deep Dive on Regularized Adjusted Plus-Minus I: Introductory Example. https://squared2020.com/2017/09/18/deep-dive-on-regularized-adjusted-plus-minus-i-introductory-example/

Matano, F., Richardson, L., Pospisil, T., Politsch, C. A., & Qin, J. (2023). Augmenting adjusted plus-minus in soccer with FIFA ratings. Journal of Quantitative Analysis in Sports, 19(1), 43-49.

Myers, D. (2020). About Box Plus/Minus (BPM). https://www.basketball-reference.com/about/bpm2.html

Singh, K. (2018). Introducing Expected Threat (xT). https://karun.in/blog/expected-threat.html