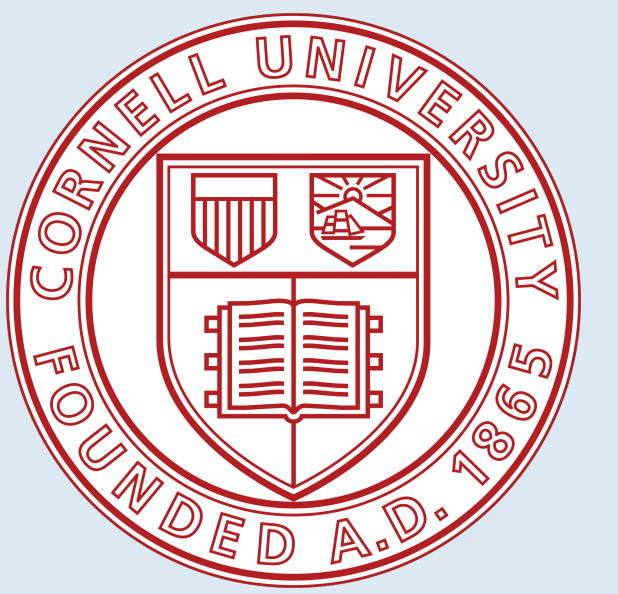# FÚTBOL FORMATION ANALYSIS USING COMPUTER SIMULATIONS

Cornell University College of Engineering
Ithaca, New York
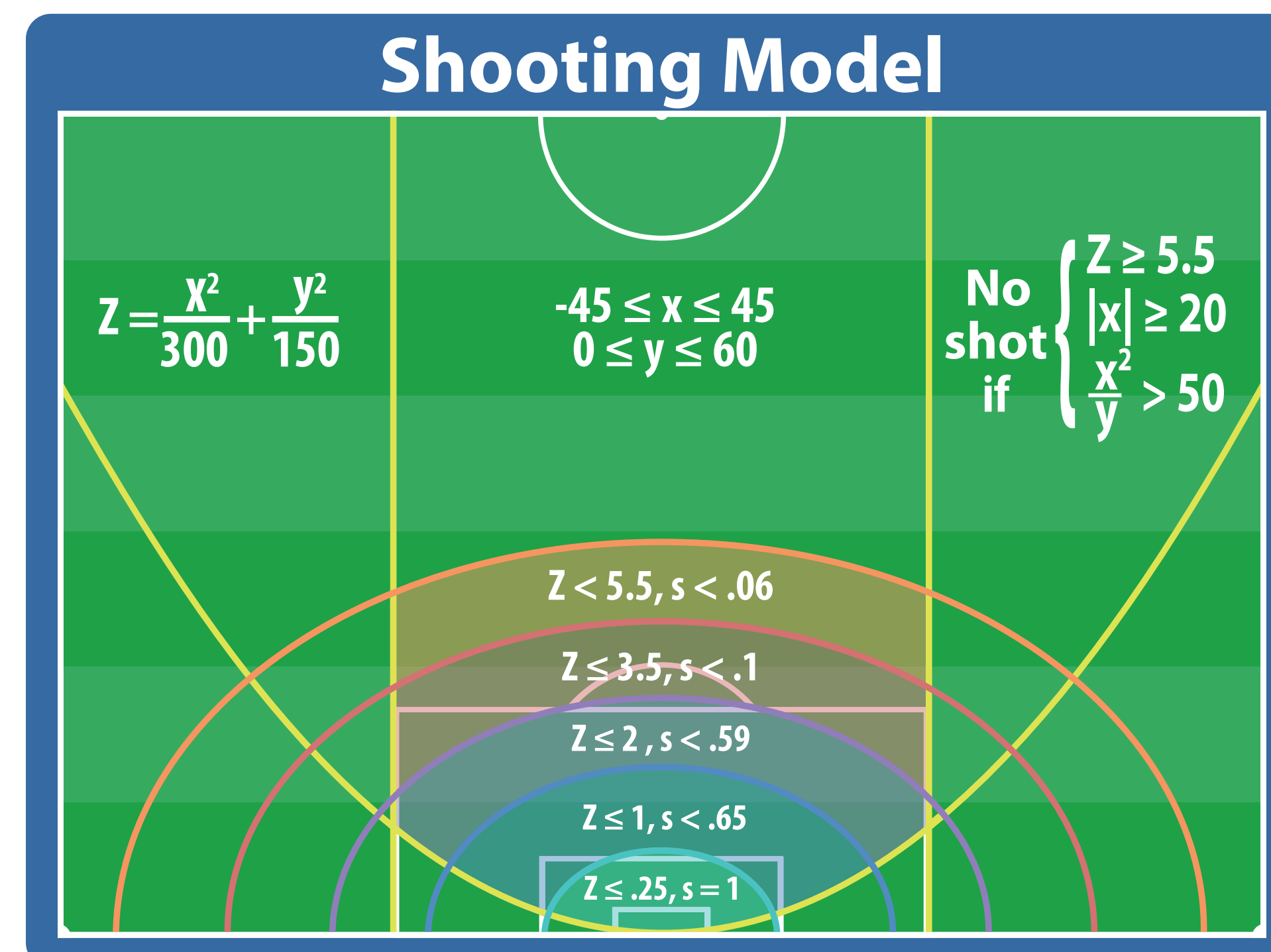Samad Arshad, Daniel Cardenas, Daniel Stabile, Thomas Taffe
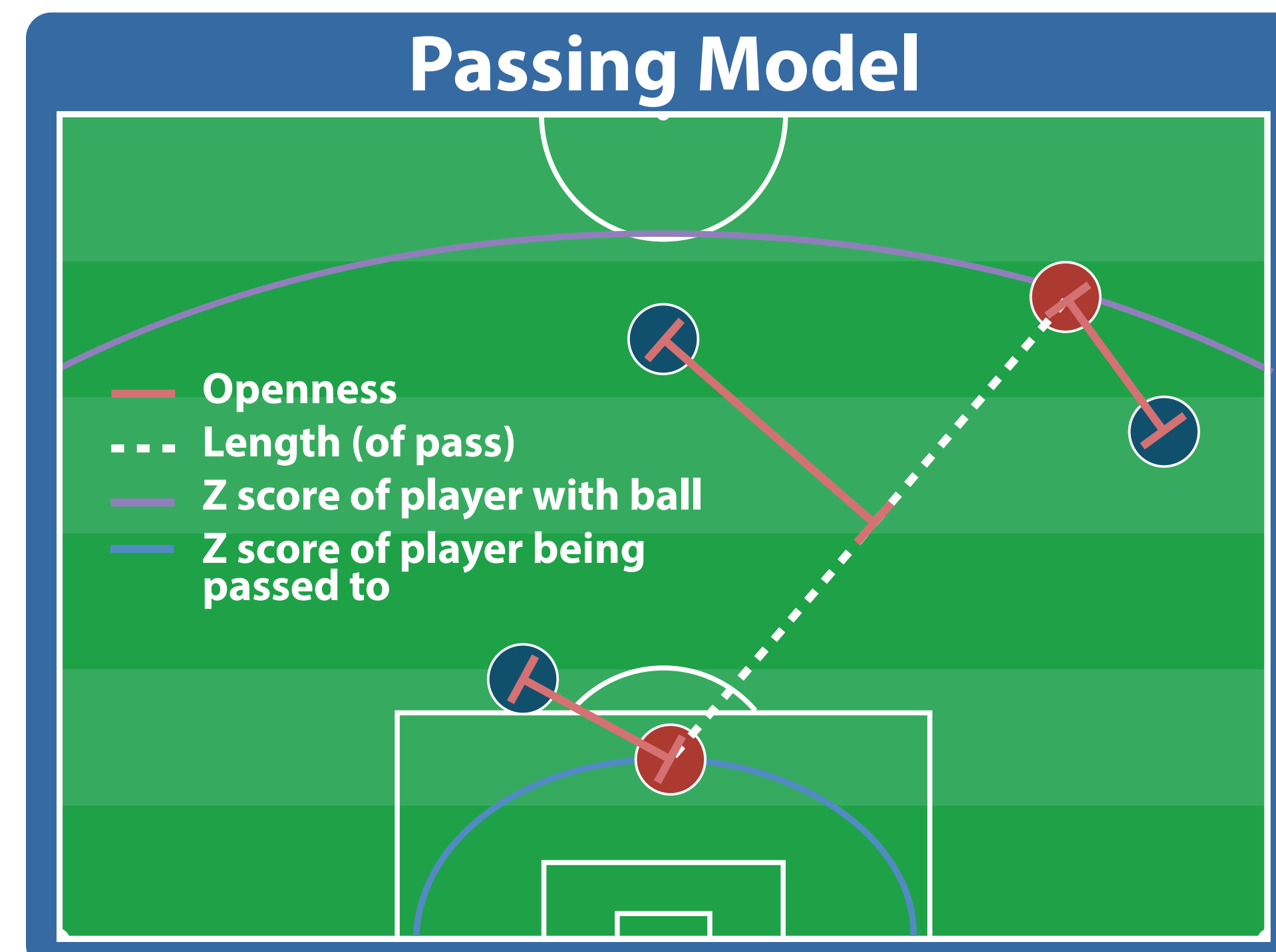
Abstract

