

Quantifying Pitch Control

William Spearman

#OptaProForum

Introduction

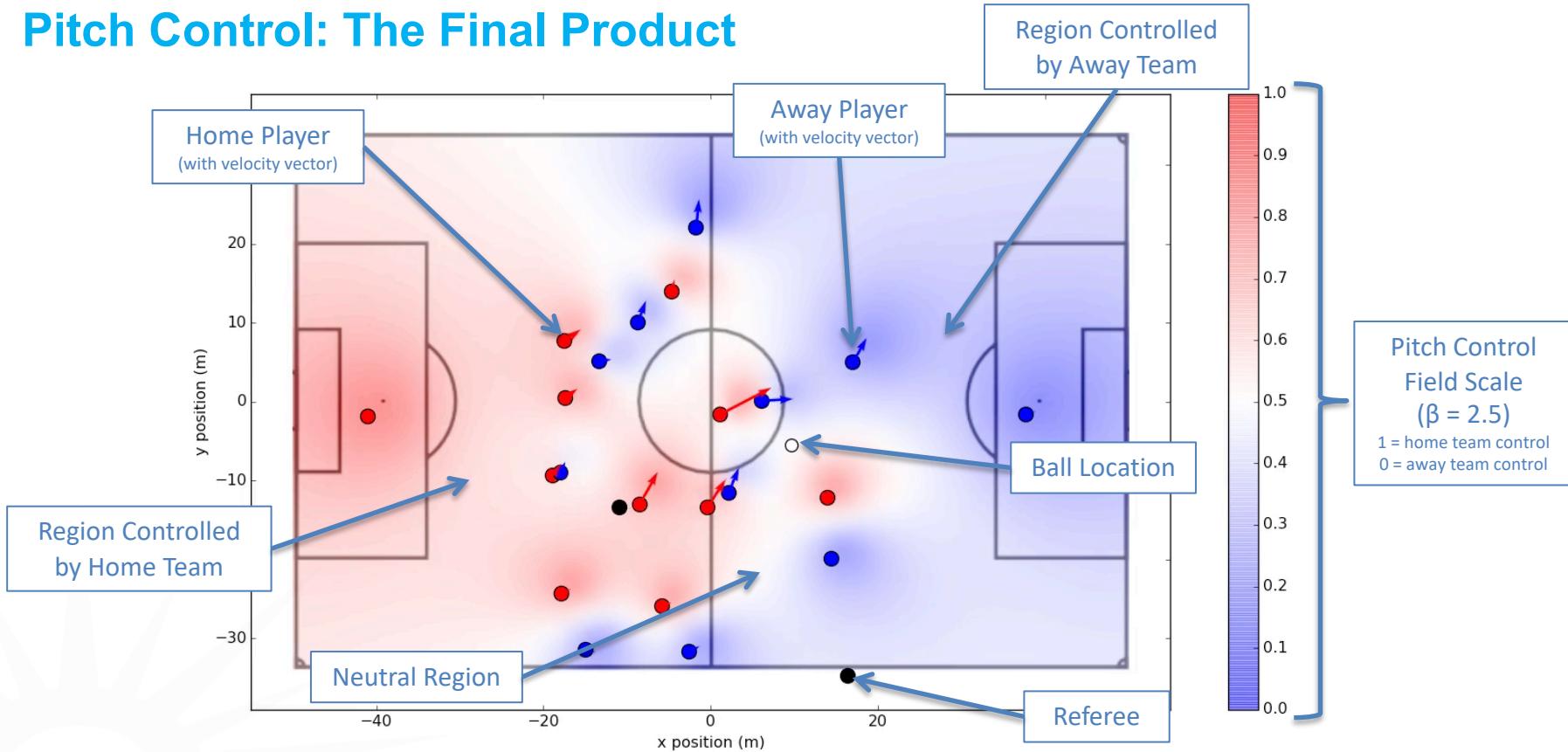
Dr. William Spearman

- Ph.D. in High Energy Particle Physics from Harvard University
 - Studied the Higgs Boson
- Works as a data scientist for Hudl
 - Not a football expert!

