Extended Model for Expected Threat in Soccer

Jirka Poropudas (SportIQ) and Ville-Pekka Inkilä (Football Assocation of Finland)

NESSIS 2021

What is the value of ball possession?

In other words, what is the probability of scoring when a team has the ball at a given location?

Possession Is The Puzzle Of Soccer Analytics. These Models Are Trying To Solve It.

By John Muller

Filed under Soccer

More people have started trying to measure the vast, muddled majority of soccer that happens in between shots. SHAUN BOTTERILL / GETTY IMAGES

Expected Threat (xT)

Originally introduced by Karun Singh in a blog post in 2018.

- Based on simple Markov chain
- Only ball event data
  - No locations for other players
- Offensive model
  - Goals for
- Short time scope
  - Next 5 events

https://karun.in/blog/expected-threat.html

Soccer possession models are gaining steam

Key soccer possession models by publication year, with type of model and possession information

NAME	CREATOR	DEBUT	METHOD	WINDOW	OFF-BALL INFORMATION
Markov Chains	S. Rudd	2011	Markov chain	One possession	Defensive states tagged in event data
Possession-Based Model	N. Mackay	2016	Logistic regression and GAM	One possession	None
Expected Threat (xT)	K. Singh	2019	Markov-like	Next 5 actions (goal for)	None
Valuing Actions by Estimating Probabilities (VAEP)	KU Leuven DTAI	2019	Gradient-boosted trees	Next 10 actions (goal for or against)	Possession history proxies
Expected Possession Value (EPV)	J. Fernández et al.	2019	Multiple models	Next goal (for or against) or end of half	Full tracking data
Possession Value (PV)	Stats Perform	2019	Gradient-boosted trees	Next 10 seconds (goal for)	Possession history proxies
Goals Added (g+)	American Soccer Analysis	2020	Gradient-boosted trees	Two possessions	Possession history proxies
On-Ball Value (OBV)	StatsBomb	2021	Gradient-boosted trees	Two possessions	Broadcast freeze frames (in development)

FiveThirtyEight

Expected threat (xT)

- "Probability of scoring a goal in next 5 events, when in possession of the ball at the given location"
- Model considers two types of events
  - Shot at goal from the current location
    - Scoring probability from a suitable expected goals (xG) model
  - Movement to another location
- The field is divided into a grid of locations
- Event probabilities are calculated from real-life event data

Expected Threat (xT)

Definition of expected threat:　　$xT = S \cdot xG + M \cdot P_M \cdot xT$

Decision model:
- **Shoot** or Move
  - Probabilities depend on location

Dynamics:
- xG
  - Scoring probability
  - Expected goals model!
- P_M

  - Move to new location
  - Transition probabilities

The numerical value of xT is solved by iteration. That is, by repeating the equation five times.

https://karun.in/blog/expected-threat.html

Expected Threat (xT)

Expected Threat (xT) = 0.136

i.e. when the team has the ball in the highlighted zone, they will score in the next **5** actions **13.6%** of the time.

https://karun.in/blog/expected-threat.html

Two-way xT model

Definition: $$xT = S \cdot xG + M \cdot P_M \cdot xT + T \cdot P_{TO} \cdot (-xT)$$

"Decision model":
- Shoot
- Move
- Turnover
  - Probabilities depend on location

Dynamics:
- xG
  - Scoring probability
- P_M
  - Move to new location
- P_{TO}
  - Move to new location following a turnover

The minus sign in last term denotes loss of possession!

