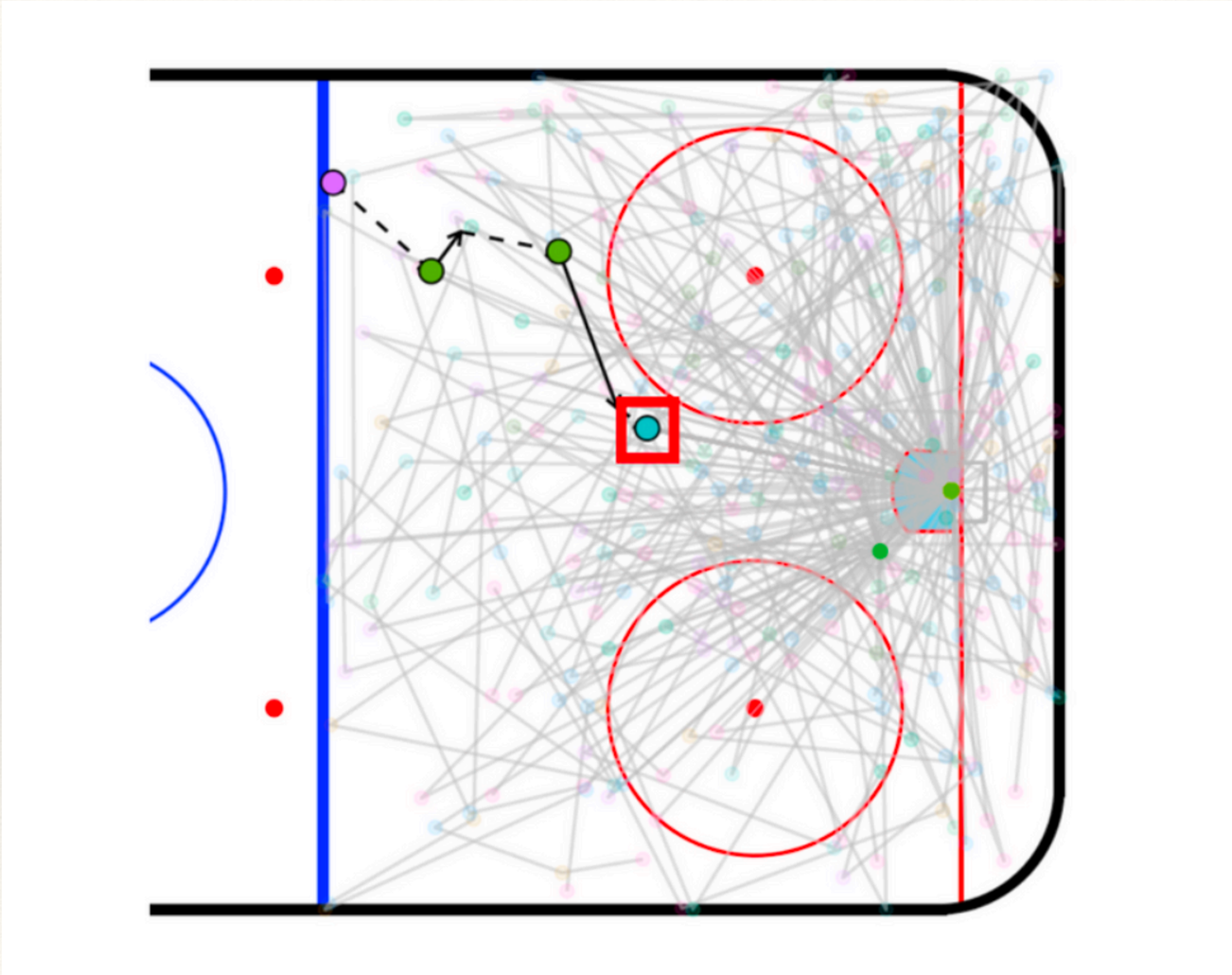
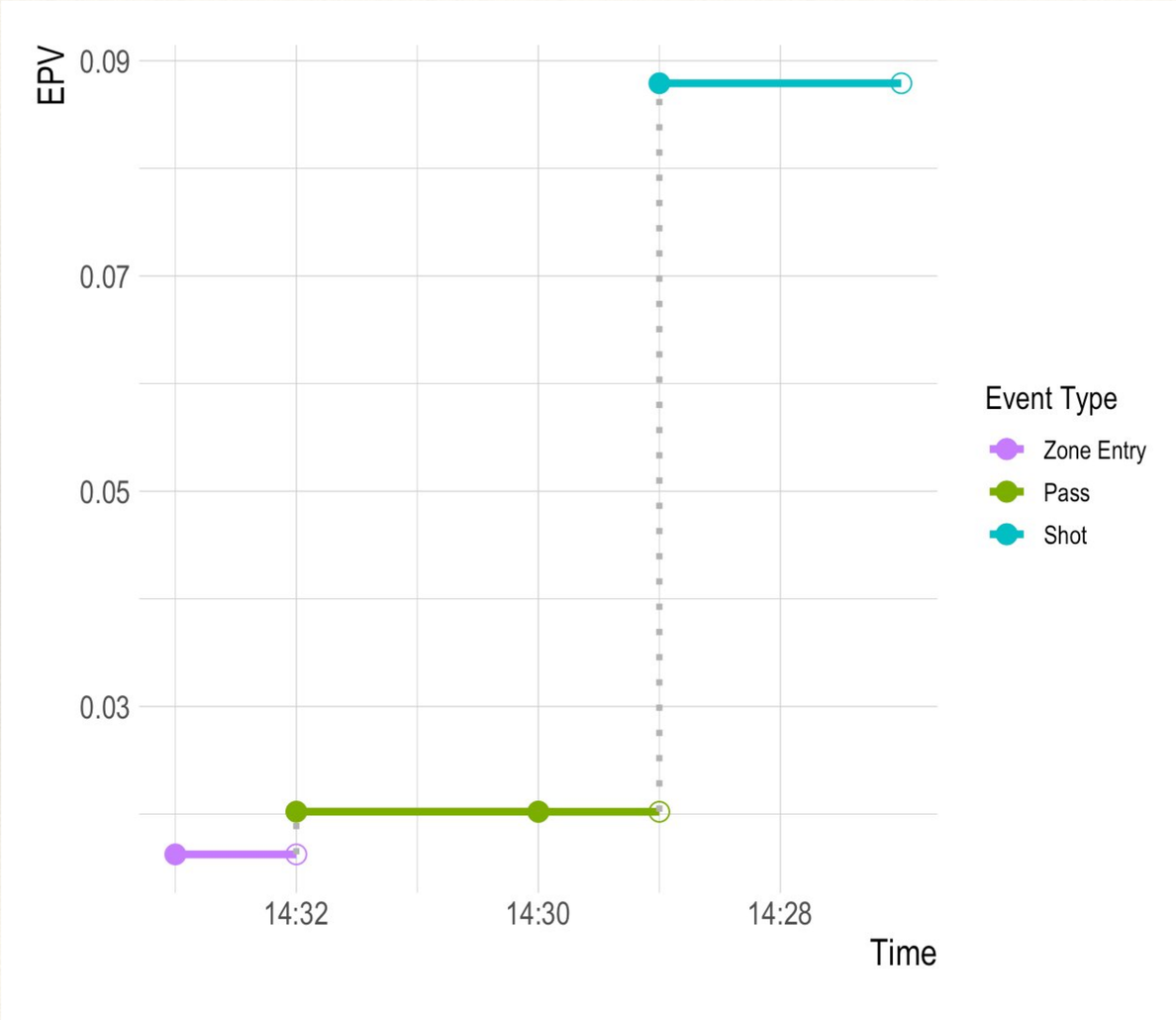
# A simulation approach to EPV for Hockey