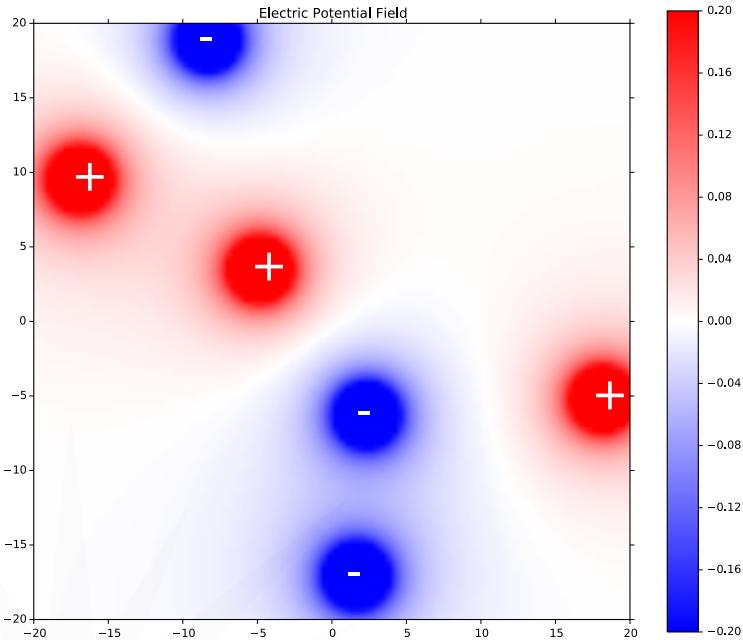


Hudl + Replay Analysis + Sportstec

Pitch Control: The Final Product



Motivation



Physics

The **electric potential field** quantifies the way charged **particles** exert a force on a test charge in **space**.

Football

We propose the development of a **pitch control field** that can be used to quantify the way football **players** control regions on the **pitch**.

What is Pitch Control?

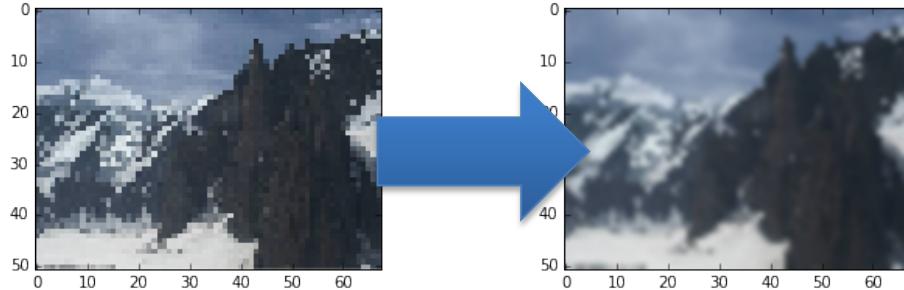
Definition

We define *pitch control field (PCF)* for location, x , as the probability that the home team will end up with possession of the ball if it were at location, x . The PCF predicts the “next possessor”.

Get there first with the most men

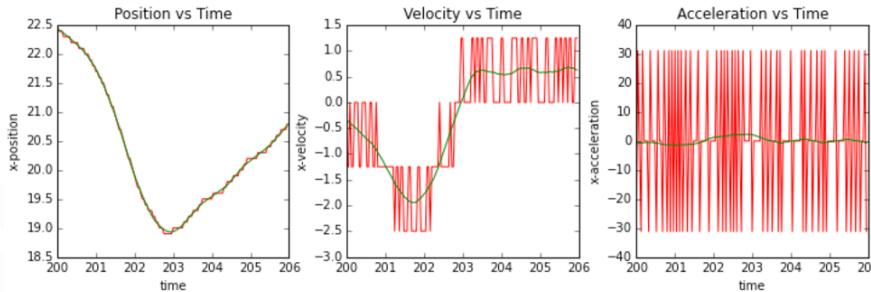
The PCF will probably depend on **the time** it takes each player to reach location, x and this time will depend on the location and velocity of each player.

