university of groningen

Universitair Medisch Centrum Groningen

Predicting Match Outcome Based on Tactical Performance in Soccer

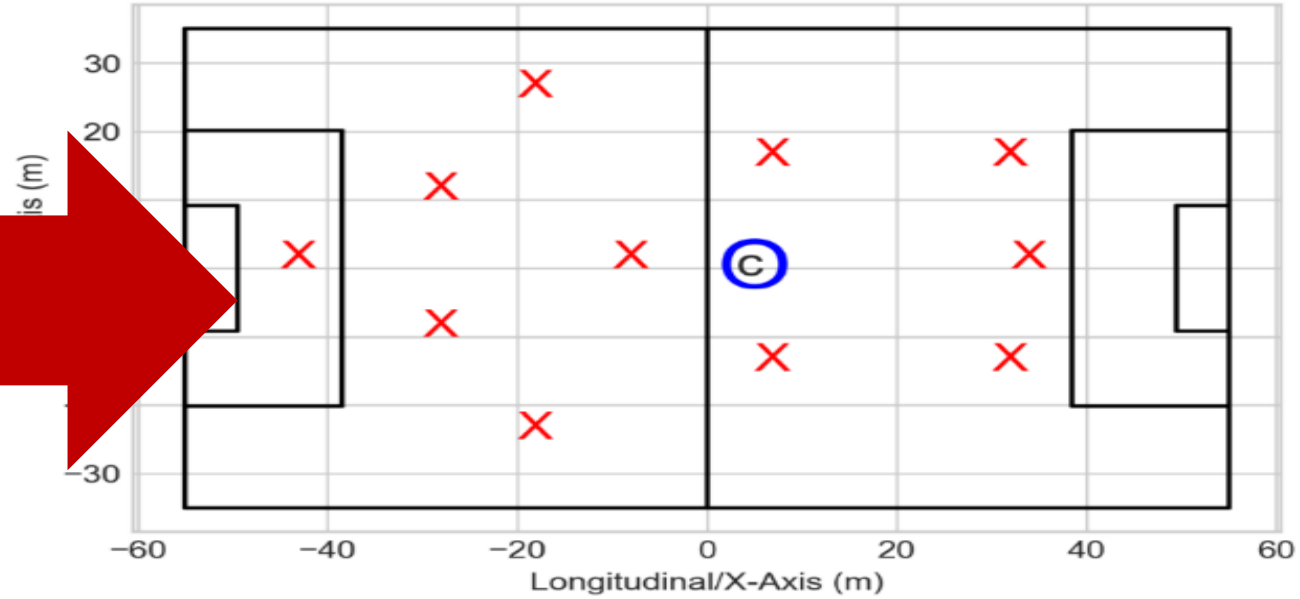
Floris Goes MSc, Dr. Matthias Kempe & Prof. Koen Lemmink

University of Groningen, UMCG, Center for Human Movement Sciences (CHMS)

f.r.goes@umcg.nl | @Florisgoes

Sport Science Institute Groningen

League-wide collection of Tracking-Data



Sports Analytics Challenges in Soccer

- Lack of Contextual Information
- Scientific/Theoretical Rationale
- Link with Performance
  - <30% research has link with performance
- Popularity vs. Relevance

Goes et al., (under review)

Sport Science Institute Groningen

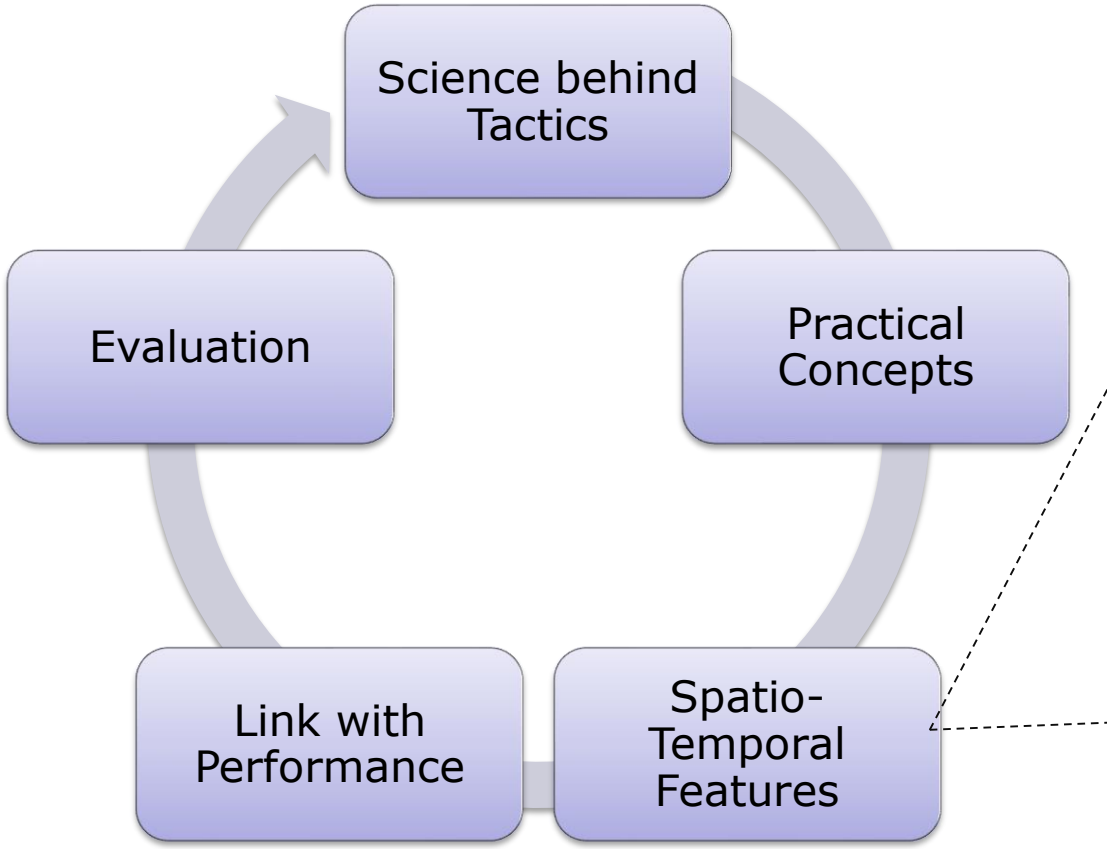
Scope of our work → Translate position tracking data into evidence based feedback on *Tactical* Performance for coaches & analysts

