Pointwise:

Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data

Dan Cervone

Joint work with Alex D'Amour and Luke Bornn

December 10, 2013

New frontiers for analyzing NBA offenses

#	W	B	#	W	B	#	W	B	#	W	B
1			7			13			19		
2			8			14			20		
3			9			15			21		
4			10			16			22		
5			11			17			23	Be7#	
6			12			18					

Table: Anderssen vs Kieseritsky, 1851

What can we learn about this chess game from "**23. Be7#**"?

New frontiers for analyzing NBA offenses

#	W	B	#	W	B	#	W	B	#	W	B
1	e4	e5	7	d3	Nh5	13	h5	Qg5	19	e5	Qxa1+
2	f4	exf4	8	Nh4	Qg5	14	Qf3	Ng8	20	Ke2	Na6
3	Bc4	Qh4+	9	Nf5	c6	15	Bxf4	Qf6	21	Nxg7+	Kd8
4	Kf1	b5	10	g4	Nf6	16	Nc3	Bc5	22	Qf6+	Nxf6
5	Bxb5	Nf6	11	Rg1	cxb5	17	Nd5	Qxb2	23	Be7#	
6	Nf3	Qh6	12	h4	Qg6	18	Bd6	Bxg1			

Table: Anderssen vs Kieseritsky, 1851

What can we learn about this chess game from "**23. Be7#**"?

Like chess matches, NBA possessions are often won/lost before the ball does/doesn't swish through the net.

Expected Possession Value (EPV)

Overview

We want to learn the value of any spatial configuration of the players/ball.

- Gain insights on players' decision-making by associating changes in this value with players' actions.
- Recognize players' offensive contributions on actions before the end of the possession, and/or away from the ball.
- Understand strategy as a function of space.

We want to learn the value of any spatial configuration of the players/ball.

- Gain insights on players' decision-making by associating changes in this value with players' actions.
- Recognize players' offensive contributions on actions before the end of the possession, and/or away from the ball.
- Understand strategy as a function of space.

EPV assigns value to a spatial configuration by considering the expected points scored by the offense conditional on this configuration.

Some notation:

- Let $t \in [0, T]$ index time during a particular team's possession of the ball.
- $Z_t \in \mathcal{Z}$ is the spatial data at time t.
- X is the number of points the team scores on its possession.

Expected Possession Value (EPV)

Definition and notation

Some notation:

- Let $t \in [0, T]$ index time during a particular team's possession of the ball.
- $Z_t \in \mathcal{Z}$ is the spatial data at time t.
- X is the number of points the team scores on its possession.

Definition

EPV at time t: $\nu_t = E[X|Z_t]$

- $X = E[X|Z_T]$ since X is known from the end of the possession Z_T.

Evaluating EPV

$$\nu_t = \int \int E[X|Z_T]P(Z_T|Z_t, T)P(T|Z_t)dZ_T dT$$

requires modeling the evolution of the possession $P(Z_{t+\epsilon}|Z_t)$.

Evaluating EPV

$$\nu_t = \int \int E[X|Z_T] P(Z_T|Z_t, T) P(T|Z_t) dZ_T dT$$

requires modeling the evolution of the possession $P(Z_{t+\epsilon}|Z_t)$.

- Z_t is high dimensional.
- Components of Z_t coupled in subtle ways.
- Basketball court is a unique spatial domain.

Evaluating EPV

$$\nu_t = \int \int E[X|Z_T]P(Z_T|Z_t, T)P(T|Z_t)dZ_T dT$$

requires modeling the evolution of the possession $P(Z_{t+\epsilon}|Z_t)$.

- Z_t is high dimensional.
- Components of Z_t coupled in subtle ways.
- Basketball court is a unique spatial domain.

Idea: Multiresolution transitions.

- Simplify the necessary probability models.
- Introduce some stationarity for computational tractability.

Multiresolution transitions

A lower resolution process

Let $S_0, S_1, \ldots, S_k$ be a state sequence representation of a possession.

- $S_i \in \mathcal{S}$, a state space discretizing $\mathcal{Z}$.
- $\mathcal{S}$ made up of:

$$\text{ballhandler} \times \text{court region} \times \text{closely defended}$$

- $\mathcal{S}$ also contains absorbing states that mark the end of a possession.

Multiresolution transitions

A lower resolution process

Let $S_0, S_1, \ldots, S_k$ be a state sequence representation of a possession.