One of the most important factors in a fútbol team's performance is the formation used. To determine the best formation, a program was created to simulate offensive pushes on half the field in 24,500 matches. Analyzing these matches revealed that the 3-5-2 offenses made the most goals, while the 4-3-3 offense made the fewest. The analysis also indicated that the strongest defense was the 3-5-2 W, making the most interceptions, while the 4-3-3 made the least. These results are applicable to entire games, as they tend to consist of mostly offensive pushes. Therefore, a formation's performance in entire games will correlate with its performance in offensive pushes. However, the applicability of the results to real fútbol games is limited by the accuracy of the simulations. For instance, a significant drawback is that our program only simulates a 2D plane, as plays in the real world often involve three dimensions. Further work on simulating entire games would allow for finer analysis of formation performance.
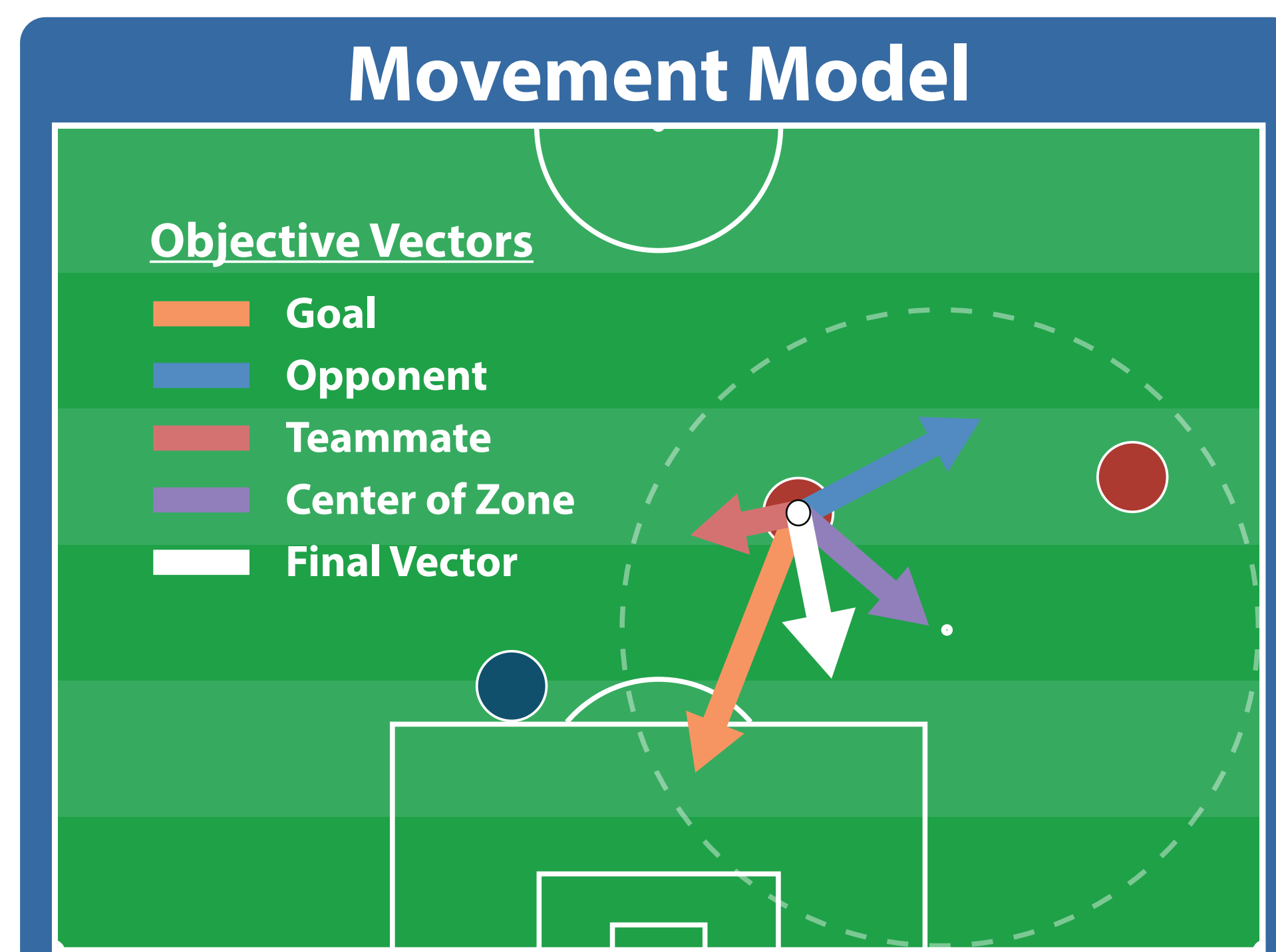
Methods