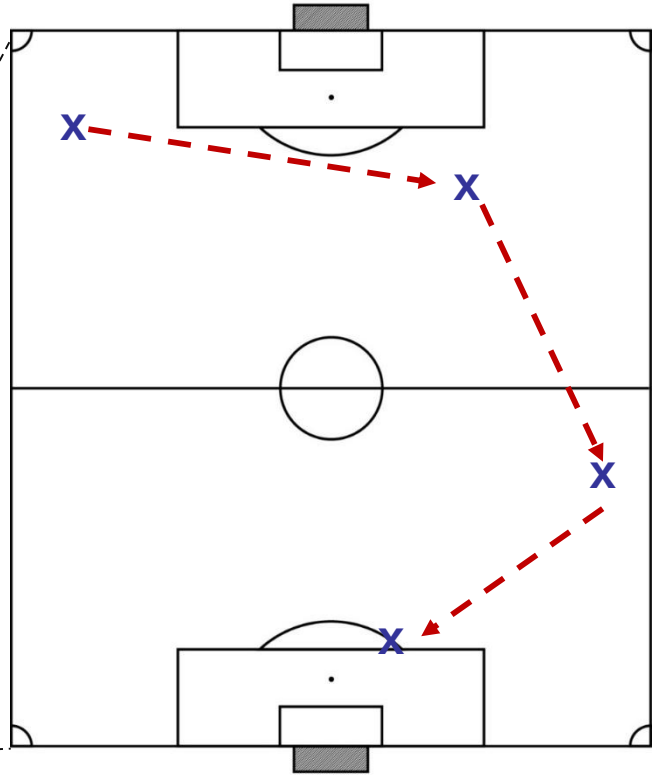
@Florisgoes

@kempe_matthias

Sport Science Institute Groningen

Constructing Features That Capture Tactical Behaviour

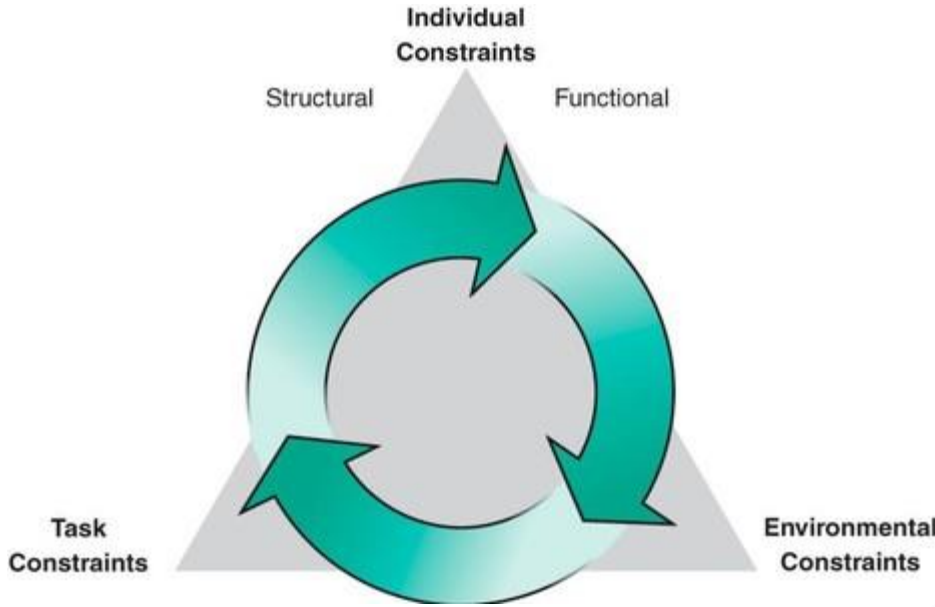




Scientific Perspective on Tactical Behaviour

- Teams are complex dynamical systems
- Players coordinate their behaviour in space and time
- Interaction with the opponent and the environment

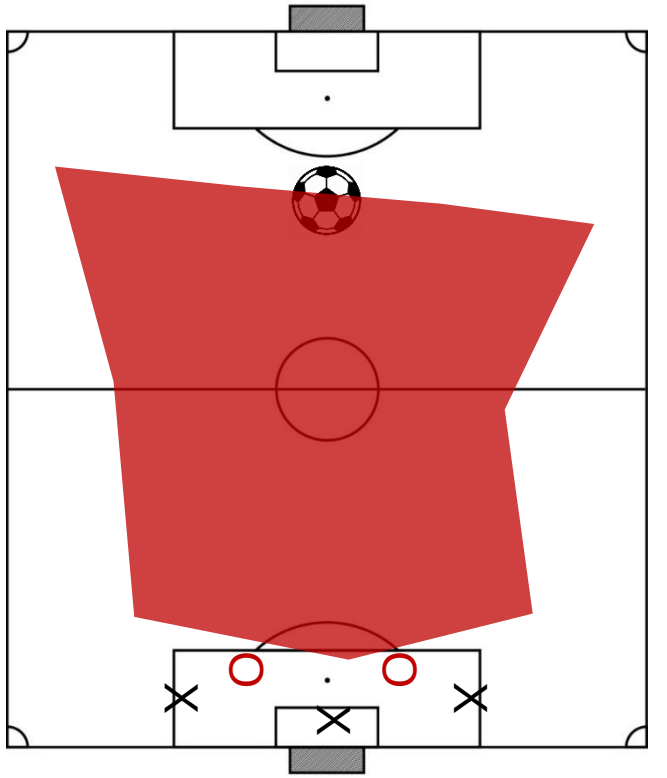
NEWELL'S THEORY OF CONSTRAINTS



Gréhaigne et al., 1997, 1999, Davids et al., 2005; Balague et al., 2013.