- $S_i \in \mathcal{S}$, a state space discretizing $\mathcal{Z}$.
- $\mathcal{S}$ made up of:

$$\text{ballhandler} \times \text{court region} \times \text{closely defended}$$

- $\mathcal{S}$ also contains absorbing states that mark the end of a possession.
- Assume $S_0, S_1, \ldots$ is a homogeneous Markov Chain with transition probability matrix $\mathbb{P}$: $\mathbb{P}_{jk} = P(S_i = s_k | S_{i-1} = s_j)$ for $s_j, s_k \in \mathcal{S}$.
- $T_i \in (0, T]$ is the time transition $S_{i-1} \to S_i$ begins.

Multiresolution transitions

A lower resolution process

Let $S_0, S_1, \ldots, S_k$ be a state sequence representation of a possession.

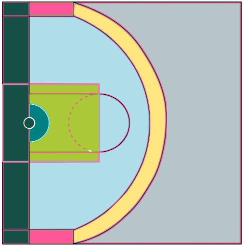
- $S_i \in \mathcal{S}$, a state space discretizing $\mathcal{Z}$.
- $\mathcal{S}$ made up of:

$$\text{ballhandler} \times \text{court region} \times \text{closely defended}$$

- $\mathcal{S}$ also contains absorbing states that mark the end of a possession.
- Assume $S_0, S_1, \ldots$ is a homogeneous Markov Chain with transition probability matrix $\mathbb{P}$: $\mathbb{P}_{jk} = P(S_i = s_k | S_{i-1} = s_j)$ for $s_j, s_k \in \mathcal{S}$.
- $T_i \in (0, T]$ is the time transition $S_{i-1} \to S_i$ begins.

Example:

i	T_i	S_i
0		Parker, center 3, unguarded
1	2.4	Leonard, corner 3, guarded
2	4.8	Duncan, key, guarded
3	6.0	Duncan, restricted area, guarded
4	6.5	Made basket



Macrotransitions are transitions $S_{i-1} \to S_i$ occurring over a longer-than-instantaneous time scale.

- $\tilde{\mathcal{S}}_i = \{s \in \mathcal{S} : S_{i-1} \to s$ represents a pass, shot event, or turnover$\}$.
- $M_t^\epsilon(s) = \{T_i \in (t, t+\epsilon]$ and $s = S_i \in \tilde{\mathcal{S}}_i$ for some $i \leq k\}$, the event that a macrotransition into state s occurs in an ϵ window.
- $M_t^\epsilon = \bigcup_{s \in \mathcal{S}} M_t^\epsilon(s)$ is the event **any** macrotransition occurs in an ϵ window.
- Macrotransition model computes $P(M_t^\epsilon(s)|Z_t)$, $P(M_t^\epsilon|Z_t)$.

Multiresolution transitions

Macro- and micro- models

Macrotransitions are transitions $S_{i-1} \to S_i$ occurring over a longer-than-instantaneous time scale.

- $\tilde{\mathcal{S}}_i = \{s \in \mathcal{S} : S_{i-1} \to s$ represents a pass, shot event, or turnover$\}$.
- $M_t^\epsilon(s) = \{T_i \in (t, t+\epsilon]$ and $s = S_i \in \tilde{\mathcal{S}}_i$ for some $i \leq k\}$, the event that a macrotransition into state s occurs in an ϵ window.
- $M_t^\epsilon = \bigcup_{s \in \mathcal{S}} M_t^\epsilon(s)$ is the event **any** macrotransition occurs in an ϵ window.
- Macrotransition model computes $P(M_t^\epsilon(s)|Z_t)$, $P(M_t^\epsilon|Z_t)$.

Microtransitions occur continuously at times $\{t : t \neq T_i$ for all i$\}$.

- Microtransitions model local evolution of the spatial configurations in the absence of macrotransitions.
- Computes $P(Z_{t+\epsilon}|!M_t^\epsilon, Z_t)$

Multiresolution transitions

Macro- and micro- models

Macrotransitions are transitions $S_{i-1} \to S_i$ occurring over a longer-than-instantaneous time scale.