Using Tracking Data



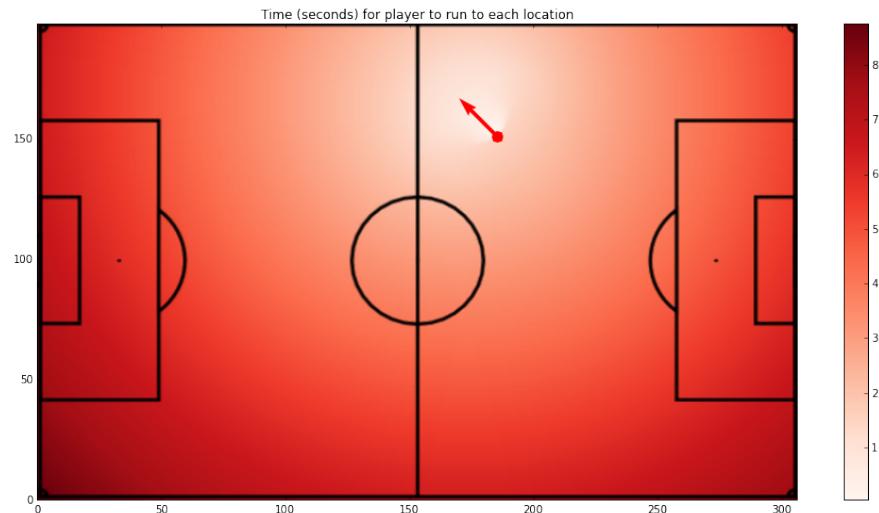
TRACAB Data

- Player/ball positions at 25 fps
- We smooth it with an S-G filter



Calculating Times Using

- Player position
- Player velocity
- Player acceleration
- Maximum player speed



Choosing a Model

Label for the i th player (1 for Home team and -1 for away team)

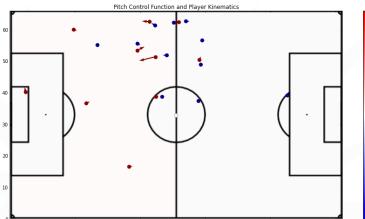
β is a parameter which indicates how much to weight being the first to the ball (Range = 0 to ∞)

$$PCF(t_i, l_i) = \left[\frac{\sum_i l_i t_i^{-\beta}}{\sum_i t_i^{-\beta}} + 1 \right] / 2$$

Time for the i th player to reach the ball

This part will be between -1 and 1

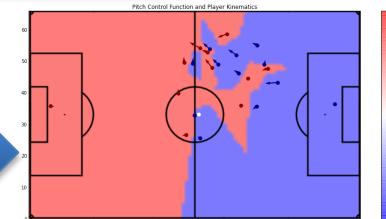
Use a naïve linear scaling to change output from -1 to 1 \rightarrow 0 to 1.



Understanding the Parameter, β

$\beta = 0$: PCF always 0.5

$\beta = \infty$: PCF is 1 if closest player is on the home team, otherwise, it's 0.



Fitting Strategy

1. Sync Opta and Tracab

- Identify when the ball is “in-play”

2. Calculate ball’s possessor

- Contested – both teams are near it
- None – no team is near the ball
- Home/Away – one team has uncontested control of the ball

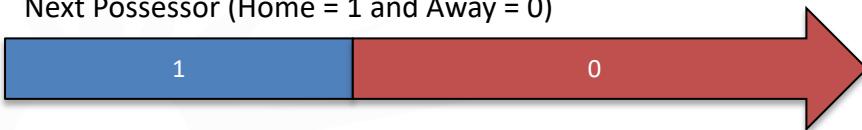
Current Possessor



3. Calculate ball’s *next* possessor

- This is done by looking forward in time to see which team gets the next uncontested possession.

Next Possessor (Home = 1 and Away = 0)



4. Choose Fit Frames

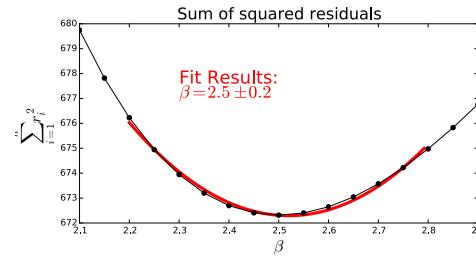
- We focus on frames where there *possession* is *contested* or *none*.
- We want there to be a clear *next possessor*.
- The next possessor is *truth* value for who gets the ball next

5. Calculate PCF at ball’s location.

- This gives us the model’s *prediction* for who gets the ball next.

6. Minimize sum of squared errors

- Calculate residual for each frame:
 - $r = x_{pcf} - x_{real}$
- Sum of the squared residuals for each frame, i :
 - $\Sigma r^2 = \sum_i (x_{pcf,i} - x_{real,i})^2$.