Practical Perspective on Tactical Behaviour

- **Principles of (Attacking) Play:**
  - Penetrate & Destabilize the defense
  - Move the ball to the goal
  - Create numerical superiority
  - Keep the field long and wide



Costa et al., 2009; Clemente et al 2014

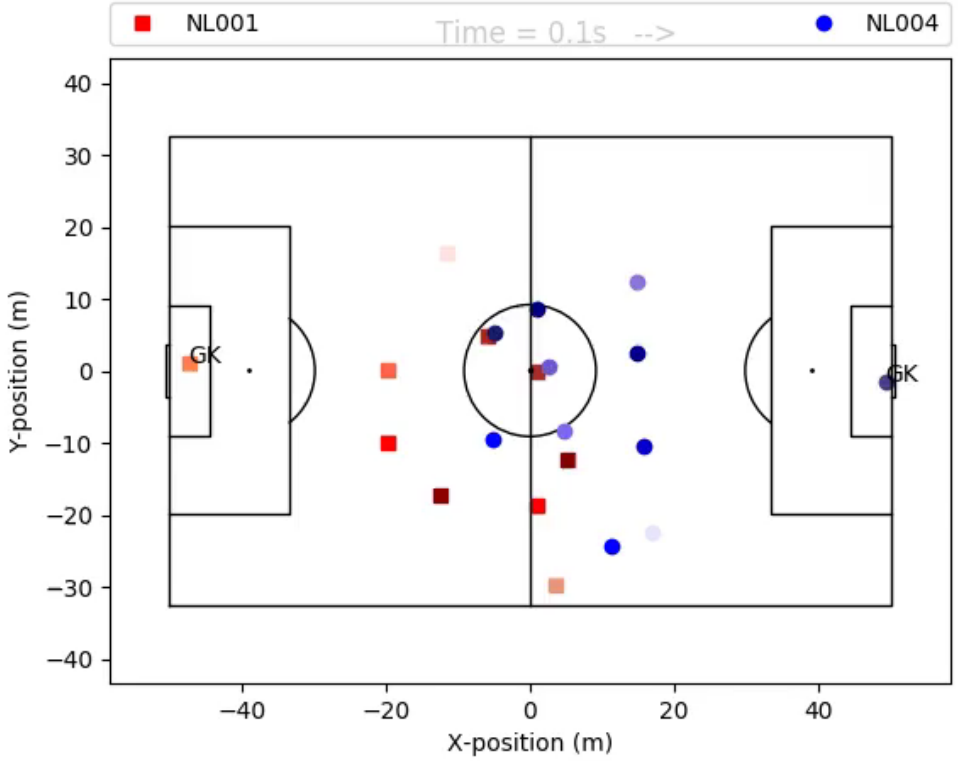
Aim → Construct tactical features inspired on existing metrics & principles of attacking play and study the link with match outcome.

Core contribution → Findings should help identify what features are relevant to provide to coaches, and should provide a basis for new developments.

METHODS

Data & Design

- Position Tracking Data
- 118 Prof. Dutch Premier League Matches
- 26 unique teams vs 2 opponents during the last 4 seasons