- $\tilde{\mathcal{S}}_i = \{s \in \mathcal{S} : S_{i-1} \to s$ represents a pass, shot event, or turnover$\}$.
- $M_t^\epsilon(s) = \{T_i \in (t, t+\epsilon]$ and $s = S_i \in \tilde{\mathcal{S}}_i$ for some $i \leq k\}$, the event that a macrotransition into state s occurs in an ϵ window.
- $M_t^\epsilon = \bigcup_{s \in \mathcal{S}} M_t^\epsilon(s)$ is the event **any** macrotransition occurs in an ϵ window.
- Macrotransition model computes $P(M_t^\epsilon(s)|Z_t)$, $P(M_t^\epsilon|Z_t)$.

Microtransitions occur continuously at times $\{t : t \neq T_i$ for all i$\}$.

- Microtransitions model local evolution of the spatial configurations in the absence of macrotransitions.
- Computes $P(Z_{t+\epsilon}|!M_t^\epsilon, Z_t)$

i	T_i	S_i
0		Parker, center 3, unguarded
1	2.4	Leonard, corner 3, guarded
2	4.8	Duncan, key, guarded
3	6.0	Duncan, restricted area, guarded
4	6.5	Made basket

Multiresolution transitions

Rewriting EPV

Assume at time t, someone possesses the ball and we are in state S_{i-1}.

$$\nu_t = \sum_{s \in \tilde{\mathcal{S}}_i} E[X|M_t^{\epsilon}(s), Z_t]P(M_t^{\epsilon}(s)|Z_t) + E[X|!M_t^{\epsilon}, Z_t]P(!M_t^{\epsilon}|Z_t)$$

Multiresolution transitions

Rewriting EPV

Assume at time t, someone possesses the ball and we are in state S_{i-1}.

$$\nu_t = \sum_{s \in \tilde{\mathcal{S}}_i} E[X|M_t^\epsilon(s), Z_t]P(M_t^\epsilon(s)|Z_t) + E[X|!M_t^\epsilon, Z_t]P(!M_t^\epsilon|Z_t)$$

Macrotransitions take $\delta \gg \epsilon$ time to complete:

- $$\begin{aligned} E[X|M_t^\epsilon(s), Z_t] &= \int E[X|Z_{t+\delta}]P(Z_{t+\delta}|M_t^\epsilon(s), Z_t)dZ_{t+\delta} \\ &\approx \int E[X|Z_{t+\delta}]P(Z_{t+\delta}|M_t^\epsilon(s))dZ_{t+\delta} \\ &= E[X|M_t^\epsilon(s)] \end{aligned}$$
- $E[X|M_t^\epsilon(s)] = E[X|S_i = s]$, which is easy to compute using $\mathbb{P}$.

Multiresolution transitions

Rewriting EPV

Assume at time t, someone possesses the ball and we are in state S_{i-1}.

$$\nu_t = \sum_{s \in \tilde{\mathcal{S}}_i} E[X|M_t^\epsilon(s), Z_t] P(M_t^\epsilon(s)|Z_t) + E[X|!M_t^\epsilon, Z_t] P(!M_t^\epsilon|Z_t)$$

Macrotransitions take $\delta \gg \epsilon$ time to complete:

- $$\begin{aligned} E[X|M_t^\epsilon(s), Z_t] &= \int E[X|Z_{t+\delta}] P(Z_{t+\delta}|M_t^\epsilon(s), Z_t) dZ_{t+\delta} \\ &\approx \int E[X|Z_{t+\delta}] P(Z_{t+\delta}|M_t^\epsilon(s)) dZ_{t+\delta} \\ &= E[X|M_t^\epsilon(s)] \end{aligned}$$
- $E[X|M_t^\epsilon(s)] = E[X|S_i = s]$, which is easy to compute using $\mathbb{P}$.

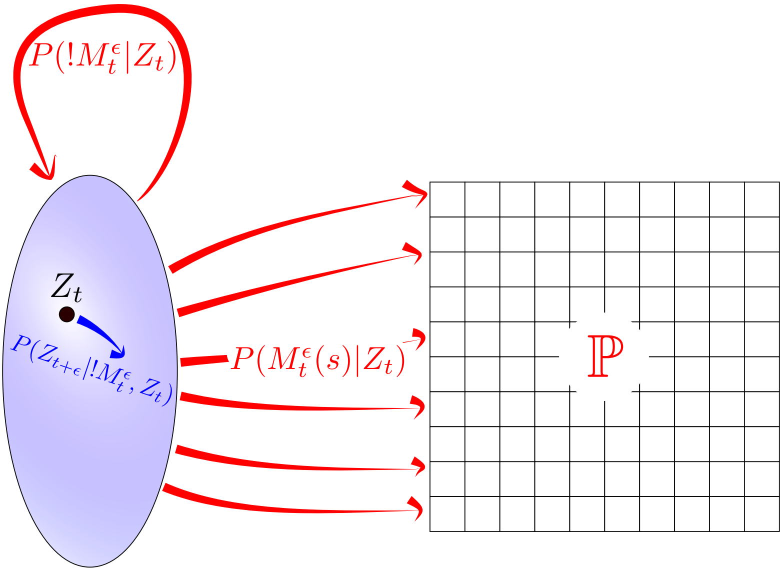
Microtransitions help with:

$$E[X|!M_t^\epsilon, Z_t] = \int E[X|Z_{t+\epsilon}] P(Z_{t+\epsilon}|!M_t^\epsilon, Z_t) dZ_{t+\epsilon}$$

- Looks like the original problem of $E[X|Z_t]$, but we only need local integration.

Multiresolution transitions

Putting it all together



With player ℓ with the ball at time t (in state S_{i-1}), there are 7 possible macrotransitions ($\tilde{\mathcal{S}}_i = \{s_1, \dots, s_7\}$).

- $S_i \to s_j$ for $j = 1, \dots, 4$ are passes to each of four teammates.
- $S_i \to s_5$ made shot; $S_i \to s_6$ missed shot.
- $S_i \to s_7$ turnover.

Macrotransition probabilities

Spatial competing risks (Prentice et al, 1978; Cox, 1959)

With player ℓ with the ball at time t (in state S_{i-1}), there are 7 possible macrotransitions ($\tilde{\mathcal{S}}_i = \{s_1, \dots, s_7\}$).

- $S_i \to s_j$ for $j = 1, \dots, 4$ are passes to each of four teammates.
- $S_i \to s_5$ made shot; $S_i \to s_6$ missed shot.
- $S_i \to s_7$ turnover.

Define the hazard for each macrotransition:

$$\lambda_j^{\ell}(t) = \lim_{\epsilon \to 0} P(M_t^{\epsilon}(s)|Z_t)/\epsilon$$

Macrotransition probabilities

Spatial competing risks (Prentice et al, 1978; Cox, 1959)

With player ℓ with the ball at time t (in state S_{i-1}), there are 7 possible macrotransitions ($\tilde{\mathcal{S}}_i = \{s_1, \ldots, s_7\}$).

- $S_i \to s_j$ for $j = 1, \ldots, 4$ are passes to each of four teammates.
- $S_i \to s_5$ made shot; $S_i \to s_6$ missed shot.
- $S_i \to s_7$ turnover.

Define the hazard for each macrotransition:

$$\lambda_j^\ell(t) = \lim_{\epsilon \to 0} P(M_t^\epsilon(s)|Z_t)/\epsilon$$

$$\log(\lambda_j^\ell(t)) = [W_j^\ell]_t' \beta_j^\ell + \varphi_j^\ell\left(z_t^\ell\right) + \left(\tilde{\varphi}_j^\ell\left(z_t^j\right)\mathbf{1}[j \leq 4]\right)$$

Macrotransition probabilities

Spatial competing risks (Prentice et al, 1978; Cox, 1959)

With player ℓ with the ball at time t (in state S_{i-1}), there are 7 possible macrotransitions ($\tilde{\mathcal{S}}_i = \{s_1, \dots, s_7\}$).

- $S_i \to s_j$ for $j = 1, \dots, 4$ are passes to each of four teammates.
- $S_i \to s_5$ made shot; $S_i \to s_6$ missed shot.
- $S_i \to s_7$ turnover.

Define the hazard for each macrotransition:

$$\lambda_j^\ell(t) = \lim_{\epsilon \to 0} P(M_t^\epsilon(s)|Z_t)/\epsilon$$

$$\log(\lambda_j^\ell(t)) = [W_j^\ell]_t' \beta_j^\ell + \varphi_j^\ell\left(z_t^\ell\right) + \left(\tilde{\varphi}_j^\ell\left(z_t^j\right)\mathbf{1}[j \leq 4]\right)$$

- $[W_j^\ell]_t$ vector of (time-varying) covariates and β_j^ℓ its coefficients.
- z_t^ℓ is player ℓ's location at time t; z_t^j is the teammate's location at time t corresponding to the pass event for s_j.
- $\varphi_j^\ell, \tilde{\varphi}_j^\ell$ are spatial random effects surfaces (GRFs).