- Matches were simulated on half a field and ended once the ball was taken by the defenders, a goal was scored, or the ball went out of bounds
- Simulation was created using Python3, PySDL2, and SDL2 (to see our project on GitHub, scan the QR code in the bottom right)

Shooting Model

$$Z = \frac{x^2}{300} + \frac{y^2}{150}$$

$$-45 \leq x \leq 45$$
$$0 \leq y \leq 60$$

No shot if $\begin{cases} Z \geq 5.5 \\ |x| \geq 20 \\ \frac{x^2}{y} > 50 \end{cases}$

$Z < 5.5, s < .06$
$Z \leq 3.5, s < .1$
$Z \leq 2, s < .59$
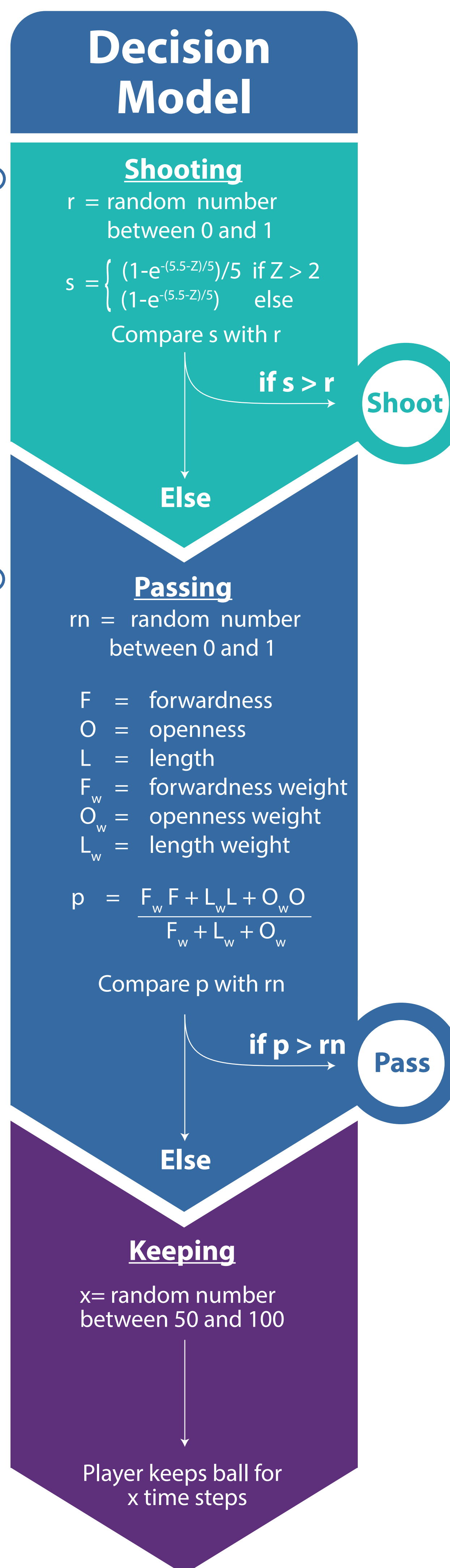$Z \leq 1, s < .65$
$Z \leq .25, s = 1$