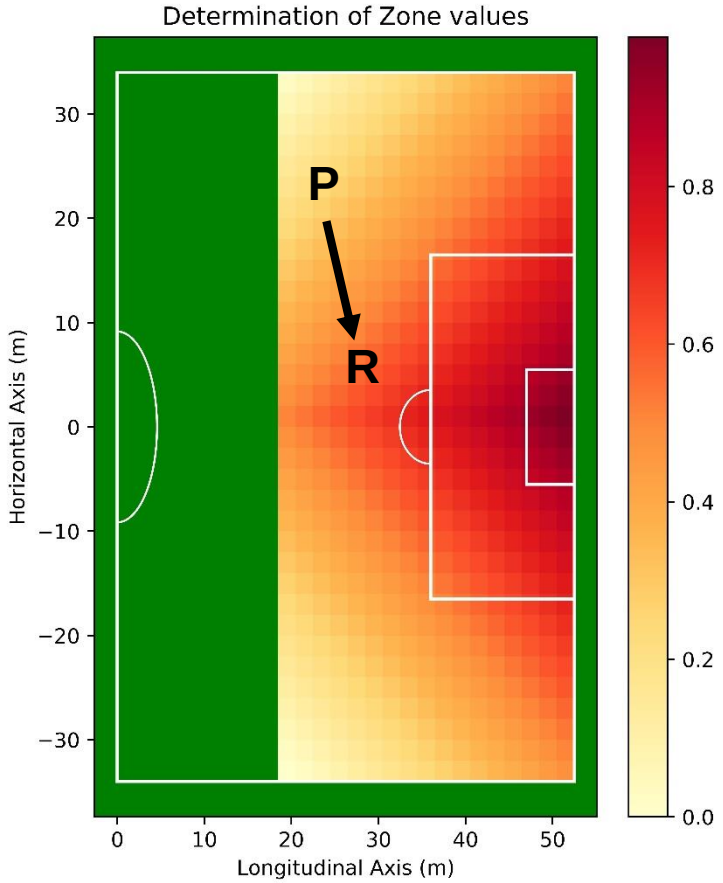


Sport Science Institute Groningen

Preprocessing

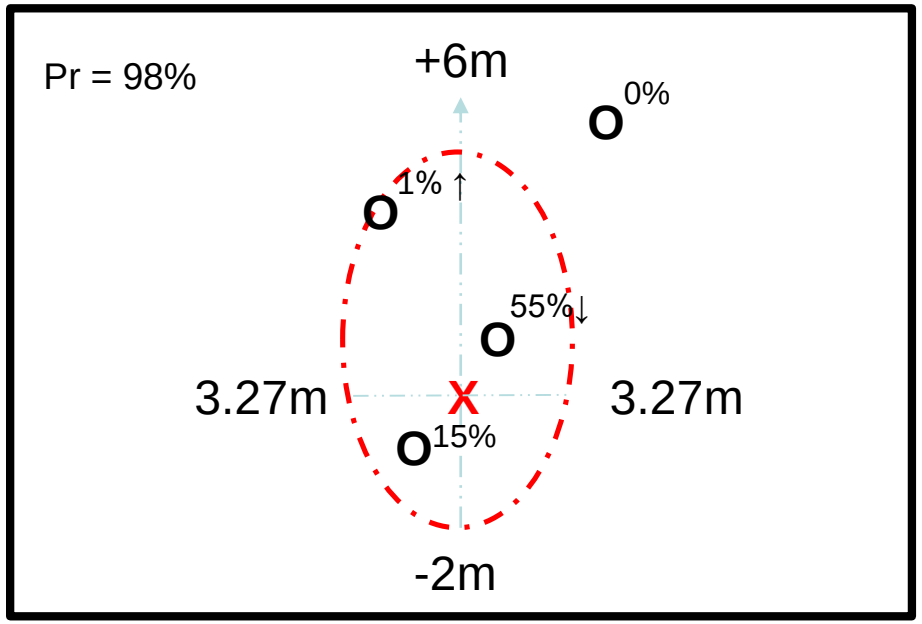
- Automatic detection of passes based on tracking data
- Feature construction
- Compute features over every pass-reception windown based on **principles of play**
- Web-Scrape match outcomes
- Exclude Draws
- N = 89 matches ending in win or lose

Zone Features



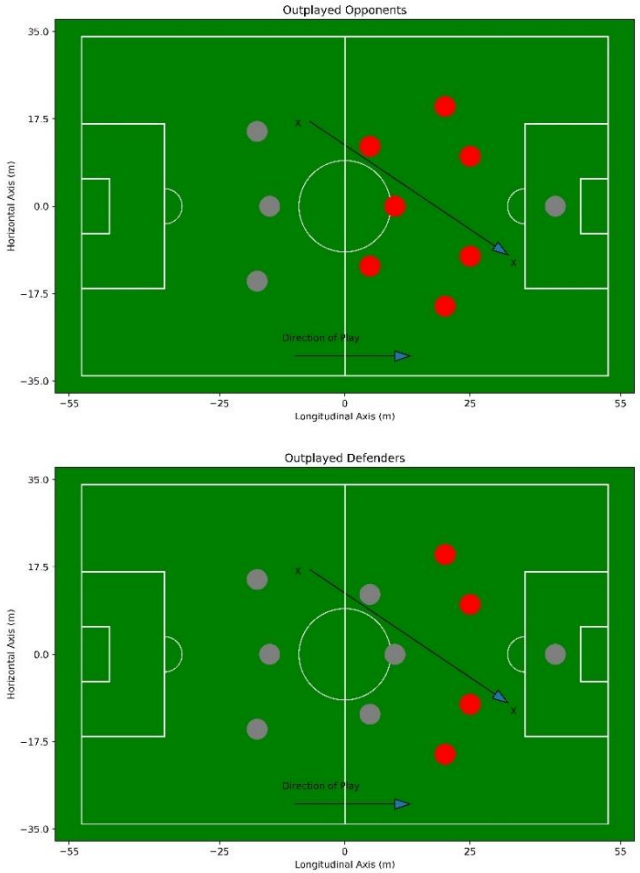
Low-Level Model → Grid-Based Values Only