Macrotransition probabilities

Parameter estimation

The likelihood for player ℓ (ℓ superscripts omitted):

$$\mathsf{Lik}(\lambda_1, \ldots, \lambda_7) = \prod_{j=1}^{7} \mathsf{Lik}(\lambda_j)$$

$$\mathsf{Lik}(\lambda_j) = \left(\prod_{i=1}^{m_j} \exp(\lambda_j(T_i^j)) \right) \exp\left(- \int_{\mathcal{T}} \exp(\lambda_j(t)) dt \right)$$

$$\text{where } \log(\lambda_j(t)) = [W_j]_t' \beta_j + \varphi_j\left(z_t\right) + \left(\tilde{\varphi}_j \left(z_t^j \right) \mathbf{1}[j \leq 4] \right)$$

Macrotransition probabilities

Parameter estimation

The likelihood for player ℓ (ℓ superscripts omitted):

$$\text{Lik}(\lambda_1, \dots, \lambda_7) = \prod_{j=1}^{7} \text{Lik}(\lambda_j)$$

$$\text{Lik}(\lambda_j) = \left(\prod_{i=1}^{m_j} \exp(\lambda_j(T_i^j)) \right) \exp\left(-\int_{\mathcal{T}} \exp(\lambda_j(t)) dt \right)$$

$$\text{where } \log(\lambda_j(t)) = [W_j]_t' \beta_j + \varphi_j(z_t) + \left(\tilde{\varphi}_j \left(z_t^j \right) \mathbf{1}[j \leq 4] \right)$$

- Unknown parameters: $\lambda_j = \{\beta_j, \varphi_j(\cdot), \tilde{\varphi}_j(\cdot)\}$ for $j = 1, \dots, 4$.
  - $\lambda_j = \{\beta_j, \varphi_j(\cdot)\}$ for $j = 5, 6, 7$
- Data: $T_i^j, i = 1, \dots, m_j$ times of macrotransition j.
  - $z_t, t \in \mathcal{T}$ tracks player's location while he has ball possession during whole season (similarly for z_t^j).
  - $[W_j]_t$ p_j-dimensional time-referenced covariates.

Macrotransition probabilities

Parameter estimation

The likelihood for player ℓ (ℓ superscripts omitted):

$$\text{Lik}(\lambda_1, \dots, \lambda_7) = \prod_{j=1}^{7} \text{Lik}(\lambda_j)$$

$$\text{Lik}(\lambda_j) = \left(\prod_{i=1}^{m_j} \exp(\lambda_j(T_i^j)) \right) \exp\left(-\int_{\mathcal{T}} \exp(\lambda_j(t)) dt \right)$$

$$\text{where } \log(\lambda_j(t)) = [W_j]_t' \beta_j + \varphi_j(z_t) + \left(\tilde{\varphi}_j \left(z_t^j \right) \mathbf{1}[j \leq 4] \right)$$

- Unknown parameters: $\lambda_j = \{\beta_j, \varphi_j(\cdot), \tilde{\varphi}_j(\cdot)\}$ for $j = 1, \dots, 4$.
    - $\lambda_j = \{\beta_j, \varphi_j(\cdot)\}$ for $j = 5, 6, 7$
- Data: $T_i^j, i = 1, \dots, m_j$ times of macrotransition j.
    - $z_t, t \in \mathcal{T}$ tracks player's location while he has ball possession during whole season (similarly for z_t^j).
    - $[W_j]_t$ p_j-dimensional time-referenced covariates.
- Priors: hierarchical priors for $\beta_j, \varphi_j, \tilde{\varphi}_j$ across ℓ.
- Many computational challenges!

Parameter estimation

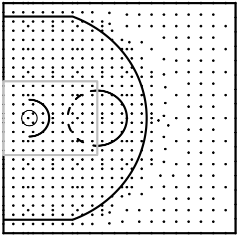
SPDE/INLA approach (Rue, 2009; Lindgren, 2011)

Functional bases for Guassian random fields.