Extended events for xT model
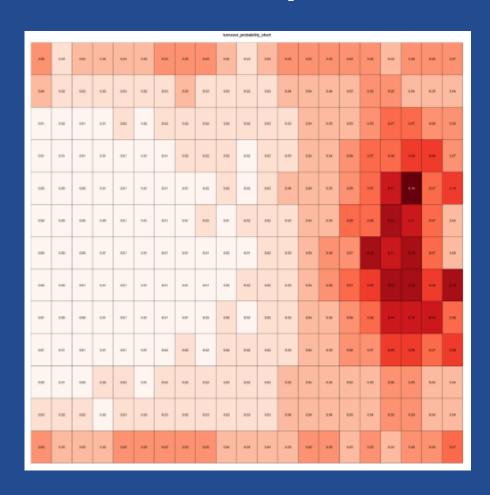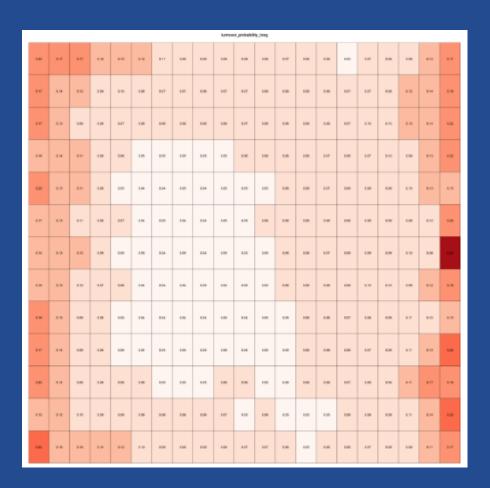
$$xT = S \cdot xG + M_S \cdot P_{MS} \cdot xT + M_L \cdot P_{ML} \cdot xT + T_S \cdot P_{TOS} \cdot (-xT) + T_L \cdot P_{TOL} \cdot (-xT)$$

"Decision model":
- Shoot
- Short or Long move
- Short or Long turnover

Dynamics:
- xG
  - Scoring probability
- P_{MS} and P_{ML}
  - Move to new location
- P_{TOS} or P_{TOL}
  - Move to new location following a turnover

The minus sign in last two terms denotes loss of possession!

Move probability

Transition probabilities for move events
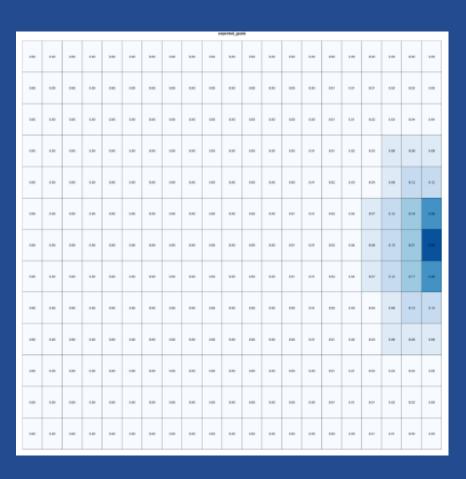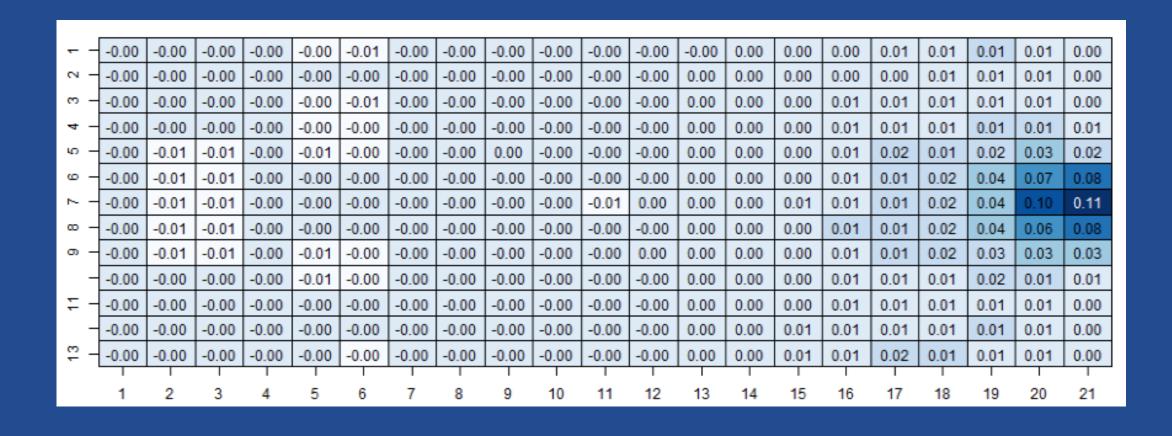
Turnover probability