High-Level Model → Account for Defensive Pressure

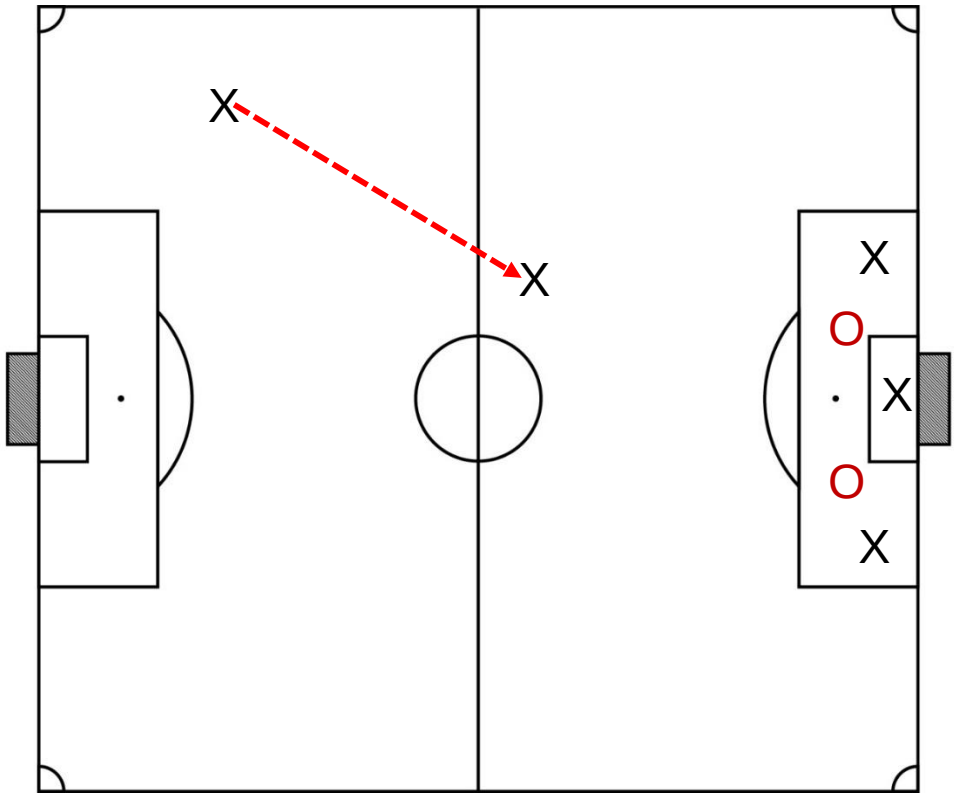


Andrienko et al., 2017

Balance Features

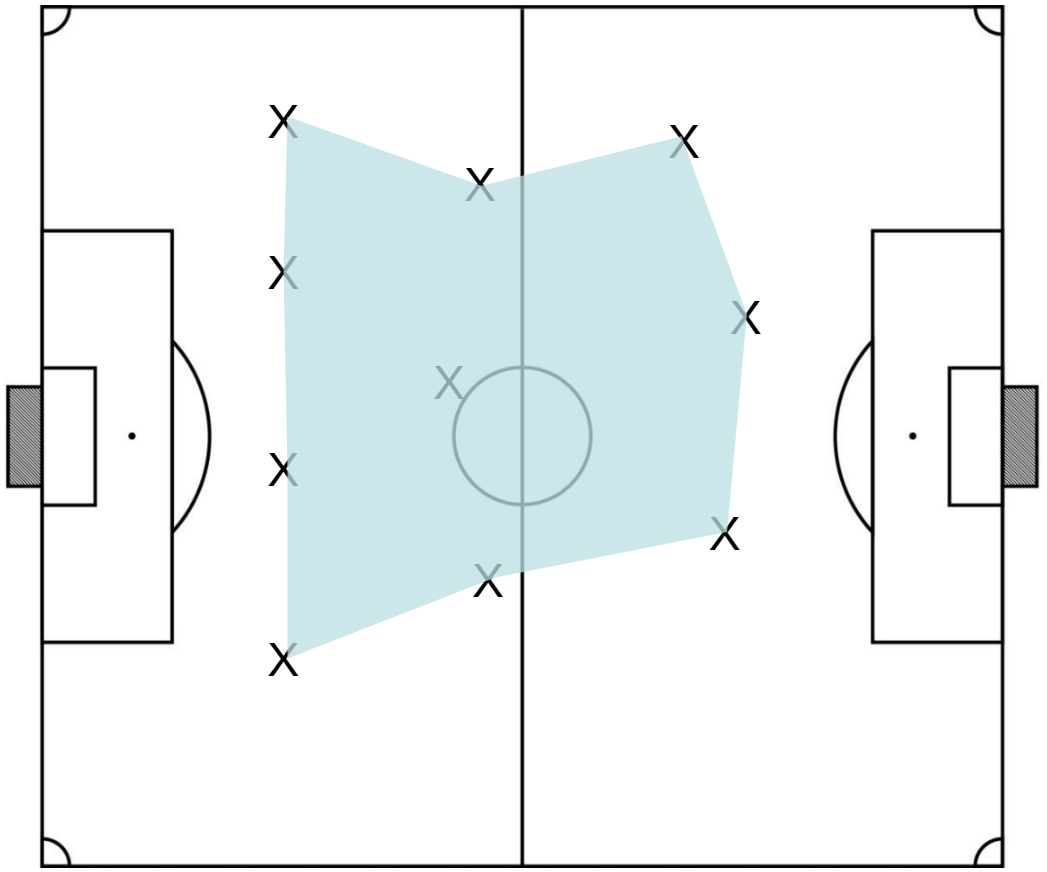


On-the-Ball



Off-the-Ball

Space Mobility Features



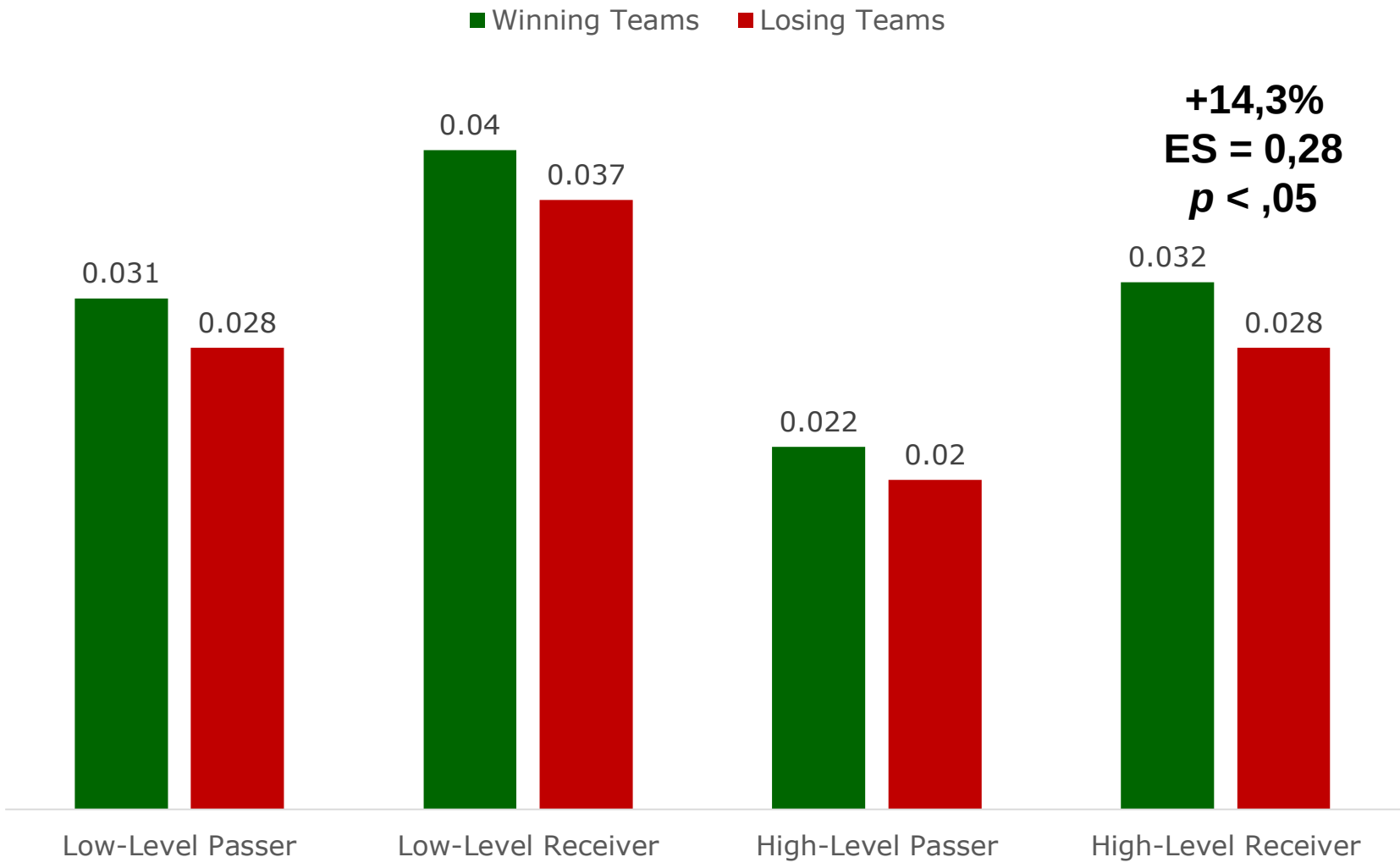
Modelling Match Outcome

- Kruskal-Wallis tests for statistical comparisons at feature level
- Cohen's *d* effect size