- Approximate infinite-dimensional $\varphi : \mathbb{R}^2 \to \mathbb{R}$ with finite $\mathbf{w} \in \mathbb{R}^k$.
- $\xi_i : \mathbb{R}^2 \to \mathbb{R}, i = 1, \dots k$ are fixed basis functions.
- $\mathbf{w} \sim N_k(0, \Sigma)$ such that for $s \in \mathbb{R}^2$

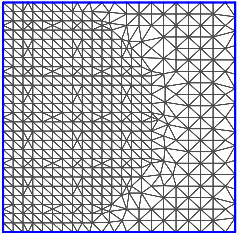
$$\varphi(x) \dot{\sim} \xi(x) = \sum_{i=1}^{k} \xi_i(x) w_i$$

Functional bases for Guassian random fields.



- Approximate infinite-dimensional $\varphi : \mathbb{R}^2 \to \mathbb{R}$ with finite $\mathbf{w} \in \mathbb{R}^k$.
- $\xi_i : \mathbb{R}^2 \to \mathbb{R}, i = 1, \ldots k$ are fixed basis functions.
- $\mathbf{w} \sim N_k(0, \Sigma)$ such that for $s \in \mathbb{R}^2$

$$\varphi(x) \dot{\sim} \xi(x) = \sum_{i=1}^{k} \xi_i(x) w_i$$

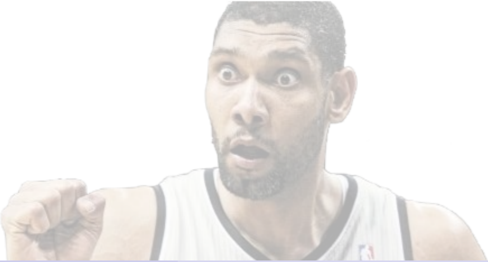


- If $\varphi(x)$ has a thrice-differentiable Matern covariance function, we can obtain accurate, easy-to-compute, GMRF approximation by considering $\{\xi_i(x), i = 1, \ldots, k\}$ piecewise linear.
- Large p log-linear model.
- Integrated Nested Laplace Approximation for latent GMRF GLMs.

Some results

Covariate coefficients

Fixed effects (β_j) for Tim Duncan:



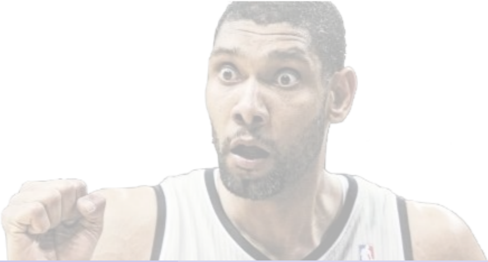
Some results

Covariate coefficients

Fixed effects (β_j) for Tim Duncan:

Covariate	Est	SE
Nearest defender	0.018	0.115
Ball movement	0.028	0.014
Dribble	1.121	0.008
Velocity (x)	0.053	0.050
Velocity (y)	-0.013	0.014
PG closest	0.806	0.019
PG second closest	-0.048	0.050
PG third closest	0.000	0.975
Defense on PG	-2.019	0.976
Constant	-5.037	0.139

Table: Fixed effect estimates and standard errors for coefficients of **pass-to-PG (mostly Parker) hazard** model



Some results

Covariate coefficients

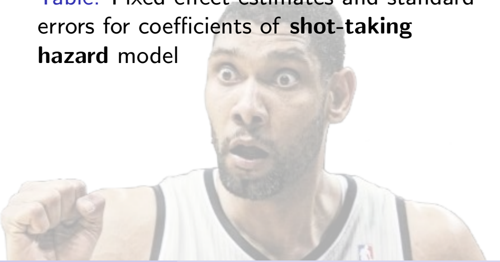
Fixed effects (β_j) for Tim Duncan:

Covariate	Est	SE
Nearest defender	0.018	0.115
Ball movement	0.028	0.014
Dribble	1.121	0.008
Velocity (x)	0.053	0.050
Velocity (y)	-0.013	0.014
PG closest	0.806	0.019
PG second closest	-0.048	0.050
PG third closest	0.000	0.975
Defense on PG	-2.019	0.976
Constant	-5.037	0.139

Table: Fixed effect estimates and standard errors for coefficients of **pass-to-PG (mostly Parker) hazard** model

Covariate	Est	SE
Nearest defender	0.116	0.076
Ball movement	0.105	0.013
Dribble	0.423	0.006
Velocity (x)	0.056	0.032
Velocity (y)	0.050	0.009
Constant	-6.793	0.011

Table: Fixed effect estimates and standard errors for coefficients of **shot-taking hazard** model