- The shooting model is based on the 2D projection of a paraboloid onto the x-y plane of the field.
- The further away a player is from middle of the goal the worse their chances are of scoring.
- Players automatically shoot within 10 yards of the center of goal
- S comes from the exponential CDF where Z is the input

Passing Model

- Openness
- Length (of pass)
- Z score of player with ball
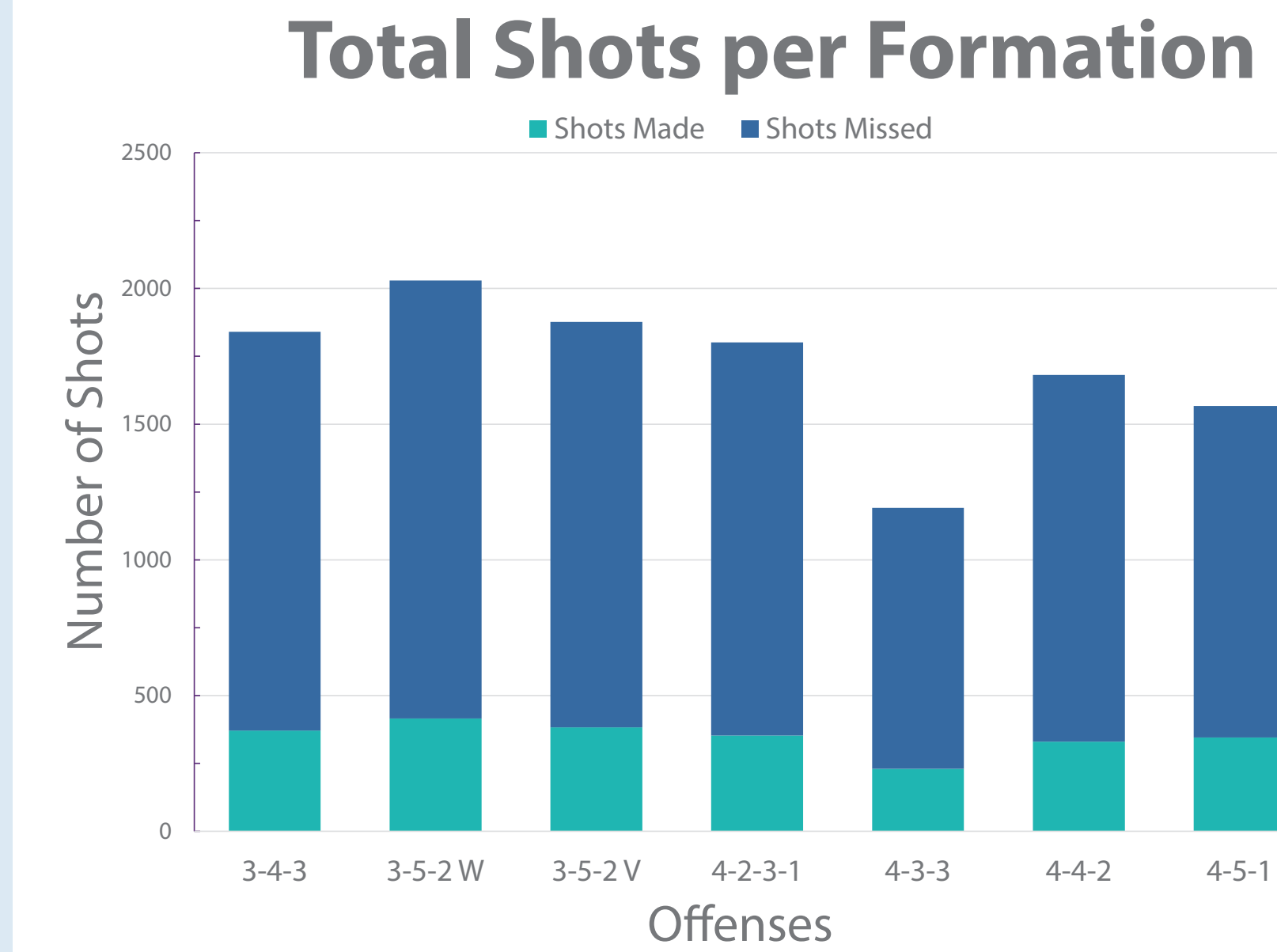- Z score of player being passed to

- If a pass is too short or too long it is less likely to happen.
- Forward passes are prioritized over backward ones
- Forwardness is determined by the players' difference in Z value from the above graph
- Open passes are ideal with the largest factor of passing being the distance a defender is from the line of the pass
- $F_w = 5, O_w = 3, L_w = 2$