- 5-fold CV Logistic Regression ML models (80-20 train-test-split)

  Model 1 → Features with Significant Differences
  Model 2 → Zone Model
  Model 3 → Balance Model
  Model 4 → Space Mobility Model

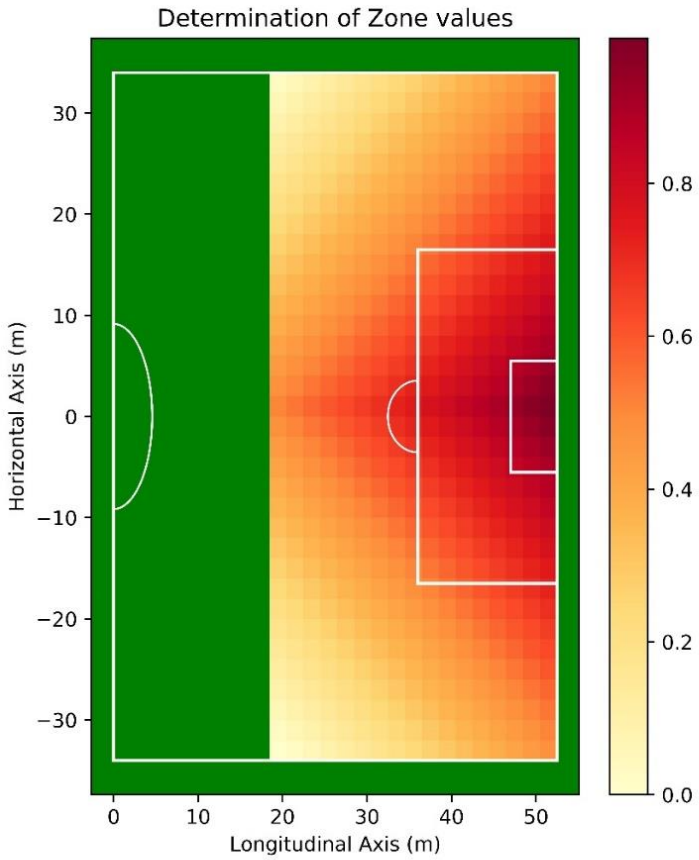
RESULTS & INTERPRETATION

Mean Zone Values Per Pass



Sport Science Institute Groningen

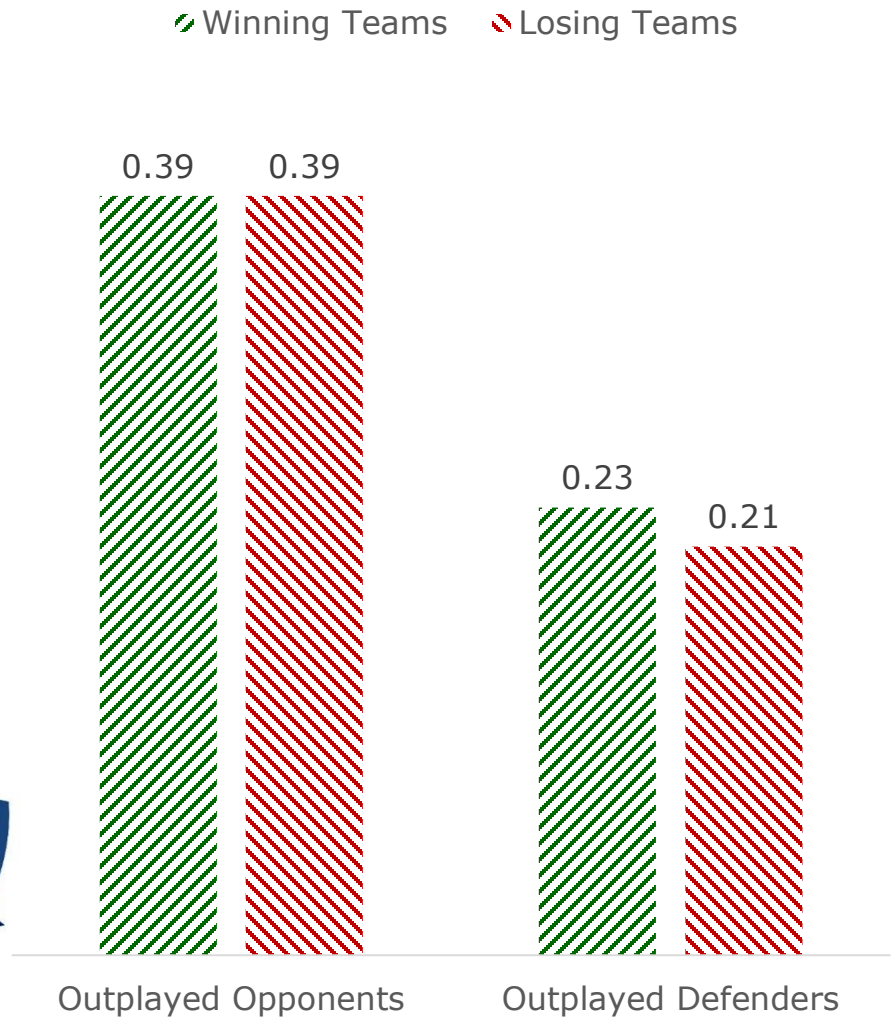
What does this tell us?