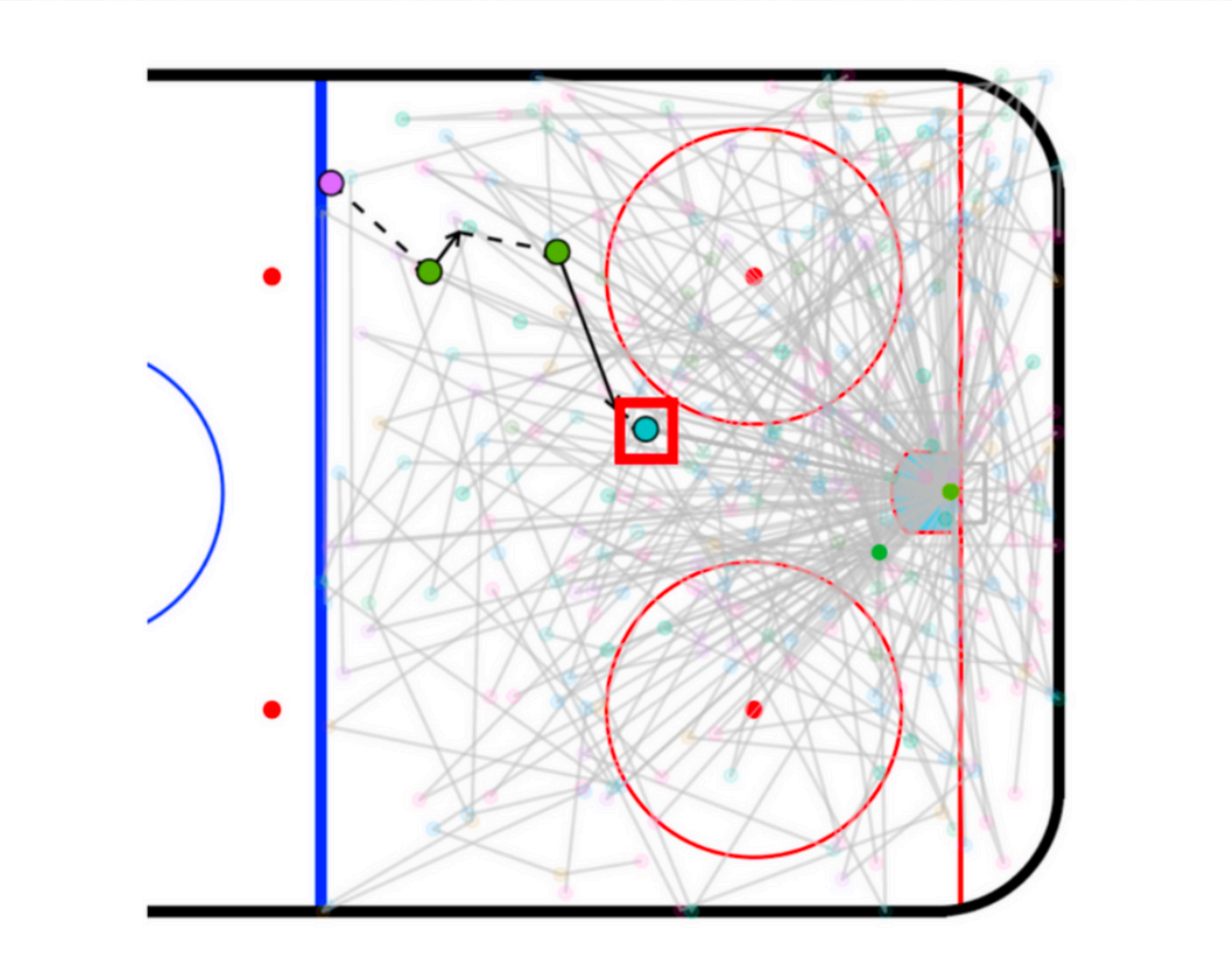
Tyrel Stokes
Brendan Kumagai
Mikael Nahabedian
Thibaud Châtel

# Expected Possession Value (EPV)



# Expected Possession Value (EPV)



# Preceding Models

## Full Tracking EPV

Basketball - Cervone et al (2014)

Soccer - Fernandez et al (2020)

## Discrete Expected Threat Models

Soccer - Singh et al (2018)

Hockey - Yu et al (2020)

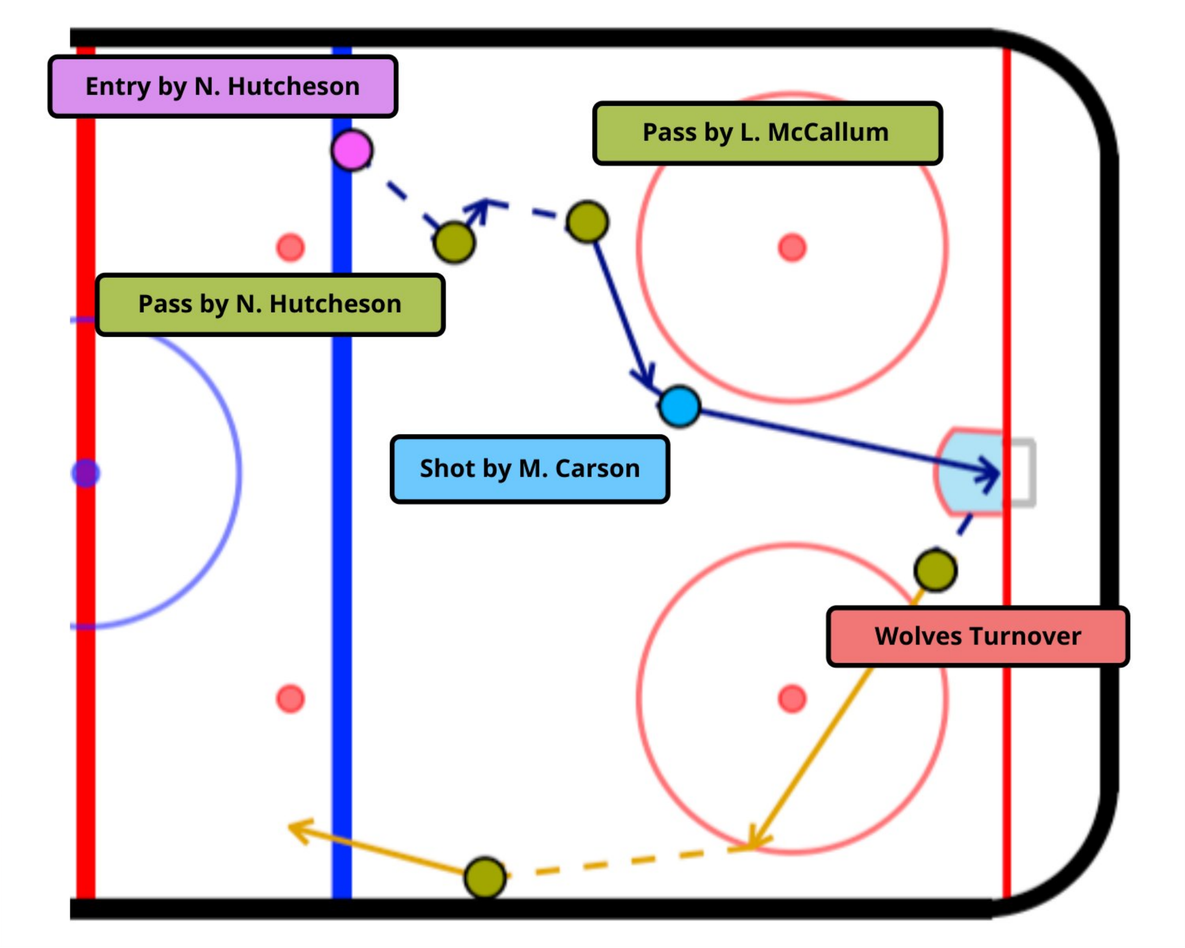
# Data Source: Big Data Cup Data (Stathletes)