Movement Model

Objective Vectors
- Goal
- Opponent
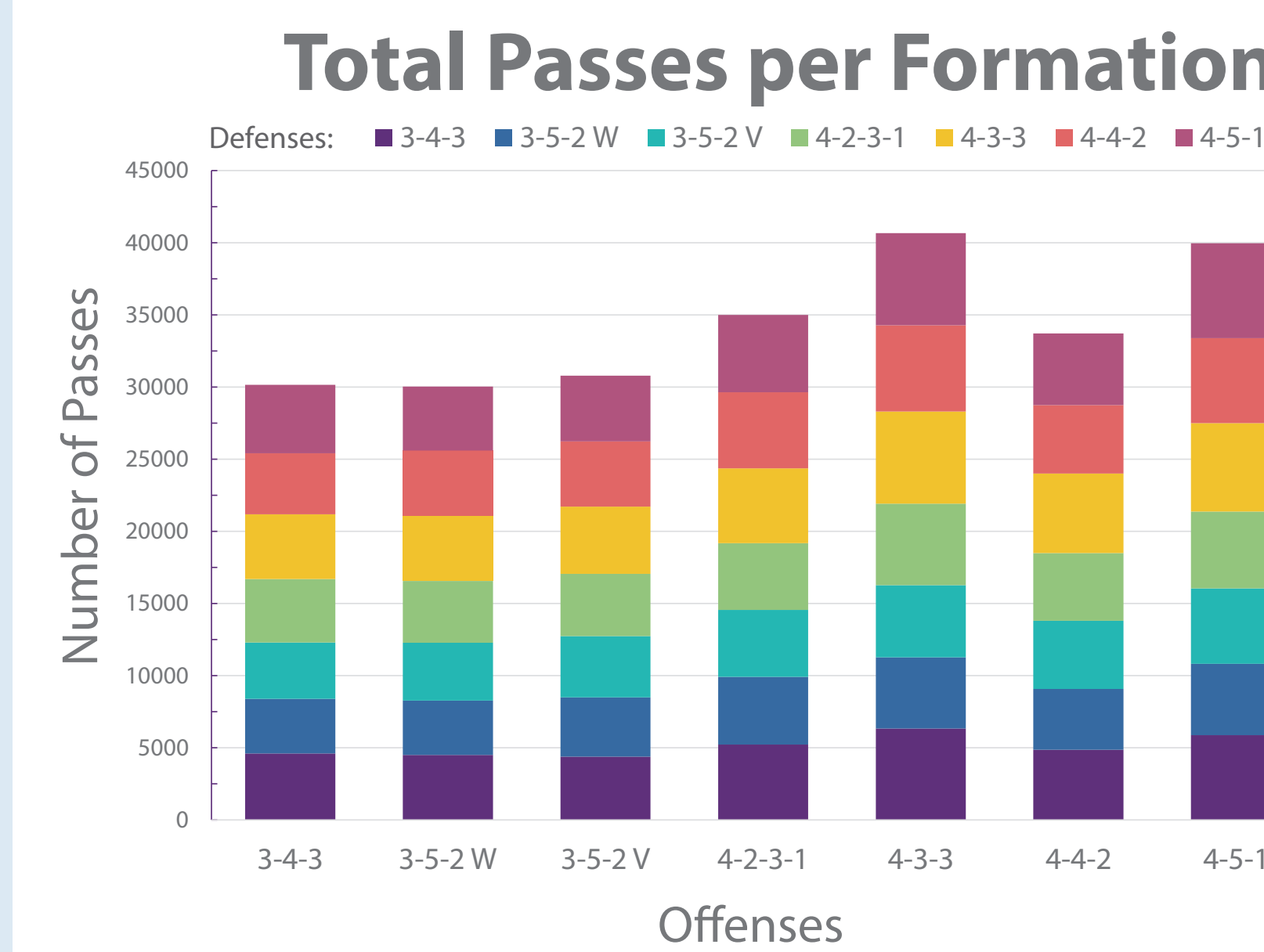- Teammate
- Center of Zone
- Final Vector