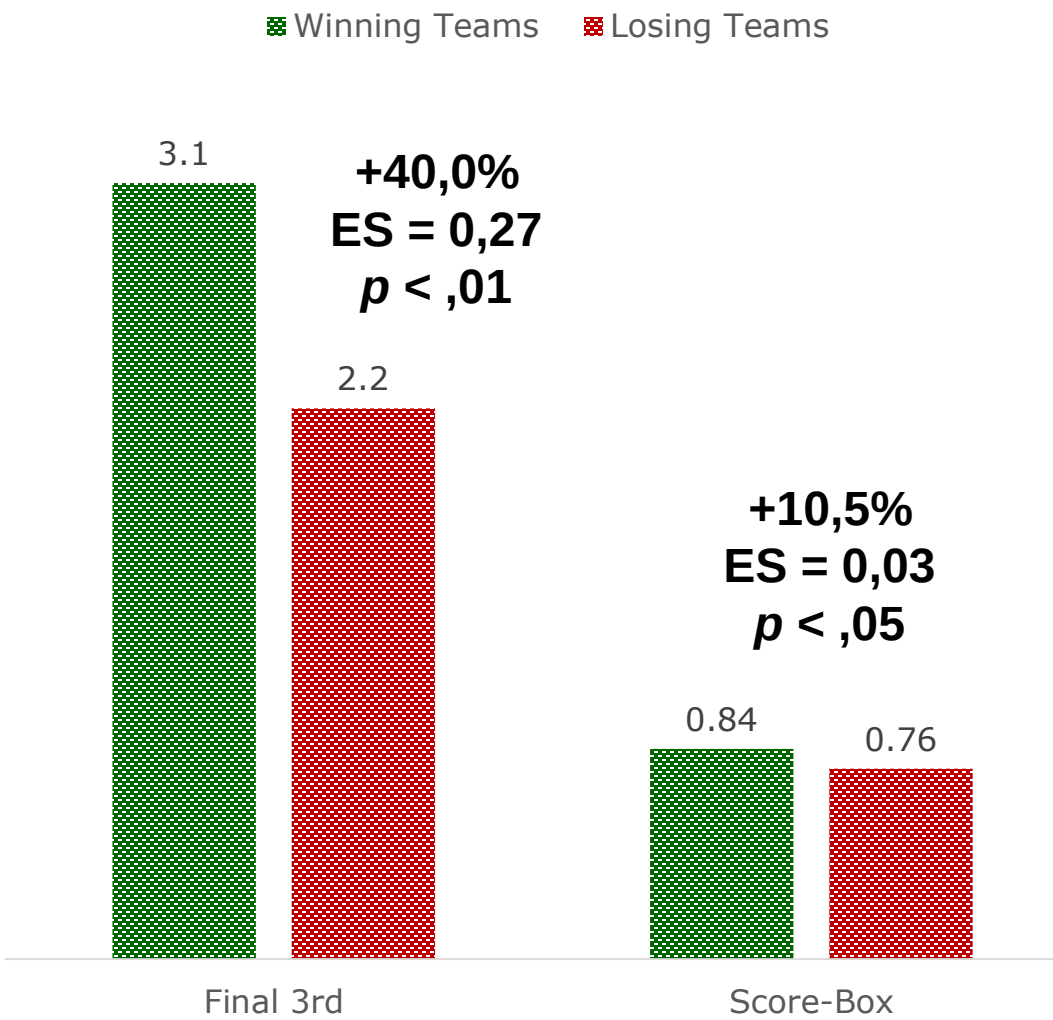
- ‘Zone Related’ Metrics are the most popular metrics in soccer (*xG, Shots, Dangerousity)*
- Weak relationship with match outcome (avg. +10% in winning teams)
- Features that account for context (tracking-based) are better predictors