- 40+ Games from the Erie Otters

- Extended Play-by-Play (Things like passes with locations (passer and receiver) beyonds usual play-by-play)

- We have the location of the puck and players involved in the actions, but NOT the other players on the ice nor their locations

- ~ 70,000 Events




| Clock | Event | Team | Player | x | y | Player 2 | Target x | Target y | Details | |
|---|---|---|---|---|---|---|---|---|---|---|
| 14:33 | Zone Entry | Sudbury Wolves | Nolan Hutcheson | 126 | 74 | Jacob Golden | | | Carried | ... |
| 14:32 | Pass | Sudbury Wolves | Nolan Hutcheson | 136 | 65 | Landon McCallum | 139 | 69 | Indirect | ... |
| 14:30 | Pass | Sudbury Wolves | Landon McCallum | 149 | 67 | Macauley Carson | 155 | 51 | Direct | ... |
| 14:30 | Shot | Sudbury Wolves | Macauley Carson | 158 | 49 | | | | Snapshot On Net | ... |
| 14:27 | Turnover | Erie Otters | Macauley Carson | 183 | 33 | | | | | ... |

# Expected Possession Value

$$EPV_t = E[X \mid \mathscr{F}_t^Z]$$

Cervone et al (2014)

What is the expected result (X) of the possession, given all of the information ($\mathscr{F}_t^{(Z)}$) up to time t

**Key Insight from Cervone et al (2014)**

1. Want to think about possessions and their outcomes as stochastic processes which evolve over space and time
2. Need to define and model how / when possessions end

# Additional Considerations

- Want to get not just expectations, but a full distribution capturing uncertainty and noise in the process ——> Posterior predictive distribution
- Goals are rare in Hockey —-> We use xG as our targeted outcome (i.e want a posterior predictive distribution for xG given appropriate covariates)

# Model Formulation

Want: Posterior Predictive Distribution for Expected Goals

$$P(xg_T \mid \mathcal{F}_t^{(Z)}) = \int \underbrace{P(xg_T \mid \mathcal{F}_t^{(Z)}, \theta)}_{\text{Likelihood}} \underbrace{P(\theta \mid \mathcal{F}_t^{(Z)})}_{\text{Posterior}} d\theta$$

T: Random variable indicating stopping state
$\mathcal{F}_t^{(Z)}$: Information state at point in space and time

Adapted from Cervone et al

# Model Formulation

Want: Posterior Predictive Distribution for Expected Goals

$$P(xg_T \mid \mathcal{F}_t^{(Z)}) = \int \underbrace{P(xg_T \mid \mathcal{F}_t^{(Z)}, \theta)}_{\text{Likelihood}} \underbrace{P(\theta \mid \mathcal{F}_t^{(Z)})}_{\text{Posterior}} d\theta$$

T: Random variable indicating stopping state
$\mathcal{F}_t^{(Z)}$: Information state at point in space and time

$$P(xg_T \mid \mathcal{F}_t^{(Z)}, \theta) = \int_{\Omega} P(xg_{T(\omega)} \mid \mathcal{F}_t^{(Z)}, \theta) P(d\omega \mid \mathcal{F}_t^{(Z)}, \theta)$$

Integrate over possible Possession paths

$\Omega$ : Total possession path
$\omega$: a particular possession

Adapted from Cervone et al

# Distribution for Possessions

To estimate distribution of possessions: $P(d\omega \mid \mathscr{F}_t^{(Z)}, \theta)$

Use forward simulation

# Distribution for Possessions

To estimate distribution of possessions: $P(d\omega \mid \mathcal{F}_t^{(Z)}, \theta)$

Use forward simulation

Need the following model types for each t and available over the space-time region of interest

$$P(xG_{t+1}, S_{t+1}, Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta) = P(xG_{t+1} \mid S_{t+1}, Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta) P(S_{t+1} \mid Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta) P(Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta)$$

# Distribution for Possessions

To estimate distribution of possessions: $P(d\omega \mid \mathcal{F}_t^{(Z)}, \theta)$

Use forward simulation

Need the following model types for each t and available over the space-time region of interest

$$P(xG_{t+1}, S_{t+1}, Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta) = \underbrace{P(xG_{t+1} \mid S_{t+1}, Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta)}_{\text{Value}} P(S_{t+1} \mid Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta) P(Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta)$$

# Distribution for Possessions

To estimate distribution of possessions: $P(d\omega \mid \mathcal{F}_t^{(Z)}, \theta)$

Use forward simulation

Need the following model types for each t and available over the space-time region of interest

$$P(xG_{t+1}, S_{t+1}, Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta) = P(xG_{t+1} \mid S_{t+1}, Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta)\underbrace{P(S_{t+1} \mid Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta)}_{\text{State / Action}}P(Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta)$$

# Distribution for Possessions

To estimate distribution of possessions: $P(d\omega \mid \mathcal{F}_t^{(Z)}, \theta)$

Use forward simulation

Need the following model types for each t and available over the space-time region of interest

$$P(xG_{t+1}, S_{t+1}, Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta) = P(xG_{t+1} \mid S_{t+1}, Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta) P(S_{t+1} \mid Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta) P(\underbrace{Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta}_{\text{Space-Time}})$$

# Distribution for Possessions

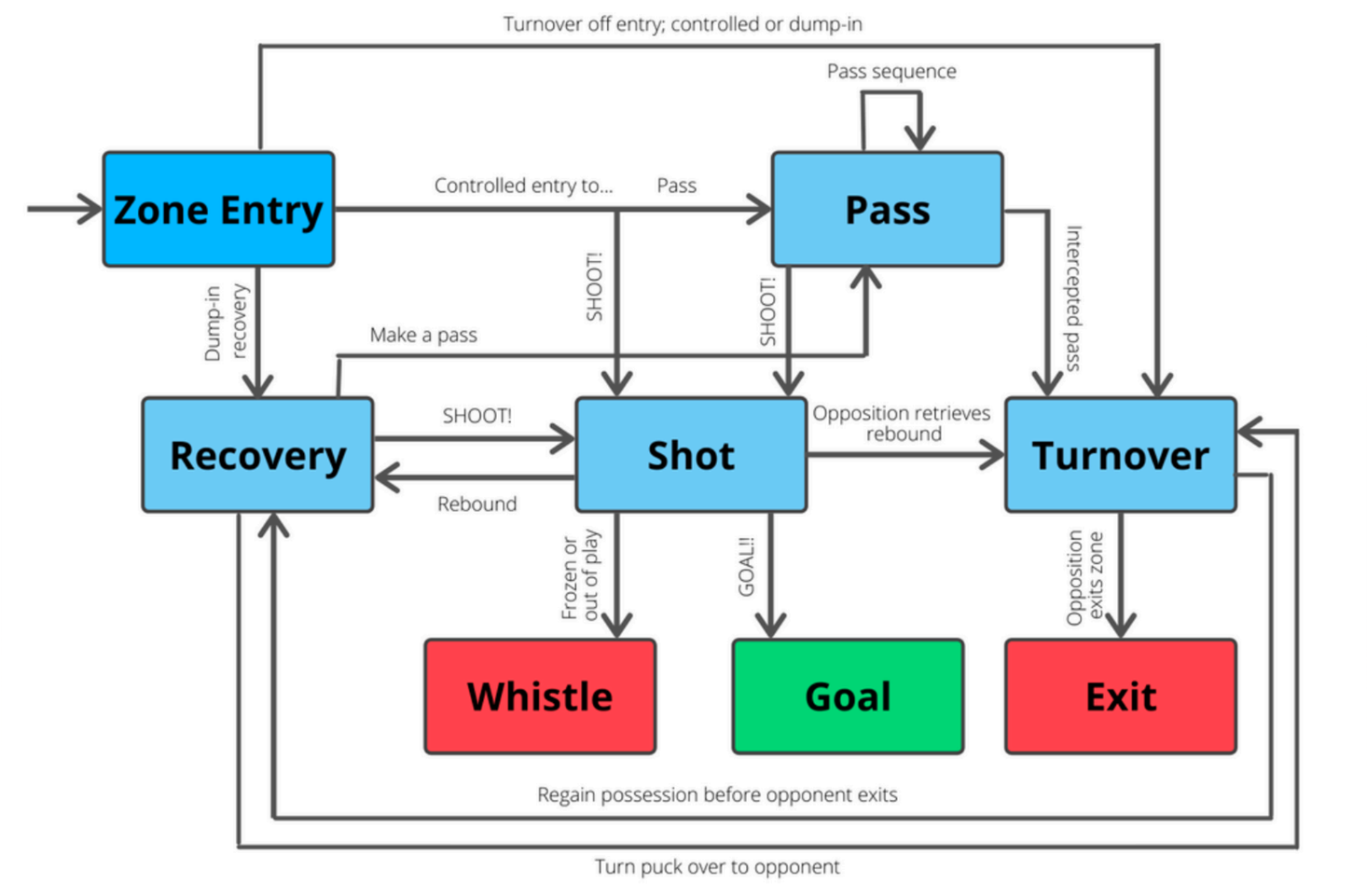
To estimate distribution of possessions: $P(d\omega \mid \mathcal{F}_t^{(Z)}, \theta)$