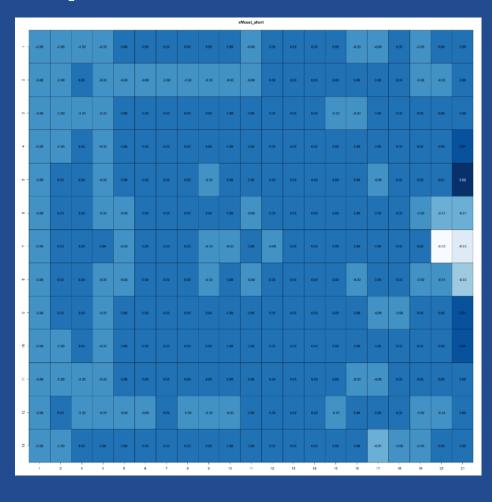# Transition probabilities for turnovers
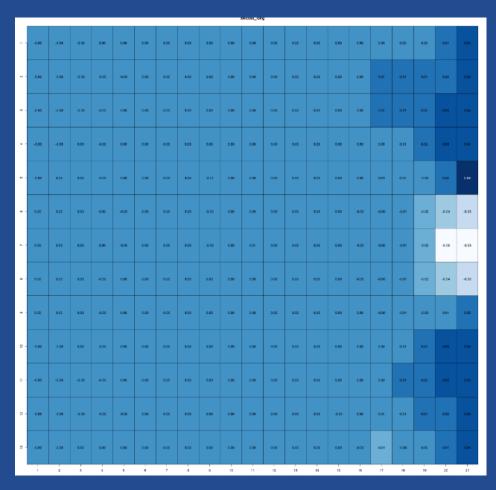
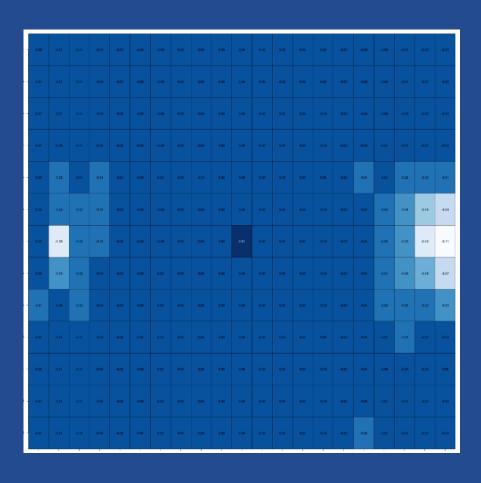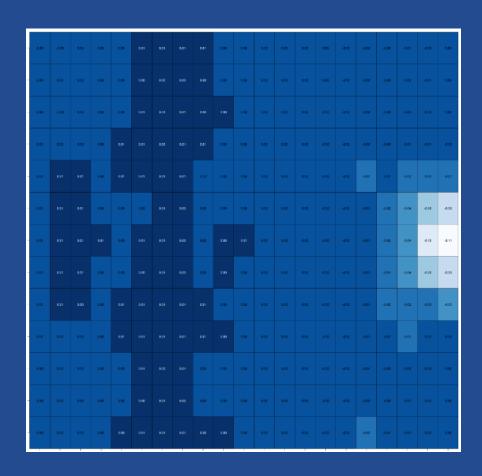# Shot probability and xG model

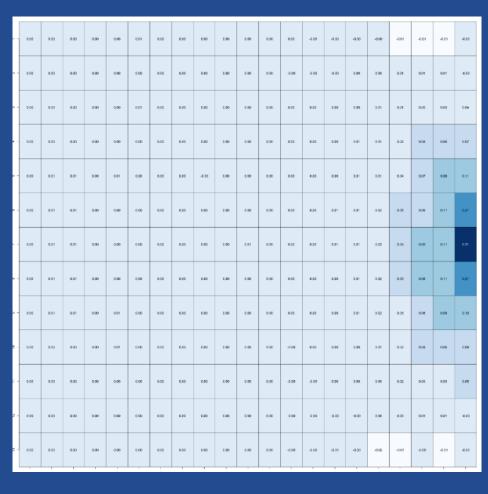# Extended xT

Expected value of movement

Expected cost of turnovers

Expected value of shot

Discussion

- Extended xT model gives a better representation of soccer dynamics
  - Two-way model includes also opponent's scoring
    - Time horizon up to the next shot
  - More detailed state transitions
    - Expected value/cost of events
- Future work
  - Rewarding players properly for their actions
    - "xT added" models should consider the expected value of actions?
  - Include player locations into xT models' state space?
    - For example, StatsBomb 360 data

Any questions or comments?

Jirka Poropudas

@hamahakkimies

jirka.poropudas@gmail.com