Sport Science Institute Groningen

Mean **On-the-Ball** Balance



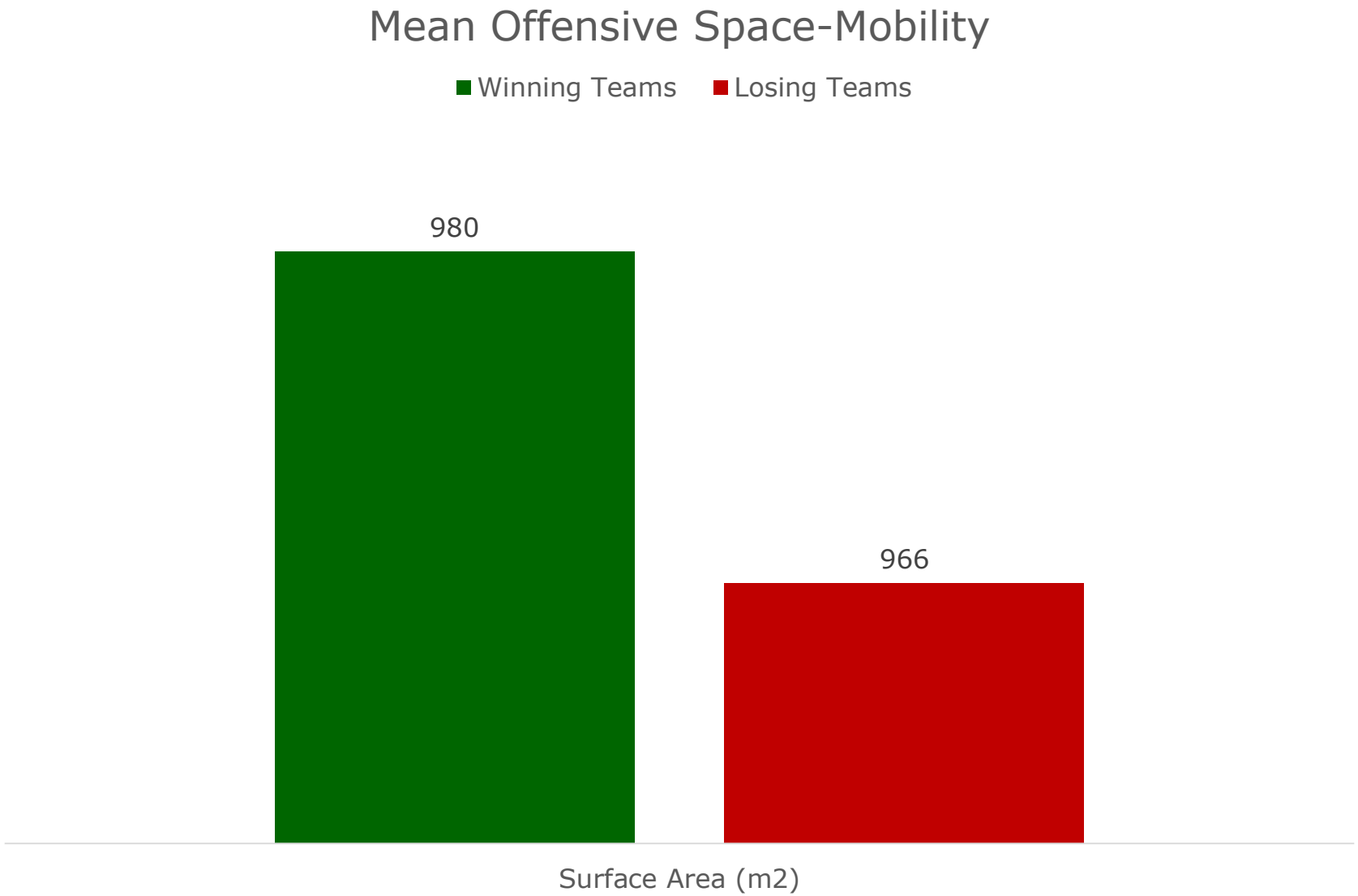
Mean **Off-the-Ball** Balance



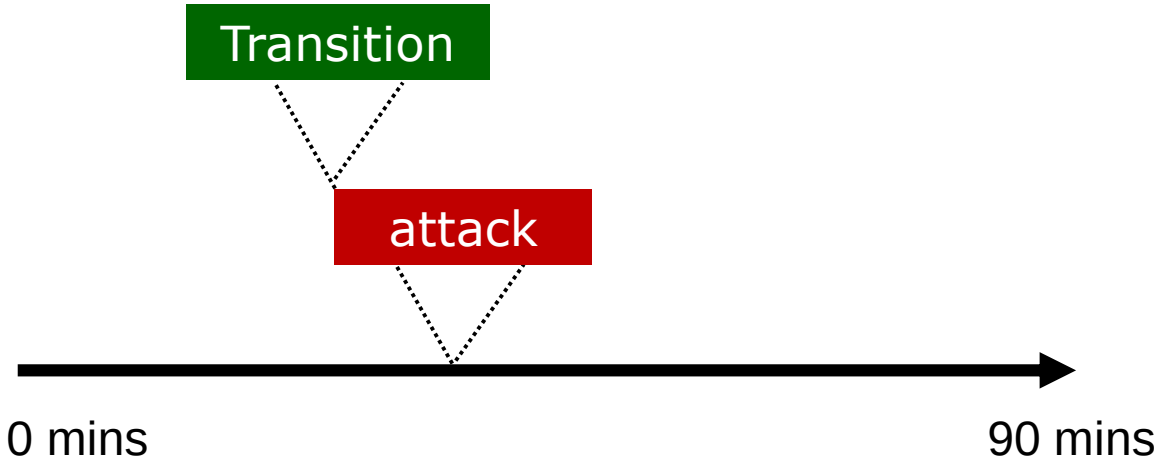
Sport Science Institute Groningen

What does this tell us?

- No relationship between outplayed opponents and winning or losing
- Instead of controlling for context this metric might actually measure context (play-style interactions)
- We are really focussed **on the ball,** but the real information seems to be somewhere else…


Mean Offensive Space-Mobility
Winning Teams
Losing Teams
980
966
Surface Area (m2)

What does this tell us?



- Strong coupling between both team's surface area's & low discriminatory power
- Does the core principle of soccer (**i.e. make the field big when you have the ball)** actually hold
- Does aggregation over a **full** attack capture the concept of *space mobility*

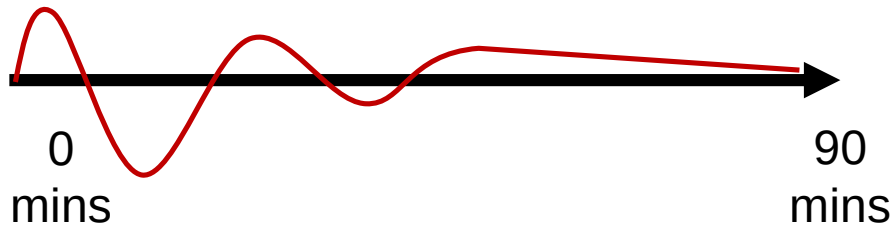
Modelling Match Outcome

	Accuracy	Log Loss
Best Predictors	64%	0,67

y = -0,0167 + 0,136 ***MeanZoneReceiver*** + 0,130 ***Final3rdBalance*** – 0,0162 ***ScoreBoxBalance***

General Discussion & Future Work

- Popular concepts often seem to be **weak** predictors of outcome
- Modelling performance & match outcome requires features that capture the complexity of interpersonal coordination (tactical behaviour)
- From post-game evaluation to in-game prediction and feedback, what are the possibilities



Thank you!

Floris Goes MSc, Dr. Matthias Kempe & Prof. Koen Lemmink

University of Groningen, UMCG, Center for Human Movement Sciences (CHMS)

f.r.goes@umcg.nl | @Florisgoes

Sport Science Institute Groningen