Use forward simulation

Need the following model types for each t and available over the space-time region of interest

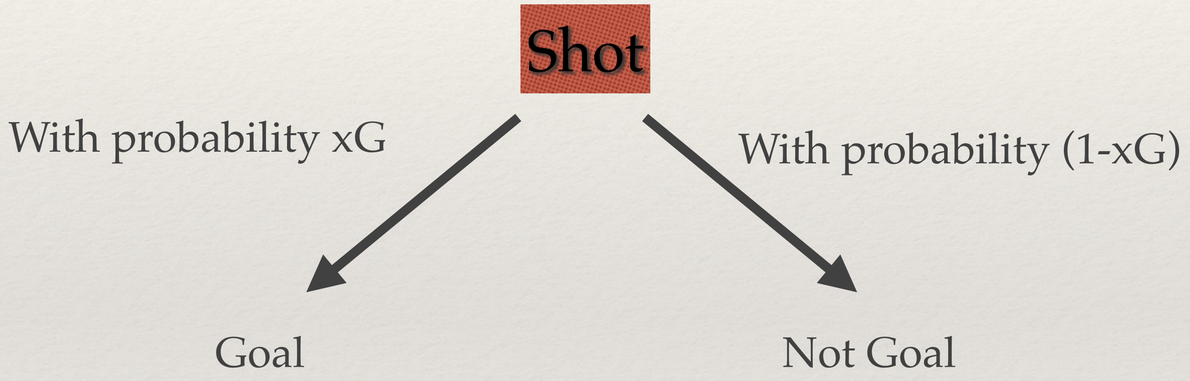
$$P(xG_{t+1}, S_{t+1}, Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta) = \underbrace{P(xG_{t+1} \mid S_{t+1}, Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta)}_{\text{Value}} \underbrace{P(S_{t+1} \mid Z_{t+1}, \mathcal{F}_t^{(Z)}, \theta)}_{\text{State / Action}} \underbrace{P(Z_{t+1} \mid \mathcal{F}_t^{(Z)}, \theta)}_{\text{Space-Time}}$$

# State/Action Models



# State/Action Models

Goals are rare, we don't treat them as an absorbing state, but branching state!




Each time a shot is reached, the xG is collect of the shot is collected for the current possession value

But we continue the simulation as though a goal was not scored and weight future xG on this path with (1-xG)

$$xG(\omega) = \sum_{i=1}^{k} xG_i * (\prod_{j=1}^{i-1} (1 - xG_j))$$

## Action-Transition Models

- Controlled Entry to Pass
- Controlled Entry to Shot
- Controlled Entry to Turnover
- Dump-in Entry to Recovery
- Dump-in Entry to Turnover
- Pass to Pass
- Pass to Recovery
- Pass to Turnover
- Recovery to Pass
- Recovery to Shot
- Recovery to Turnover
- Shot to Recovery
- Shot to Turnover
- Shot to Whistle
- Turnover to Exit
- Turnover to Recovery

16+ Action models needed to represent the Markov diagram

Since we have added a few to separate shots on net and missed shots
Since missed shots and shots on goal have different values and different
Properties in terms of where we expect the next events to occur.

