Move it or lose it: Exploring the relation of defensive disruptiveness and team success.

Kempe, Matthias & Goes, Floris,
University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, The Netherlands

Contact: m.kempe@umcg.nl
Starting Point: How to measure tactical performance?

State of the art = Notational Analysis
New Wave: Approaches using tracking data

Space control (Rein et al, 2017)

Dangerousity (Link et al, 2016)

All Goal related
The Groningen Approach- Soccer as a dynamic system?!
Creating Space = Dynamic System?
Variability of inter-team distances associated with match events in elite-standard soccer

WOUTER FRENCKEN\(^1\), HARRO DE POEL\(^1\), CHRIS VISSCHER\(^1\), & KOEN LEHMINK\(^2,3\)

\(^1\)UMCG, University of Groningen, Center for Human Movement Sciences, Groningen, Netherlands; and \(^2\)Hanze University of Applied Sciences, School of Sport Studies, Groningen, Netherlands
New approach: Create moments of bad synchronization (disruption)

1. Idea: Player Movement  = Disruption
2. Idea: Change in Structure  = Disruption
3. Idea: Change in Subunits  = Disruption
1. Idea

(Indicted) Movement → Disruption = Passing Performance

I-MOV

\[
\begin{align*}
I-\text{Mov} & = \frac{(\sum|X_{10}^1 - X_{10}^1| + \ldots + |X_{10}^n - X_{10}^n|)}{n} \\
I-\text{MovY} & = \frac{(\sum|Y_{10}^1 - Y_{10}^1| + \ldots + |Y_{10}^n - Y_{10}^n|)}{n} \\
I-\text{Mov} & = I-\text{MovX} + I-\text{MovY}
\end{align*}
\]
Validating I-Mov

Differentiate Passes & Players

Connects to Player Performance
Complex Approach: Idea 2 & 3

Disruptiveness characteristics
- Centroids
- Surface areas
- Spread

Disruptiveness Variables
- $C_X$
- $C_Y$
- $C_{XDEF}$
- $C_{YDEF}$
- $C_{XMID}$
- $C_{YMID}$
- $C_{XATT}$
- $C_{YATT}$
- $S_{area}$
- $S_F$

PCA

$$D-Def = |PC1| + |PC2| + |PC3|$$
Validating D-Def

\[ D-\text{Def} = |PC1| + |PC2| + |PC3| \]

Differentiate Passes & Players

Lateral

Longitudinal
Issues with D-Def

› Use of arbitrary 3 second window

› Use of starting formations for subunit calculation
New Study – fixing Issues + explore relation to team success

1. Issue: 3s window $\rightarrow$ normalize in change (m) per second

2. Issue: Arbitrary formations for subunit calculation
   - Spilt possession in attacking or defending
   - K-Means for 1.half of a game ($n_{clusters} = 3$) (Bialkowski et al, 2014)
   - Assign every player to one cluster for every timeframe of the game
Study design

- **Study design**
  - 89 Eredivisie Matches (with a winner)
  - Preprocessing Data
    - Calculation of I-Mov & D-Def (+ principal components)
  - Statistics:
    - \( T \)-test (winning vs. losing) for all variables
    - Logistic regression using (winning vs. losing) with differentiating variables (Train-Test split 80%-20%)

---

**Data Processing**

- **Data Filtering**
- **Data Anonymization**
  - J. Example \( \rightarrow \) X. F8gHoIP
- **Notational Data**
  - \(|T(ms)| \rightarrow \|Player\| \rightarrow \|Event\|
    - 5000 23  AST
    - 9530 21  GOAL

---

**Event Detection**

**Spatial Aggregation**
Results

Table 1 - Descriptive statistics winning and losing teams (*: p = .05 ⁿ: p < .05, ⁿ: p < .01)

<table>
<thead>
<tr>
<th></th>
<th>Wins (N = 89)</th>
<th>Losses (N = 89)</th>
<th>Mean Diff.</th>
<th>Effect Size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual Movement (I-Mov)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I-Mov-X (Mean)</td>
<td>0.866m ± 0.673m</td>
<td>0.515m ± 0.675m</td>
<td>+68.1%</td>
<td>0.52 **</td>
</tr>
<tr>
<td>I-Mov-Y (Mean)</td>
<td>0.772m ± 0.600m</td>
<td>0.451m ± 0.591m</td>
<td>+71.2%</td>
<td>0.54 **</td>
</tr>
<tr>
<td>I-Mov (Mean)</td>
<td>1.638m ± 1.268m</td>
<td>0.966m ± 1.265m</td>
<td>+69.6%</td>
<td>0.53 **</td>
</tr>
<tr>
<td><strong>Defensive Disruptiveness (D-Def)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC1 (Mean)</td>
<td>0.018 ± 0.015</td>
<td>0.013 ± 0.022</td>
<td>+34.1%</td>
<td>0.24 *</td>
</tr>
<tr>
<td>PC2 (Mean)</td>
<td>0.010 ± 0.013</td>
<td>0.014 ± 0.033</td>
<td>-23.6%</td>
<td>-0.13</td>
</tr>
<tr>
<td>PC3 (Mean)</td>
<td>-0.026 ± 0.022</td>
<td>-0.021 ± 0.022</td>
<td>-25.5%</td>
<td>-0.25 *</td>
</tr>
<tr>
<td>D-Def (Mean)</td>
<td>0.474 ± 0.048</td>
<td>0.484 ± 0.072</td>
<td>-2.0%</td>
<td>-0.16</td>
</tr>
</tbody>
</table>
Results

› 5-fold cross-validated logistic Regression

Outcome = -0.146 + 0.689 \( I\text{-Mov} \text{Mean} \) + 0.172 \( PC1 \text{Mean} \) - 0.592 \( PC3 \text{Mean} \)

› predicts 69.4 % match outcomes correctly
Discussion

› Solve the previous issues in our model

› Reconsider the lateral component of D-Def

› Spatial-temporal variables can predict match outcome
  • So far good predictors for match outcome were goals & shots on goal’s (Lago-Penas et al., 2010)
  • Passing parameters were poor game by game predictors (Collet, 2013)
Conclusion

› Key Performance Indicators build on spatial- temporal variables are good predictors for match outcome

› Spatial- temporal variables have the opportunity to model team & individual interactions

**Stop using:**

- purely goal related KPI´s
- Event data
Thanks to:

Floris Goes  Dr. Rens Meerhoff
Koen Lemmink

@kempe_matthias  Contact: m.kempe@umcg.nl
For comparison of D-Def vs other performance indicators

Wednesday 12:10 Session Inplay Prediction

“Predicting Match outcome in professional Dutch football using tactical performance metrics computed from position tracking data”

Floris Goes