- Offensive player movement is slightly more complex than defender movement
- Other than the vectors stated in the graphic offenders have wing and random vectors
- Wing vectors move players on the side 1/3 of the field forward and those on the opposite move inward
- Lastly all players shift with the ball changing the location of their zones
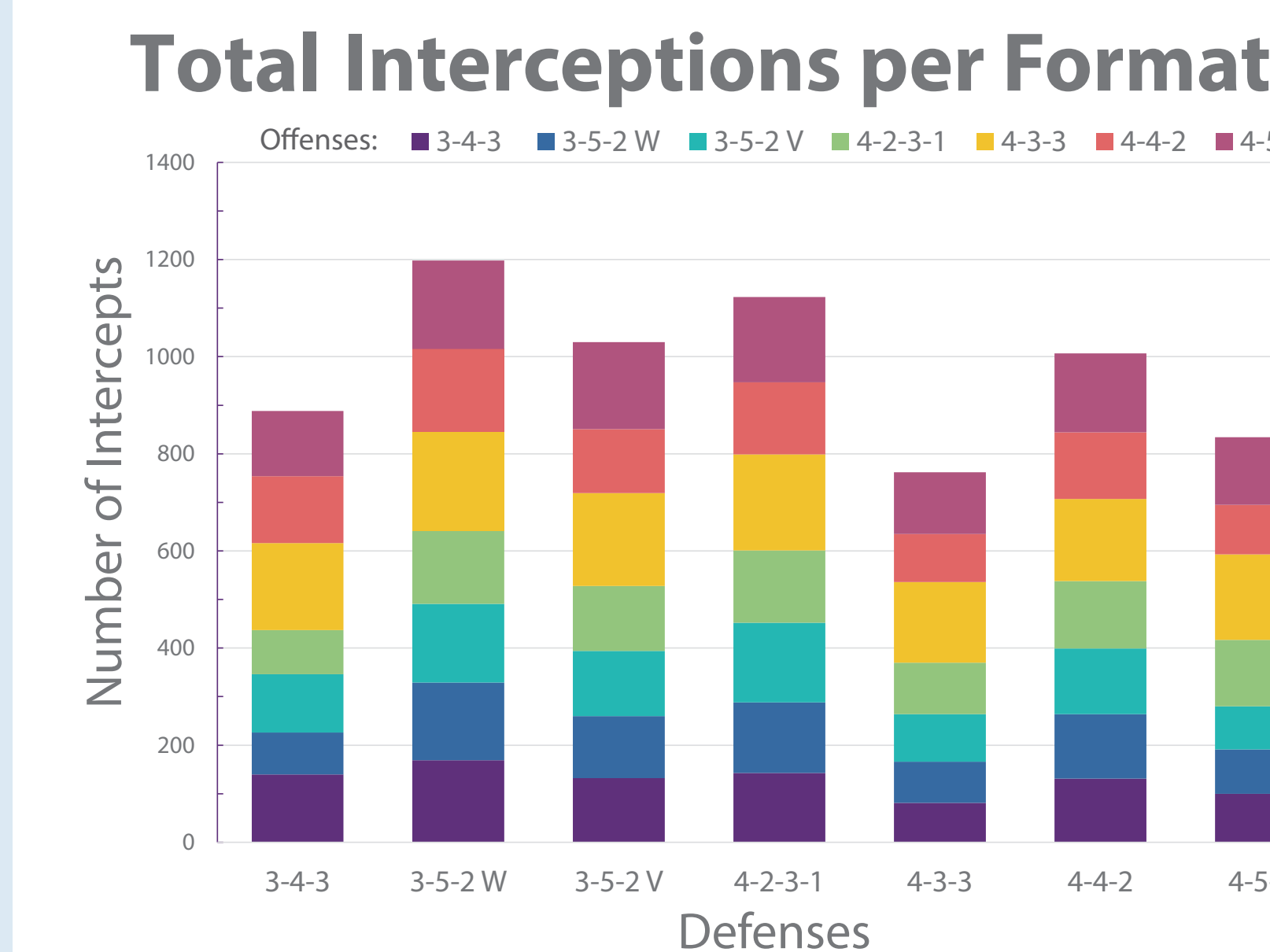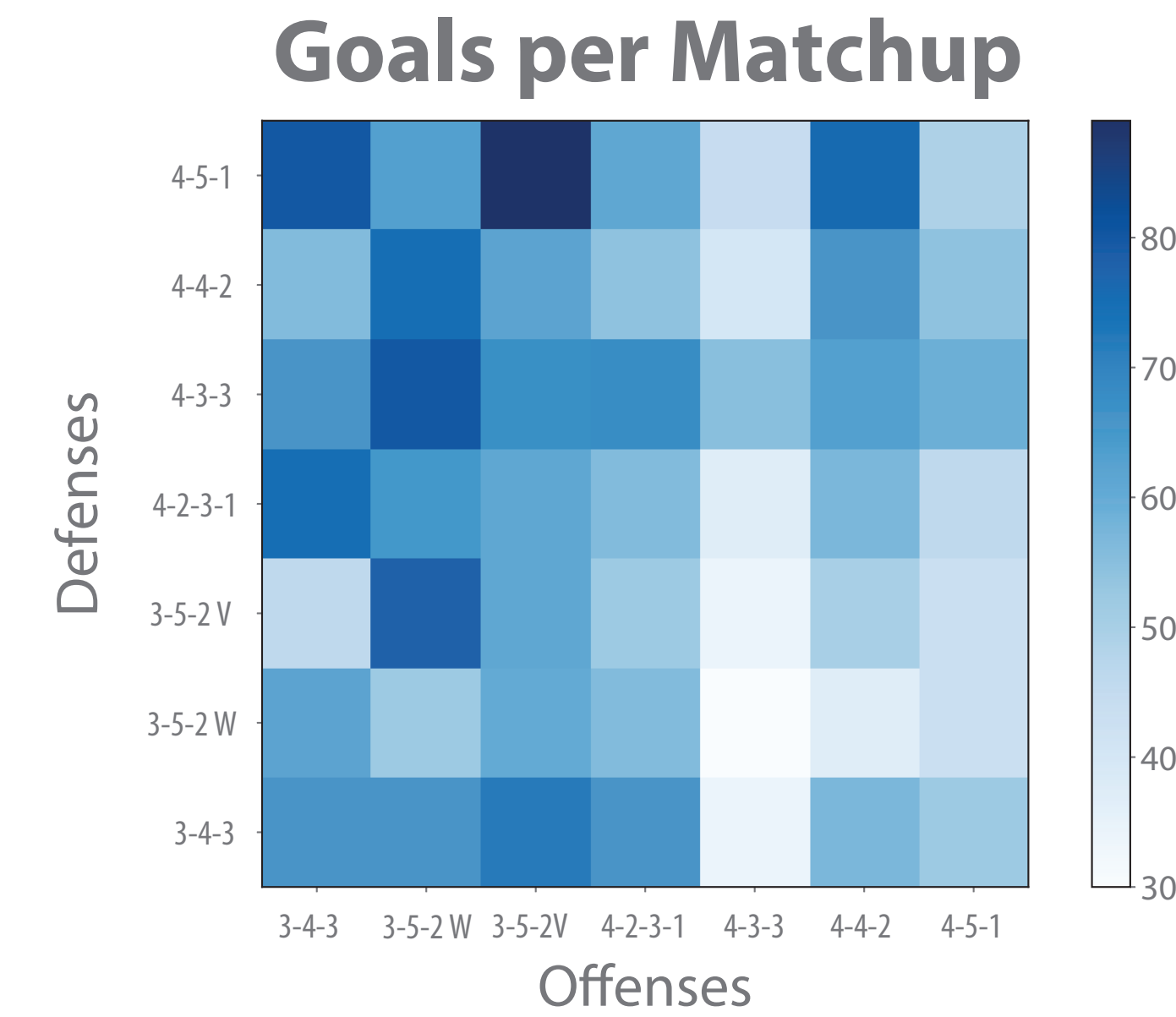
Decision Model

Shooting
r = random number between 0 and 1

$$s = \begin{cases} (1-e^{-(5.5-Z)/5})/5 & \text{if } Z > 2 \\ (1-e^{-(5.5-Z)/5}) & \text{else} \end{cases}$$

Compare s with r

if s > r → **Shoot**

Else

Passing
rn = random number between 0 and 1

F = forwardness
O = openness
L = length
F_w = forwardness weight
O_w = openness weight
L_w = length weight

$$p = \frac{F_w F + L_w L + O_w O}{F_w + L_w + O_w}$$

Compare p with rn

if p > rn → **Pass**

Else

Keeping
x= random number between 50 and 100

Player keeps ball for x time steps

Results

Total Shots per Formation

The figure above shows the number of total shots as either made or missed for each offense versus all defenses.