**1. Actual Spatial Density (Model Input)**

Recovery to Pass

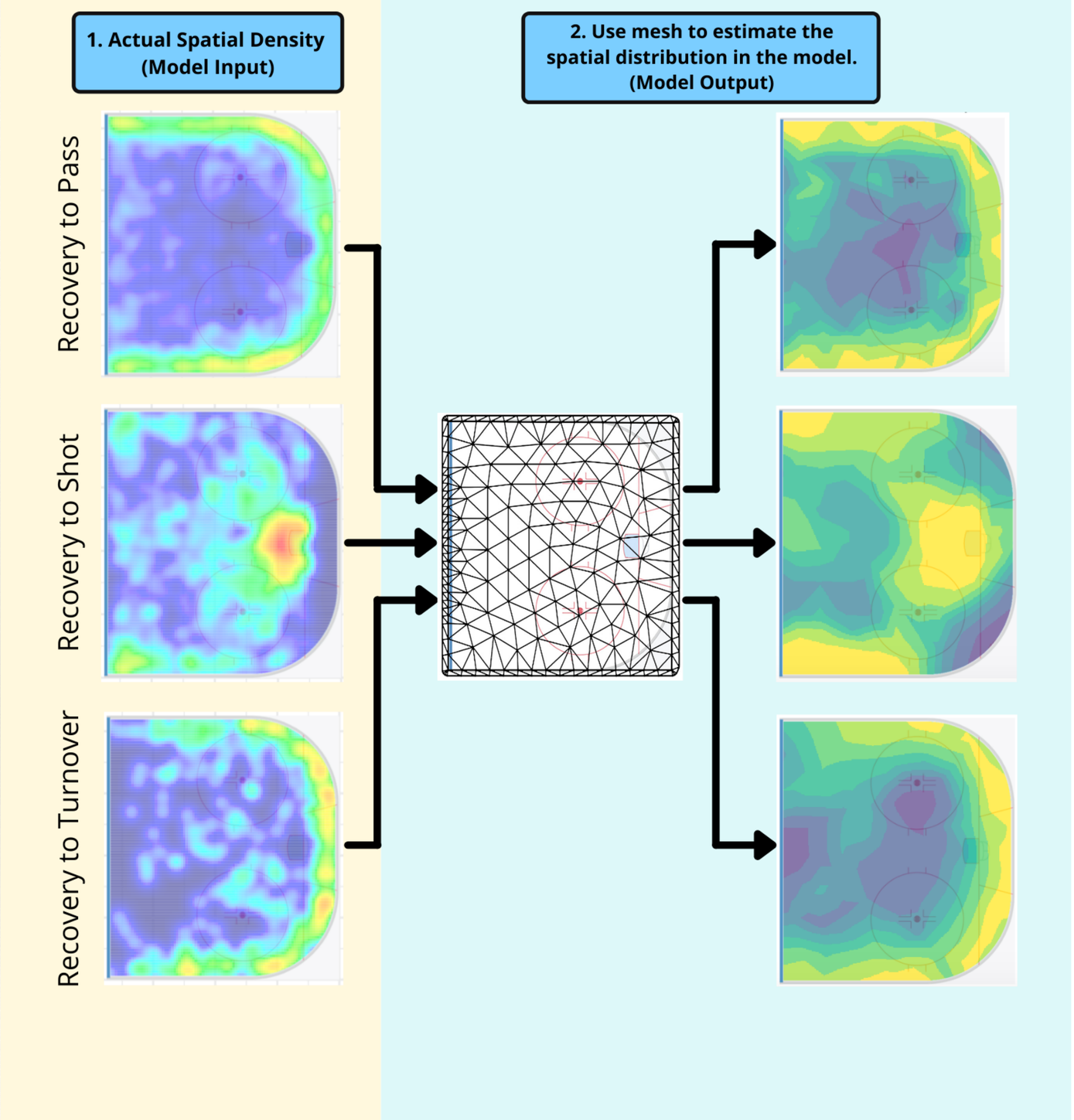


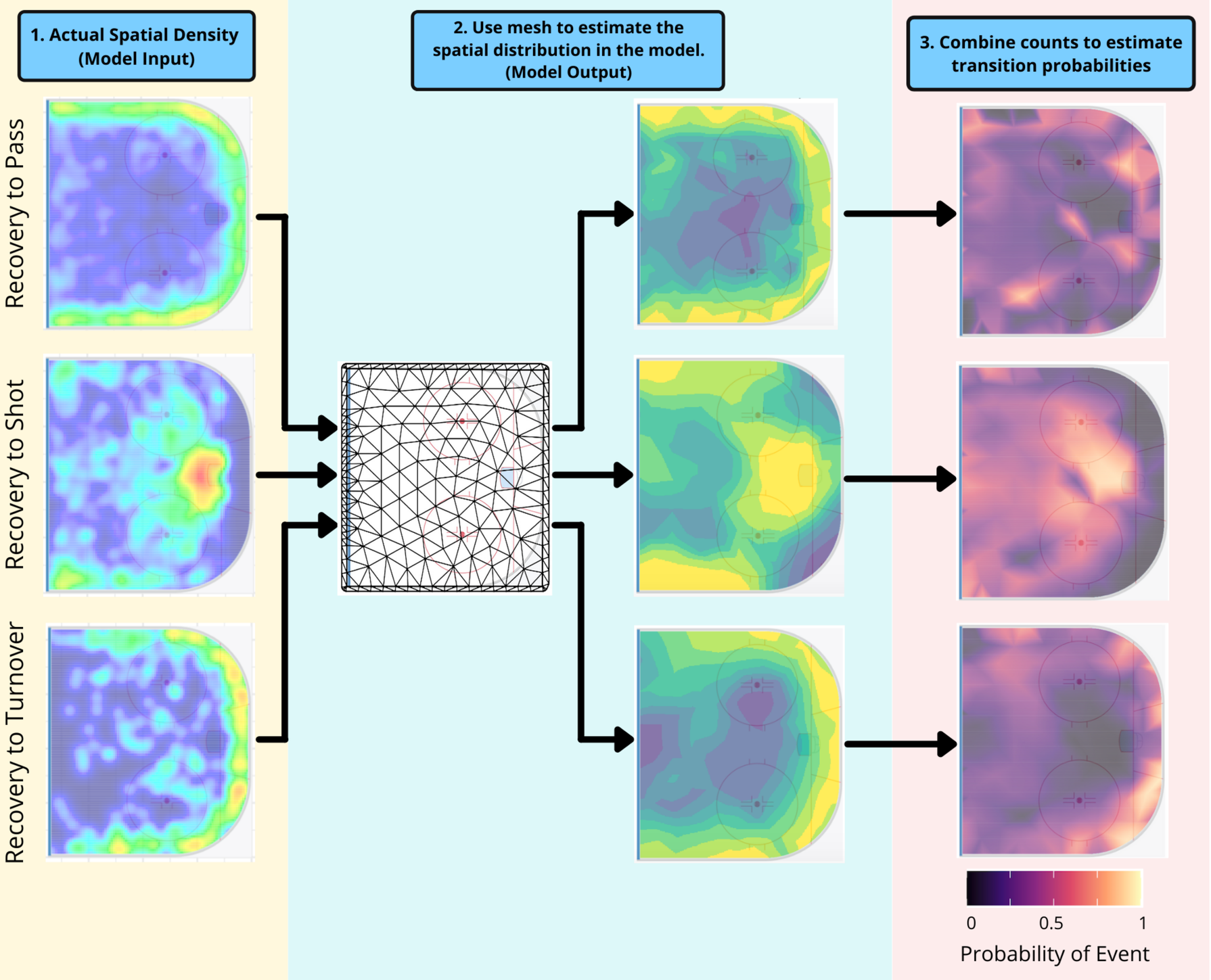
Recovery to Shot



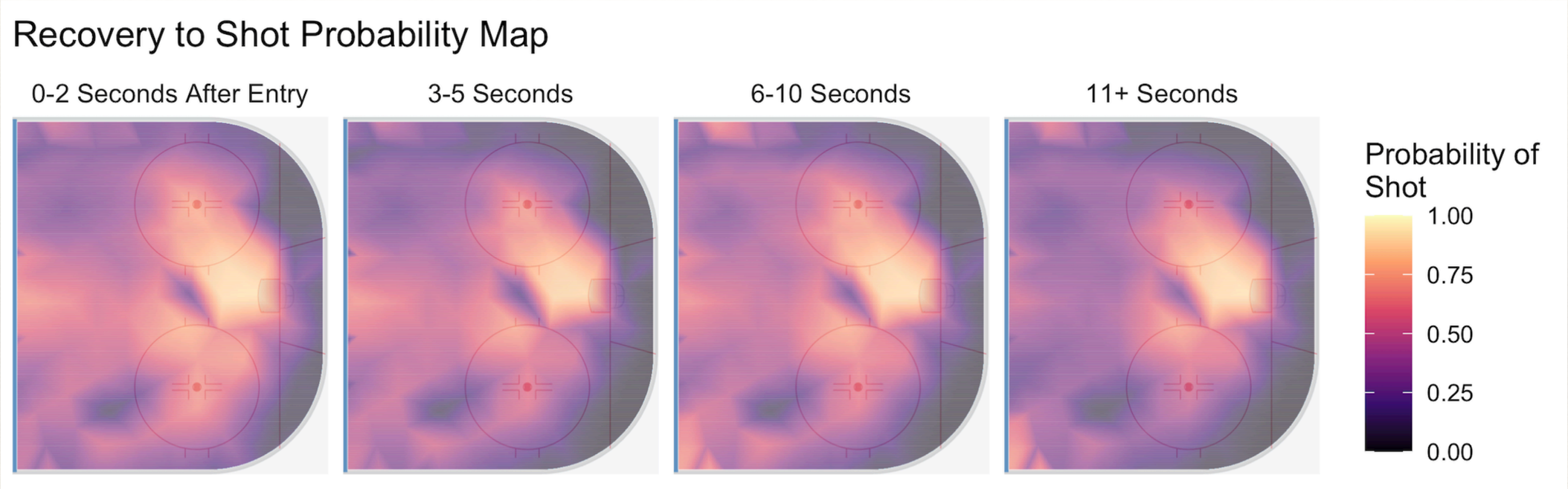
Recovery to Turnover

1. Actual Spatial Density (Model Input)
2. Use mesh to estimate the spatial distribution in the model. (Model Output)
Recovery to Pass
Recovery to Shot
Recovery to Turnover


1. Actual Spatial Density (Model Input)
2. Use mesh to estimate the spatial distribution in the model. (Model Output)
Recovery to Pass
Recovery to Shot
Recovery to Turnover


1. Actual Spatial Density (Model Input)
2. Use mesh to estimate the spatial distribution in the model. (Model Output)
3. Combine counts to estimate transition probabilities
Recovery to Pass
Recovery to Shot
Recovery to Turnover
0
0.5
1
Probability of Event

