Entrepreneurship and Soccer Analytics

Howard Hamilton
Founder, Soccermetrics Research & Consulting, LLC

30 September 2011
From Blog to Startup

A post: “Moneyball and Soccer”
A new soccer stat analysis blog
Started attracting interest!
Transition to startup company
Overview

What is soccer analytics?
Why soccer analytics?
What are its challenges?
What is its current state?
What is its future?
What is soccer analytics?

Quantitative assessment of events related to the sport of soccer
Analytics Infrastructure

Data Model → Data

Database → Extraction

Input → Analysis

Output → Post-Process

Deliverable
Why soccer analytics?

Personnel payments are a billion-dollar endeavor

Summer transfer fee payments in English Premier League, 2003-2009
(millions of British pounds)

Source: Deloitte
Why soccer analytics?

Systematic assessment of player and team value
Why soccer analytics?

Maximize value of investments
Create marginal performance gains

Winning
Losing
Why soccer analytics?

External pressures

Competition
Regulations
What are its limitations?

Not a match predictor
Not a “coach-in-a-box”
Initial assumptions
No, I'm not into betting!

Choose your customer:

- Betting industry
- Soccer industry

Morality
Legality

A point to prove
What makes soccer analytics hard?

Data Problem
Limited
Neutral
Summary

Analysis Problem
Complex invasion team sport
## Soccer Data Companies

<table>
<thead>
<tr>
<th>Company</th>
<th>Location</th>
<th>Specialty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opta Sports</td>
<td>London</td>
<td>Live sports data feeds</td>
</tr>
<tr>
<td>Prozone (now Amisco)</td>
<td>Leeds, UK &amp; Nice, France</td>
<td>Real-time player and event tracking</td>
</tr>
<tr>
<td>Tracab</td>
<td>Stockholm</td>
<td>Real-time 3D player and event tracking</td>
</tr>
<tr>
<td>Match Analysis</td>
<td>San Francisco</td>
<td>In-match player and event tracking</td>
</tr>
</tbody>
</table>
Complex Invasion Team Sports

Cooperative action

Move object to defended position

Offense/defense interdependent

Continuous/segmented play

Specialized or hybrid playing roles

Examples
More Examples
What makes soccer unique?

Cooperative and interdependent actions
What makes soccer unique?

Hybrid offensive and defensive roles
What makes soccer unique?

High level of continuous play
Soccer Analytics Challenges

Technical
Difficult to formulate
Difficult to solve

Data
Sparse
Restricted access

Acceptance
Cultural resistance
Technical Challenges

What is soccer, essentially?

- Stochastic processes
- Nonlinear dynamical system
- Run by people

Attempt to study rationally an irrational process

- Difficult to formulate
- Difficult to solve

How good is “good enough?”
Data Challenges

Sparse and neutral summary statistics

Internal data creates historical data
Nonexistent before 1990s
Proprietary
Open Source Data Model

Basic and comprehensive database

Flexible
Extensible

Open-source strategy

Public relations
Open standards
Revenue sources
Acceptance Challenges

Communicating technical information to decision-maker

“We've always done it this way”

“You can't explain everything by statistics”

“What do a bunch of eggheads know about football, anyway?”

“What do a bunch of Americans know about football, anyway?”
Interest vs. Money

Opta Sports' league-specific Twitter feeds

Europe: Where the money is

USA: Where the stats culture is

Source: OnFooty.com (Sarah Rudd)
Improving the Climate for Analytics

Education process required for analysts and end-users

Customers don't always know what they want vis-à-vis analytics

Analysts need to communicate ideas to customers in non-technical language

Be open about the limitations

Danger of overselling
Soccer Analytics: Current State

Clubs have more in-match data

Current level of analysis:
- Micro-level events (tabulations/averages)
- Video overlay

Clubs use small proportion of match data
Soccer Analytics: Current State

Debate over utility of some metrics

Distance run over a match
Time of possession
Number of tackles
Goals-against average

Difficult to develop metrics for “non-glamour” positions

Complex problem, slow progress
<table>
<thead>
<tr>
<th>Blog Title</th>
<th>Author</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soccer by the Numbers</td>
<td>Chris Anderson</td>
<td>Ithaca, NY</td>
</tr>
<tr>
<td>On Football</td>
<td>Sarah Rudd</td>
<td>Seattle, WA</td>
</tr>
<tr>
<td>Soccer Statistically</td>
<td>Ford Bohrmann</td>
<td>Seattle, WA</td>
</tr>
<tr>
<td>11tegen11 (11 on 11)</td>
<td>?</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Rational Football</td>
<td>?</td>
<td>Denmark</td>
</tr>
</tbody>
</table>

**SoccerAnalysts.com**

An aggregation of soccer analytics websites and blogs
## Soccer Analytics Firms

<table>
<thead>
<tr>
<th>Company</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>StatDNA</td>
<td>Chicago, IL</td>
</tr>
<tr>
<td>Chimu Solutions</td>
<td>Chicago, IL</td>
</tr>
<tr>
<td>PowerStats</td>
<td>San Diego, CA</td>
</tr>
<tr>
<td>AIMFútbol</td>
<td>Santiago, Chile</td>
</tr>
<tr>
<td>Sports Statistical Reviews Ltd</td>
<td>Southampton, UK</td>
</tr>
</tbody>
</table>
Future of Soccer Analytics

Integration with other aspects of sports performance

Formation of a truly global transfer market
Developing a Sports Analytics Startup

Funding
Market
Product
Infrastructure
Network
First Client
But before you start...

Passion absolutely required!
Bootstrapping essential
Part-time vs. Full-time
“Market Value” vs. Reality
Interesting vs. Compelling
Exit? What exit??
Summary

Challenges in soccer analytics

Nature of the game

Traditional attitudes

Proceed from first principles

Be open

The history of sports analytics continues to be written
Thank you!

Howard H. Hamilton, Ph.D.
Founder
www.soccermetrics.net
info@soccermetrics.net
Twitter: @soccermetrics