Goals per Matchup

The figure above is a heatmap of goals scored by each offense against each defense over the 500 simulations.

Rank	Offense	Defense
1	3-5-2 W	3-5-2 W
2	3-5-2 V	3-5-2 V
3	3-4-3	4-2-3-1
4	4-2-3-1	4-4-2
5	4-4-2	3-4-3
6	4-5-1	4-3-3
7	4-3-3	4-5-1

The chart above ranks offenses from 1-7 by goals scored and defenses from 1-7 by least goals conceded.

Total Passes per Formation
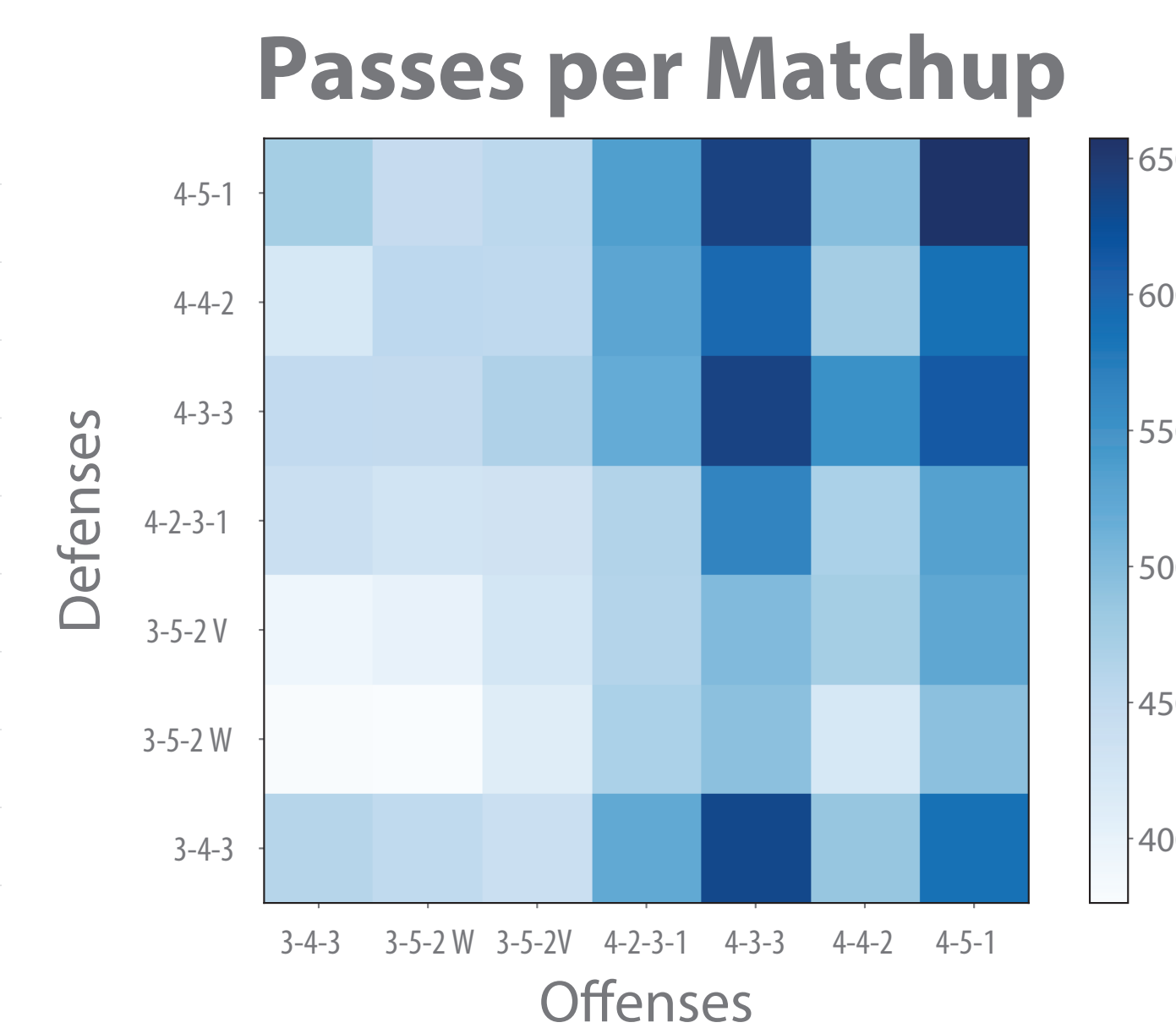

The figure above shows the number of passes made out of 3500 simulations (500 per matchup) for each offense versus all defenses.

Passes per Matchup

The figure above is a heatmap of passes made by each offense against each defense over the 500 simulations.

Rank	Offense	Defense
1	4-3-3	3-5-2 W
2	4-5-1	3-5-2 V
3	4-2-3-1	4-2-3-1
4	4-4-2	4-4-2
5	3-5-2 V	3-4-3
6	3-4-3	4-3-3
7	3-5-2 W	4-5-1

The chart above ranks offenses from 1-7 by passes made and defenses from 1-7 by least passes conceded.