# Time Component




AR(1) Process

# Time Component

**Rush State** ⟷ **Established Offense**

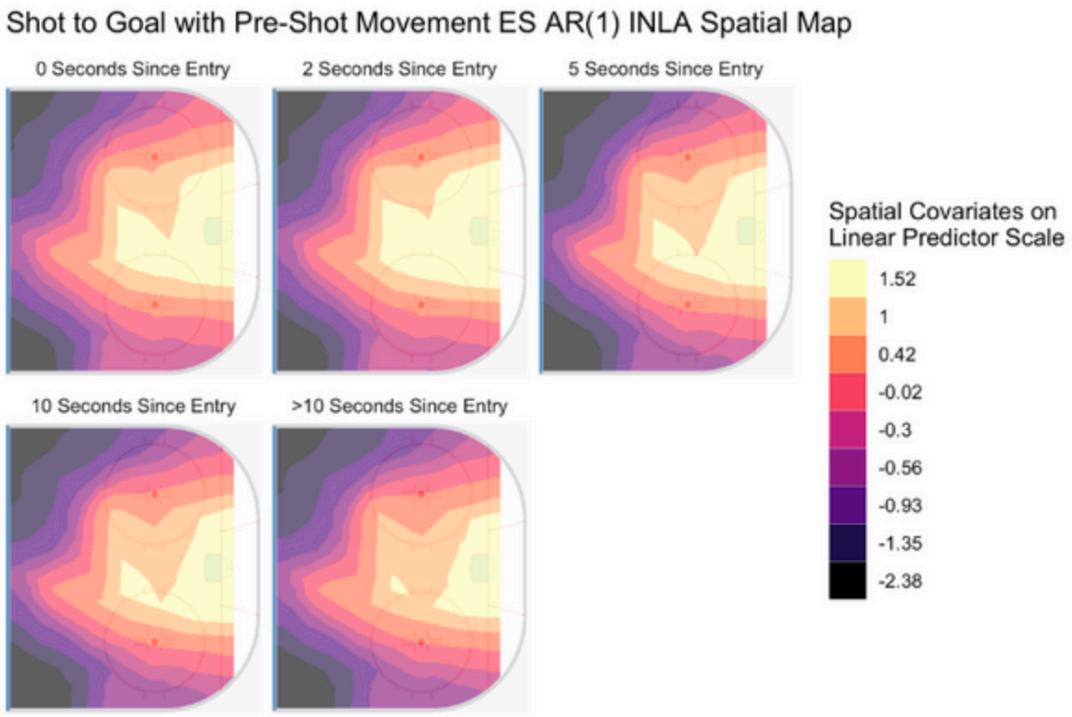
Looser structures, odd-man situations
Valuable space easier to gain and occupy

5 on 5 with very structured roles (especially defensively)

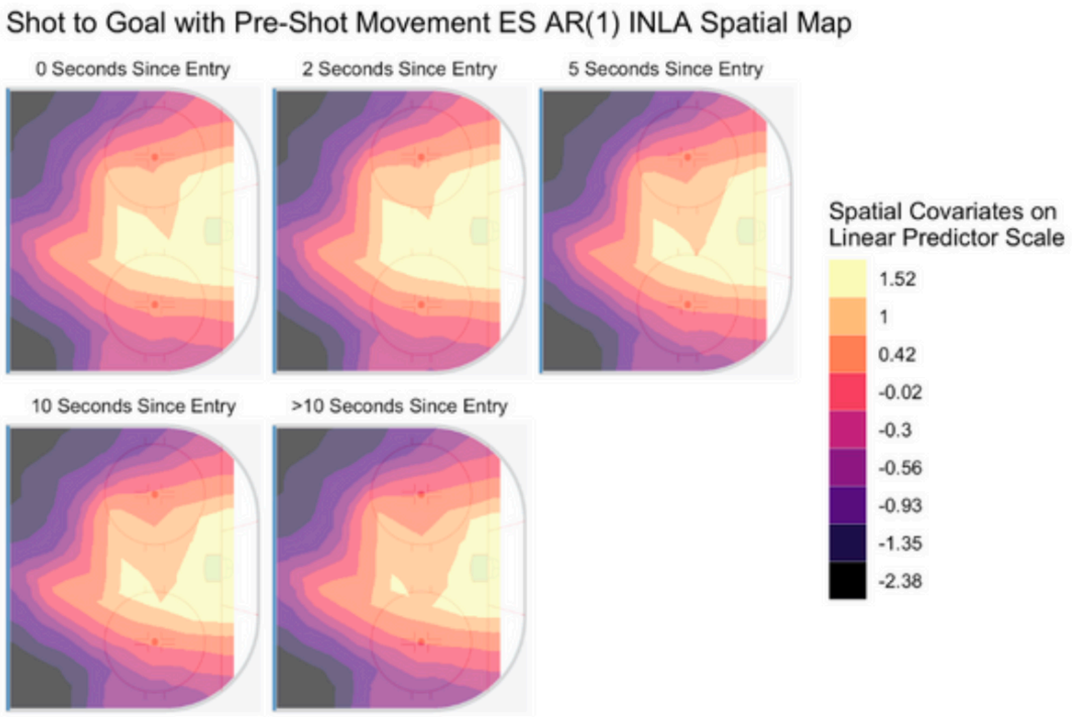
## Value Models

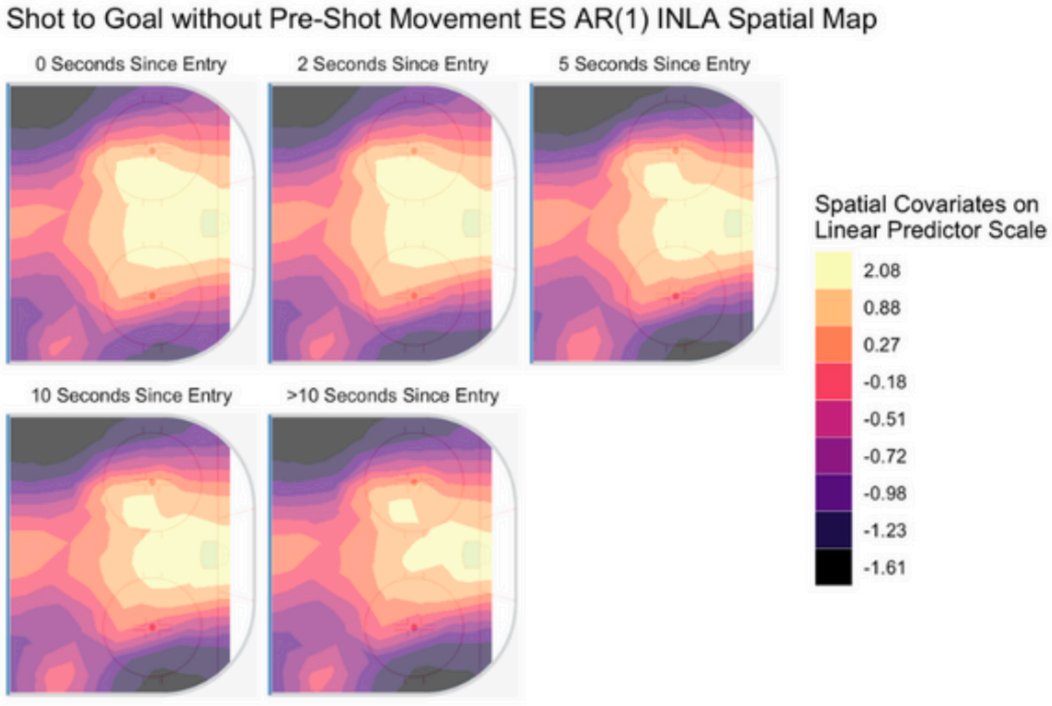
- Expected Goals with Preceding Pass
- Expected Goals without Preceding Pass