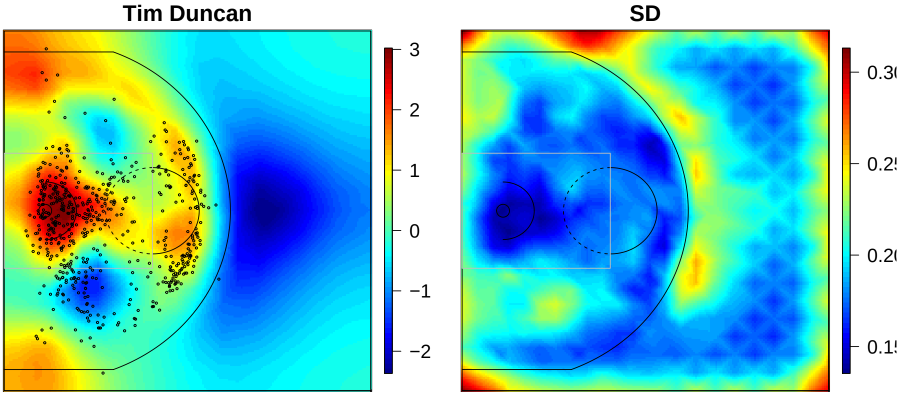


Figure: Spatial random effect surface for Tim Duncan's shot-taking hazard

Some results

Spatial random effect surfaces

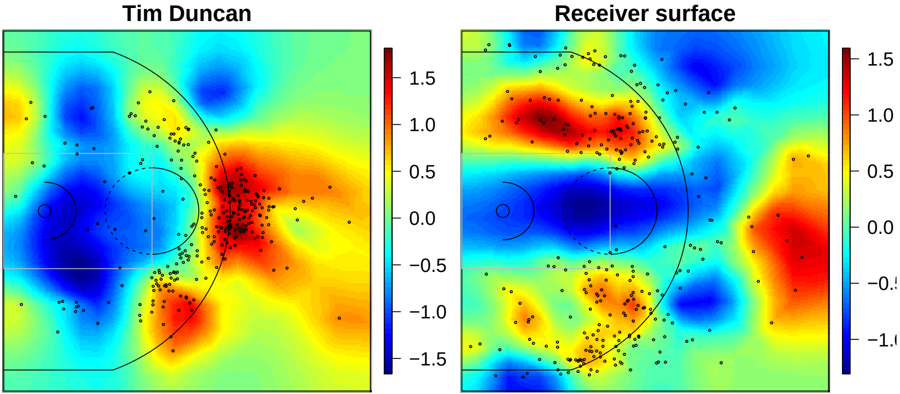


Figure: Spatial random effect surface for Tim Duncan's pass-to-PG hazard

The future of EPV

Quantifying players' decision-making

We gain insight on players' decision-making by tracking EPV across during ball movement

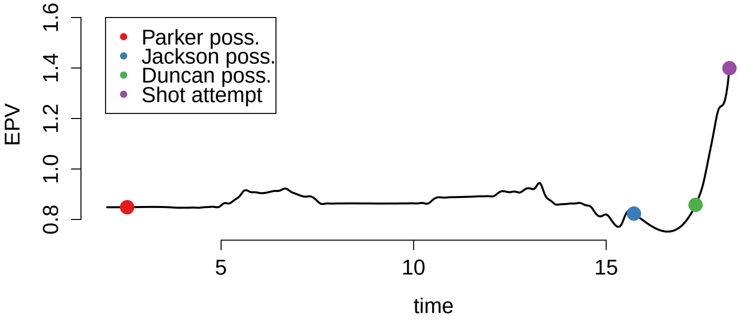
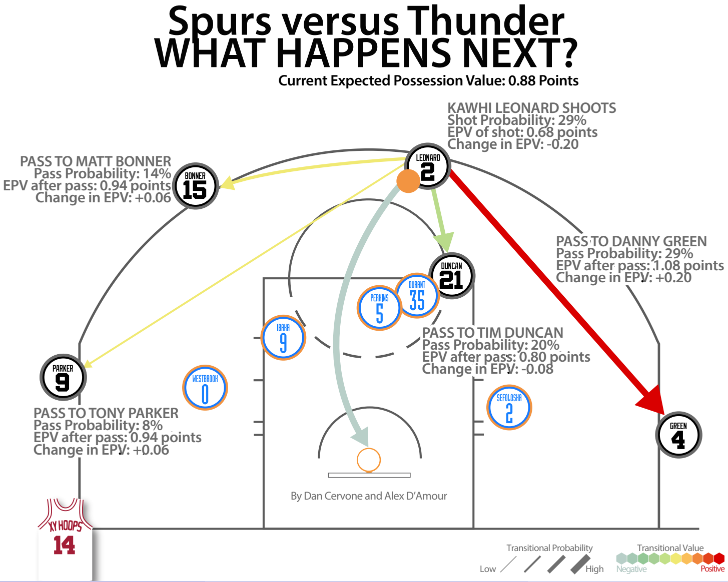