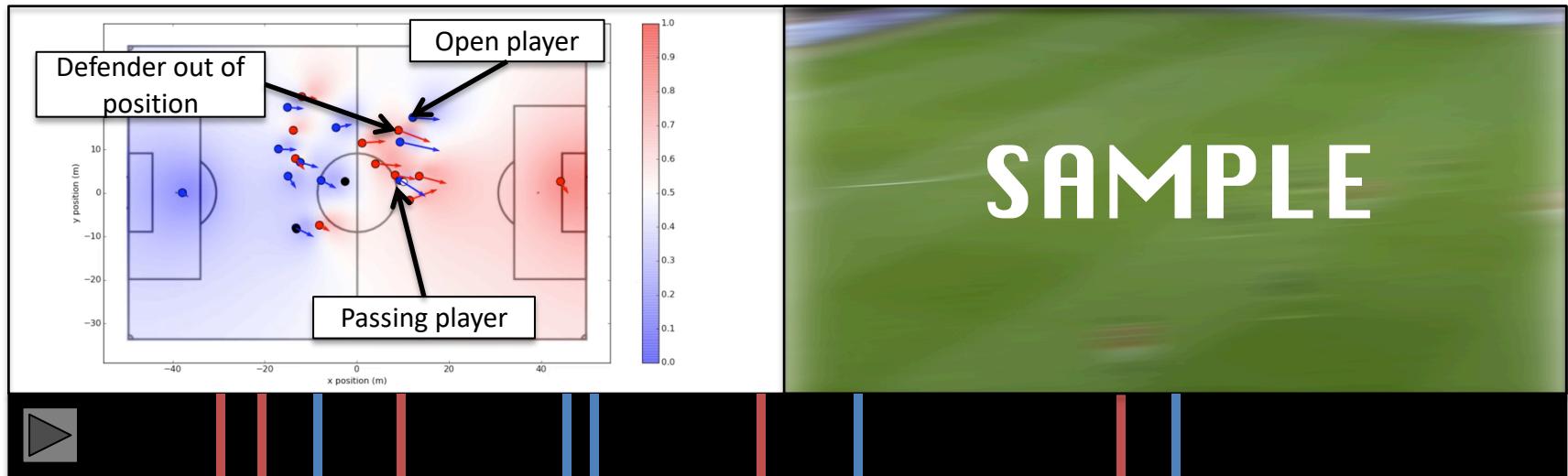
Applications: A New Way To Watch Film

Watch Alongside Film

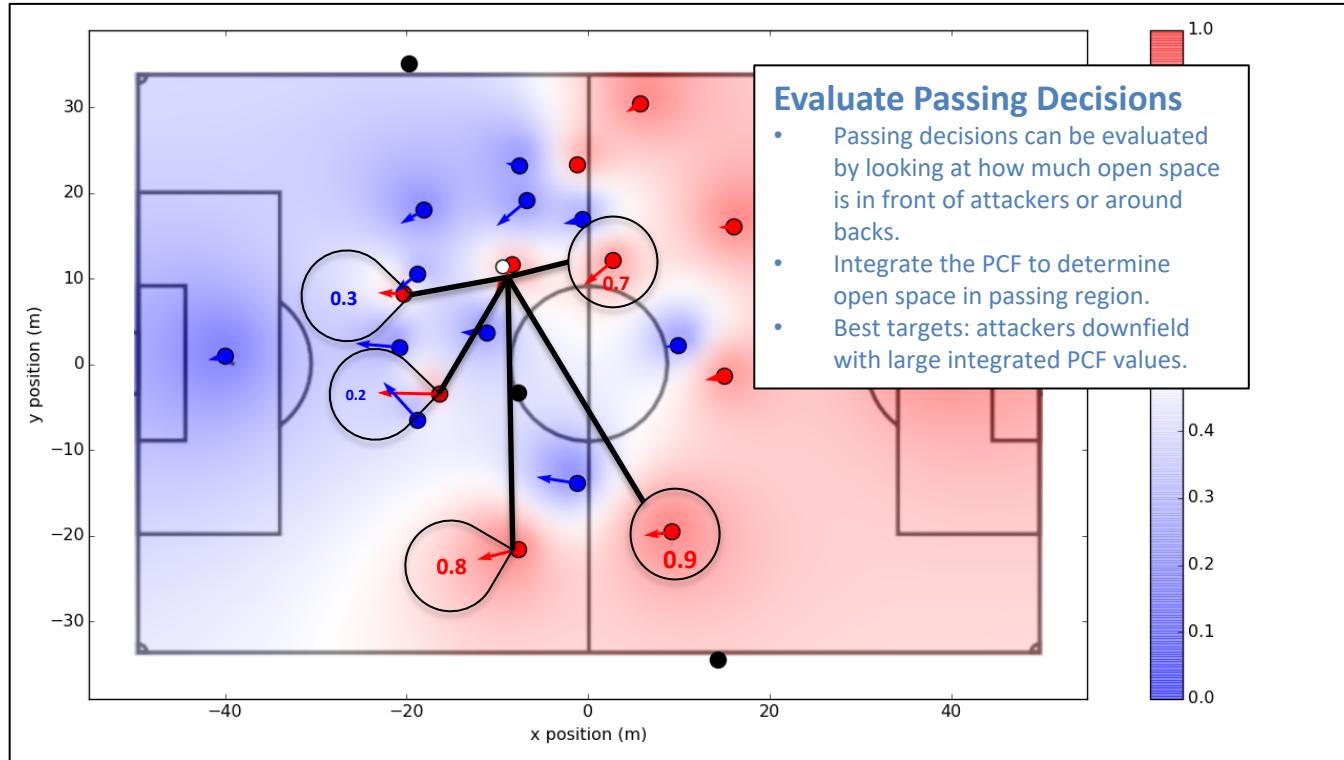
- Watch animated PCF along with video.
- PCF makes it possible to visualize space in-between units and space behind units.