# Expected Goals Models



# Expected Goals Models



## Space, Time and Context Models

- x-Location
- y-Location
- Time to Next Event

- Pass Target x-Location
- Pass Target y-Location
- Direct/Indiret Pass

- Traffic with Preceding Pass
- Traffic without Preceding Pass
- One-Timer
- Deflected Shot
- Missed Shot
- Blocked Shot

- Zone Entry Decision (Controlled or Dumped)

## Space, Time and Context Models

- **x-Location**
- **y-Location**
- Time to Next Event

- Pass Target x-Location
- Pass Target y-Location
- Direct/Indiret Pass

- Traffic with Preceding Pass
- Traffic without Preceding Pass
- One-Timer
- Deflected Shot
- Missed Shot
- Blocked Shot

- Zone Entry Decision (Controlled or Dumped)

# Location Models most important and difficult

## Space, Time and Context Models

- **x-Location**
- **y-Location**
- Time to Next Event

- Pass Target x-Location
- Pass Target y-Location
- Direct/Indiret Pass

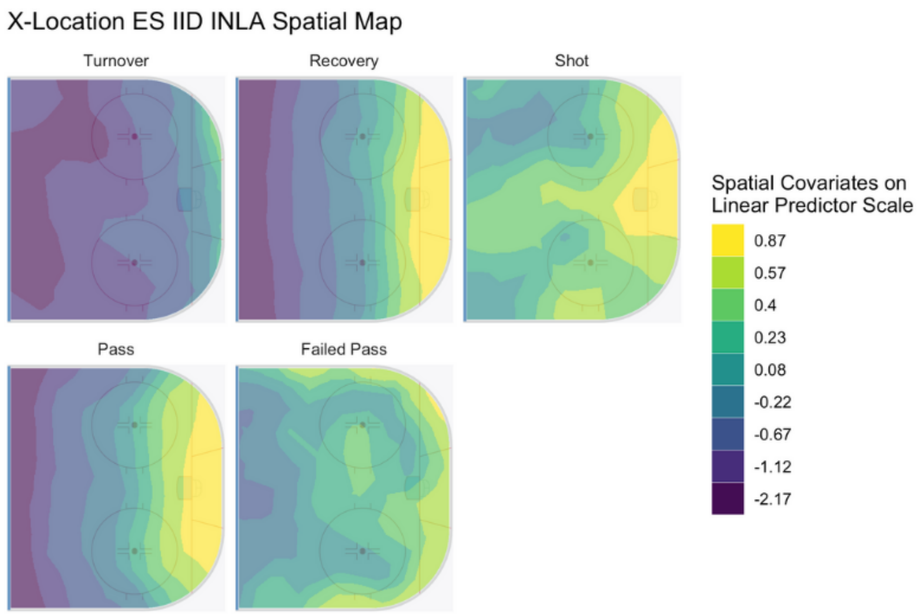
- Traffic with Preceding Pass
- Traffic without Preceding Pass
- One-Timer
- Deflected Shot
- Missed Shot
- Blocked Shot

- Zone Entry Decision (Controlled or Dumped)

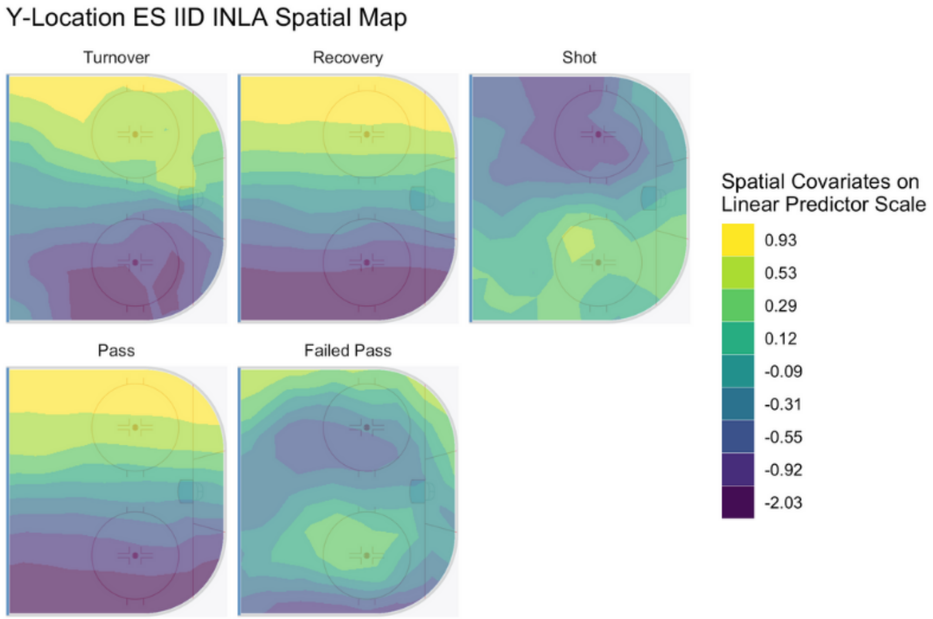
# Location Models most important and difficult

- Simulated paths depend heavily on locations because (X,Y) location is most important covariate for nearly all other models
- Want to fit X,Y distribution jointly
- The dimensions of the rink are irregular. Things like borders and location of the net introduce irregularities

# LOCATION MODELS




***Spatial effect of our model to predict the x-coordinate of the next event (yellow implies closer to goal line, navy implies closer to blue line).***




***Spatial effect of our model to predict the y-coordinate of the next event (yellow implies closer to goal line, navy implies closer to blue line).***

## Space, Time and Context Models

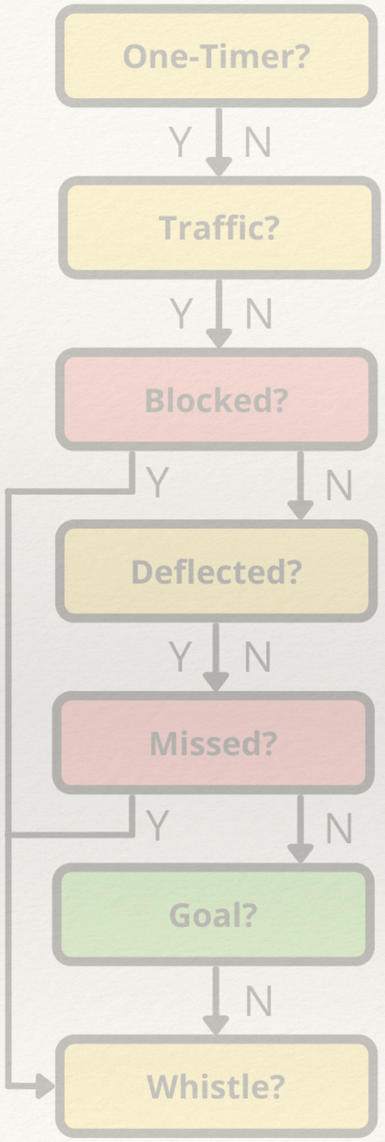
- x-Location
- y-Location
- Time to Next Event

- Pass Target x-Location
- Pass Target y-Location
- Direct/Indiret Pass

- **Traffic**
- **One-Timer**
- **Deflected Shot**
- **Missed Shot**
- **Blocked Shot**

- Zone Entry Decision (Controlled or Dumped)

| Model | Covariates *All models include Time since last pass/shot/entry/faceoff; Behind Net?* |
|---|---|
| **One-Timer?** | **Basics** |
| **Traffic?** | Basics; **One-Timer?** |
| **Blocked?** | Basics; One-Timer?; **Traffic?** |
| **Deflected?** | Basics; One-Timer?; Traffic? |
| **Missed?** | Basics; One-Timer?; Traffic?; **Deflected?** |
| **Goal?** | Basics; One-Timer?; Traffic?; Deflected? |
| **Whistle?** | Basics; One-Timer?; Traffic?; Deflected?; **Missed?; Blocked?** |