Figure: EPV of a possession during Thunder-Spurs game. Colored dots are macrotransitions (passes and shots)


Spurs versus Thunder
WHAT HAPPENS NEXT?
Current Expected Possession Value: 0.88 Points
KAWHI LEONARD SHOOTS
Shot Probability: 29%
EPV of shot: 0.68 points
Change in EPV: -0.20
PASS TO MATT BONNER
Pass Probability: 14%
EPV after pass: 0.94 points
Change in EPV: +0.06
PASS TO DANNY GREEN
Pass Probability: 29%
EPV after pass: 1.08 points
Change in EPV: +0.20
PASS TO TIM DUNCAN
Pass Probability: 20%
EPV after pass: 0.80 points
Change in EPV: -0.08
PASS TO TONY PARKER
Pass Probability: 8%
EPV after pass: 0.94 points
Change in EPV: +0.06
LEONARD 2
BONNER 15
DUNCAN 21
PARKER 9
GREEN 4
DURANT 35
PERKINS 5
IBAKA 9
WESTBROOK 0
SEFOLOSHA 2
By Dan Cervone and Alex D'Amour
NY HOOPS 14
Transitional Probability
Low
High
Transitional Value
Negative
Positive

The future of EPV

Areas of further work

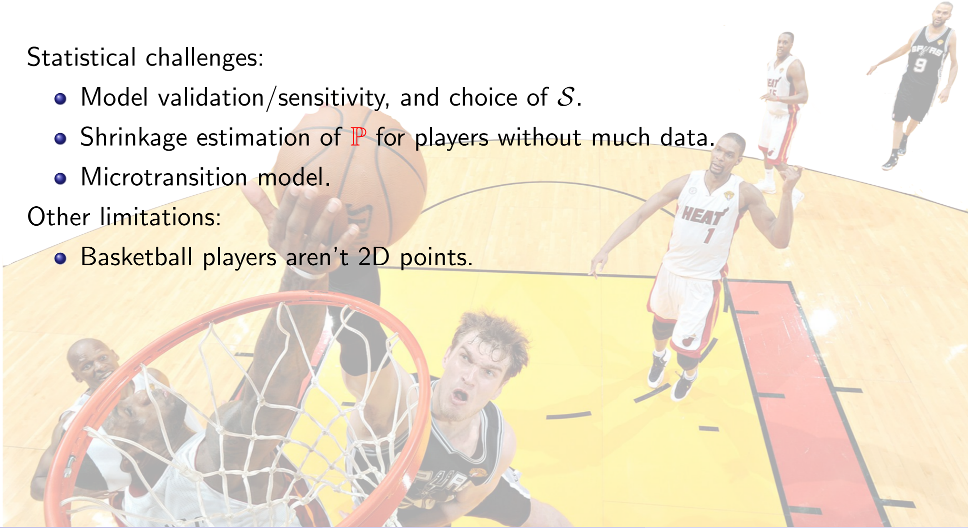
Statistical challenges:

- Model validation/sensitivity, and choice of $\mathcal{S}$.
- Shrinkage estimation of $\mathbb{P}$ for players without much data.
- Microtransition model.

Other limitations:

The future of EPV

Areas of further work



Statistical challenges:

- Model validation/sensitivity, and choice of $\mathcal{S}$.
- Shrinkage estimation of $\mathbb{P}$ for players without much data.
- Microtransition model.

Other limitations:

- Basketball players aren't 2D points.

The future of EPV

Areas of further work

Statistical challenges:

- Model validation/sensitivity, and choice of $\mathcal{S}$.
- Shrinkage estimation of $\mathbb{P}$ for players without much data.
- Microtransition model.

Other limitations:

- Basketball players aren't 2D points.
- Our models are still too limited regarding certain motifs.
- Interpretability of EPV.