Identify

- Defensive players who are out of position (even if the mistake doesn't result in a goal)
- Missed offensive opportunities.



Applications: A New Way To Watch Film



Applications: A New Metric for Player Performance

Identify Controlling Players

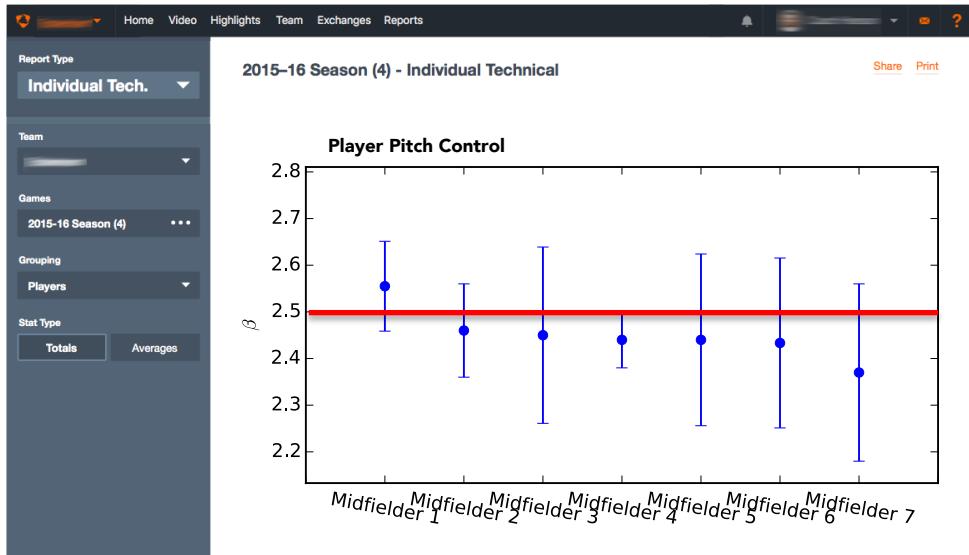
- Certain players are more capable of contesting for the football than others.
 - This should appear if we calculate their “player-specific” beta values.
 - In other words, how much does their presence improve their team’s chance of gaining possession of the ball above the average.

How?

- This is done by fitting β_i separately for each player, i .

$$PCF(t_i, l_i) = \left[\frac{\sum_i l_i t_i^{-\beta_i}}{\sum_i t_i^{-\beta_i}} + 1 \right] / 2$$

- Need more statistics to make inferences about specific player trends.
- β only matters w.r.t. other players. No common sense interpretation.