Total Interceptions per Formation
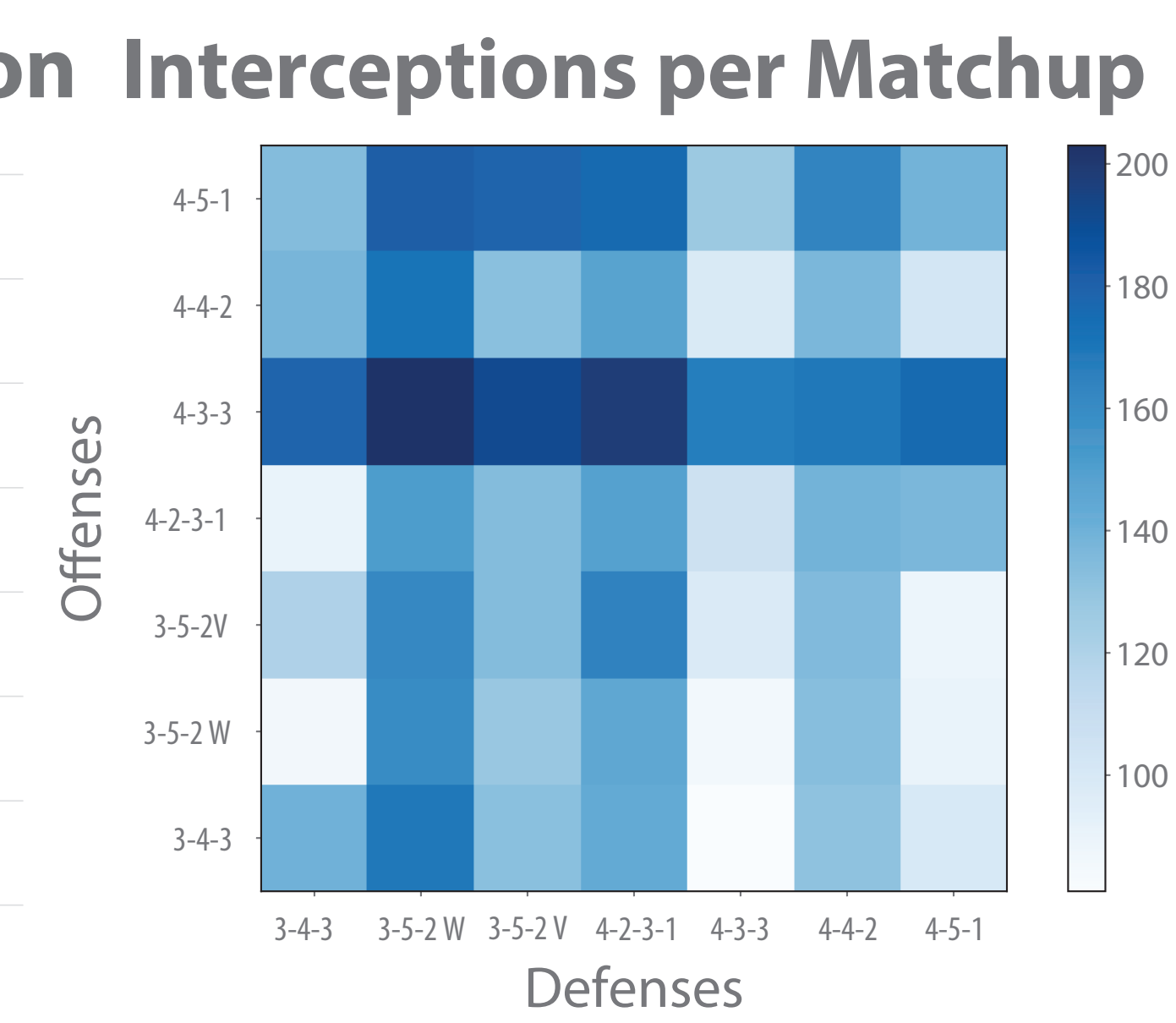

The figure above shows the number of interceptions made out of 3500 simulations (500 per matchup) for each defense versus all offenses.

Interceptions per Matchup

The figure above is a heatmap of interceptions by each defense against each offense over the 500 simulations.

Rank	Offense	Defense
1	3-5-2 W	3-5-2 W
2	4-2-3-1	4-2-3-1
3	3-4-3	3-5-2 V
4	3-5-2 V	4-4-2
5	4-4-2	3-4-3
6	4-5-1	4-5-1
7	4-3-3	4-3-3

This chart ranks defenses from 1-7 by interceptions made and offenses from 1-7 by least interceptions conceded

Conclusion

- Passes and goals are inversely proportional
- 3-5-2 W is the best overall and in almost all statistical categories is followed closely by its V variant in overall stats
- Comparisons between the 3-5-2 W and 4-2-3-1 vs 3-5-2 V and 4-5-1 shows that arranging 5 midfielders in a W performs better than a V
- Formations with central forwards (odd number of forwards) got the least interceptions
- 4-2-3-1 underscored compared to its otherwise high performance in other statistics and the 2018 World Cup
- Midfield 4 defenses perform very similarly across all statistics
- 4-3-3 and 4-5-1 were the worst by a large margin, as in the 2018 World Cup
- Defenses that forced the offenders to keep the ball also intercepted the most passes

Acknowledgements

References

Lauren, M. K., Quarrie, K. L., & Galligan, D. P. (2013). Insights from the Application of an Agent-Based Computer Simulation as a Coaching Tool for Top-Level Rugby Union. International Journal of Sports Science & Coaching, 8(3), 493-504.

Soccer (FIFA) Field Dimensions & Layout. (n.d.). Retrieved July 5, 2018, from https://www.sportscourtdimensions.com/soccer/doi:10.1260/1747-9541.8.3.493

GitHub Project Files