Results

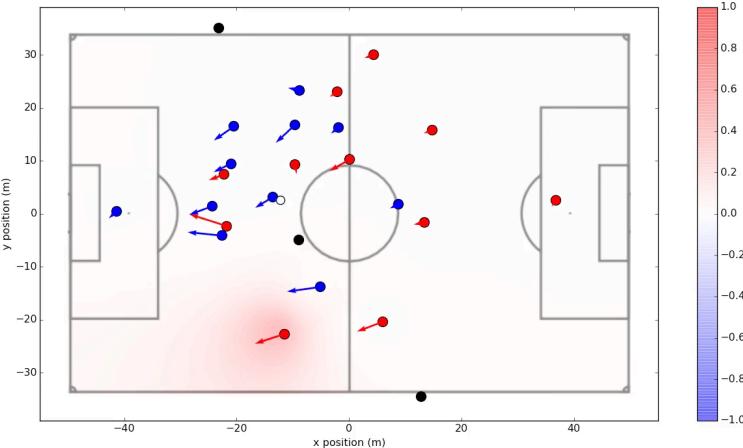
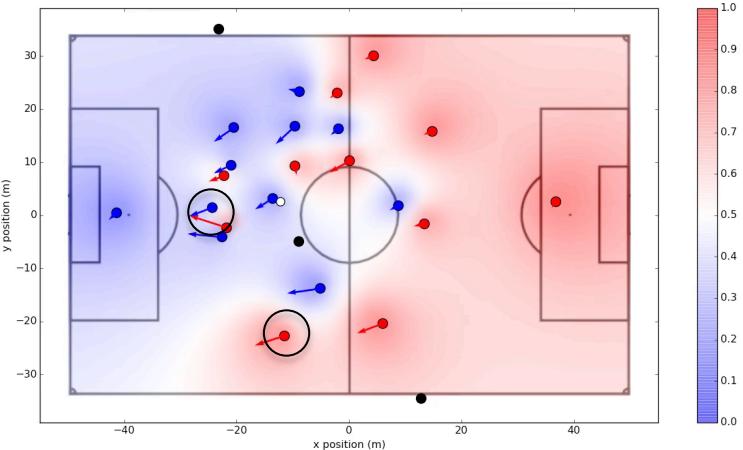
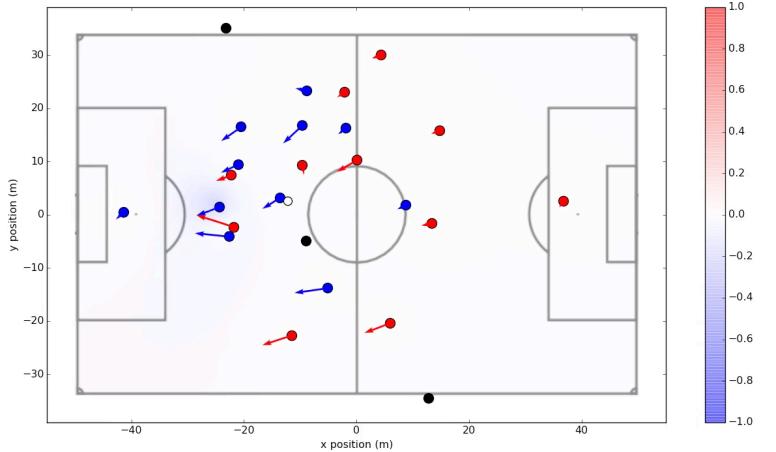
- Above we see the fitted beta value for specific midfielders when compared to the league average.
- Some are more controlling while others exert less impact.
- Error bars show the standard deviation among four games.

Applications: Player Positioning

Quantify the Effect of Player Positioning

- How much impact does a player's position have on his team's control of the pitch?
- In other words, how different would the PCF be if a certain player weren't on the pitch?

$$\Delta PCF(j, t_i, l_i) = \left[\frac{\sum_i l_i t_i^{-\beta_i}}{\sum_i t_i^{-\beta_i}} - \frac{\sum_{i \neq j} l_i t_i^{-\beta_i}}{\sum_{i \neq j} t_i^{-\beta_i}} \right] / 2$$



Thank you

William Spearman

#OptaProForum

Special Thanks:

Armen Badeer
Austin Basye
Nathan DeMaria
Keenan Hawekotte
Sam Lloyd
John McGuigan
Paul Pop
Markus Woodson