## Space, Time and Context Models

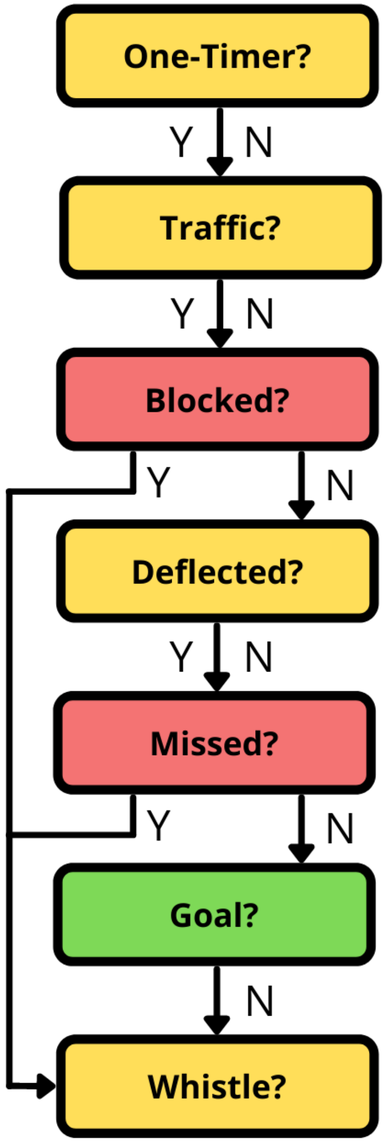
- x-Location
- y-Location
- Time to Next Event

- Pass Target x-Location
- Pass Target y-Location
- Direct/Indiret Pass

- **Traffic**
- **One-Timer**
- **Deflected Shot**
- **Missed Shot**
- **Blocked Shot**

- Zone Entry Decision (Controlled or Dumped)

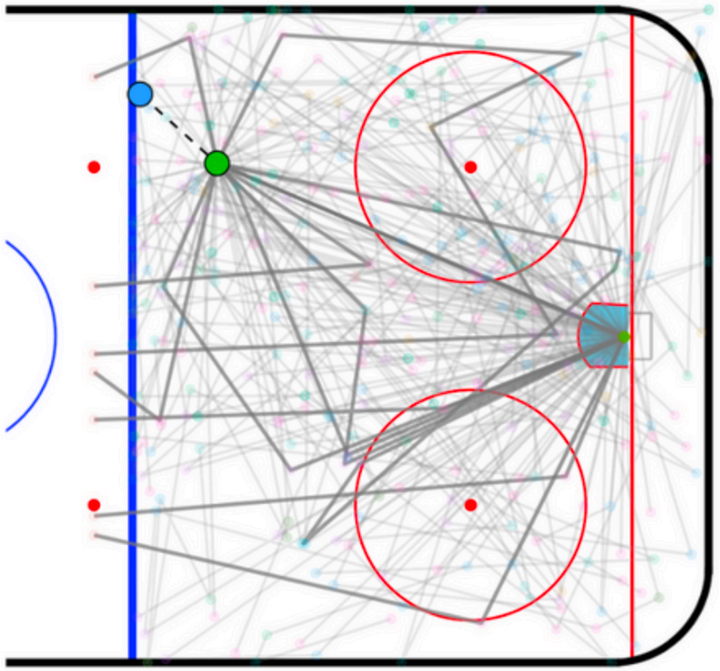
| Model | Covariates *All models include Time since last pass/shot/entry/faceoff; Behind Net?* |
|---|---|
| **One-Timer?** | **Basics** |
| **Traffic?** | Basics; **One-Timer?** |
| **Blocked?** | Basics; One-Timer?; **Traffic?** |
| **Deflected?** | Basics; One-Timer?; Traffic? |
| **Missed?** | Basics; One-Timer?; Traffic?; **Deflected?** |
| **Goal?** | Basics; One-Timer?; Traffic?; Deflected? |
| **Whistle?** | Basics; One-Timer?; Traffic?; Deflected?; **Missed?; Blocked?** |

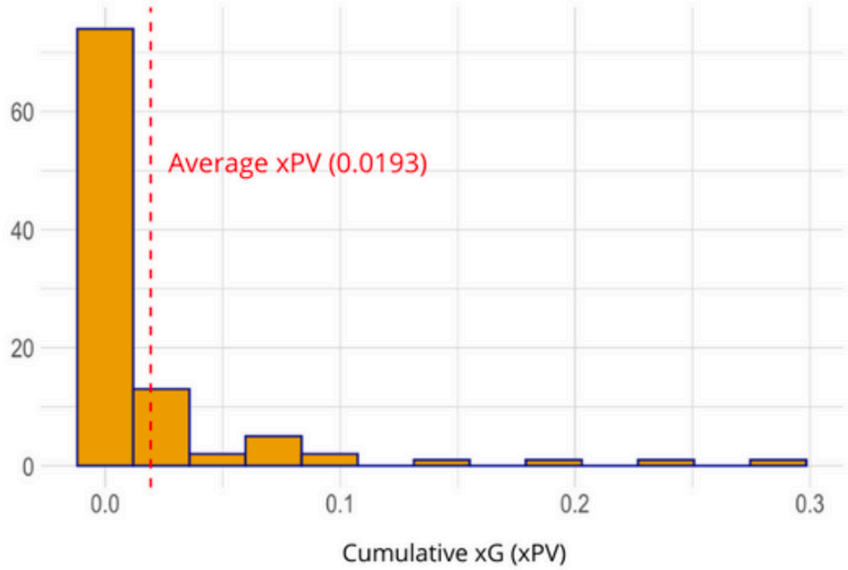


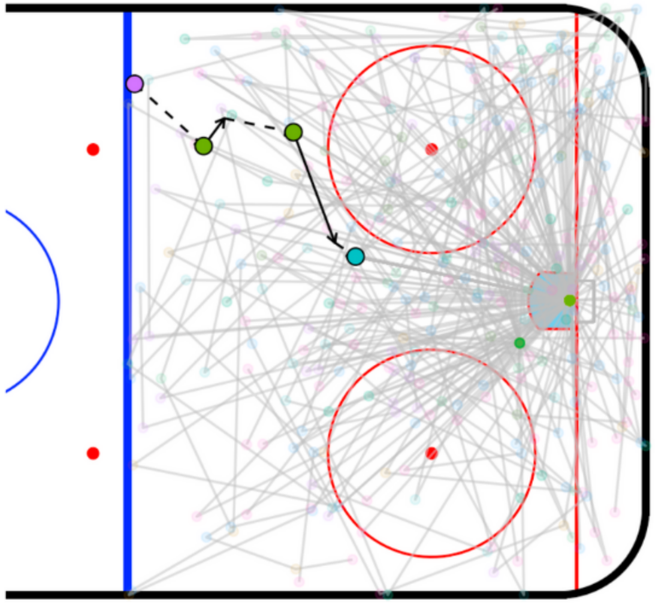
# HOW CAN WE ESTIMATE VALUE AT ANY MOMENT WITH THIS MODEL?

**1. Run simulations from current point in space and time**

**2. Get distribution of cumulative xG from the simulations to get Expected Possession Value (xPV)**

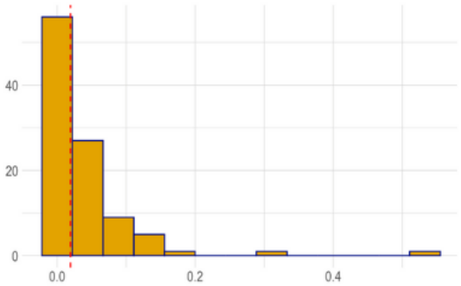




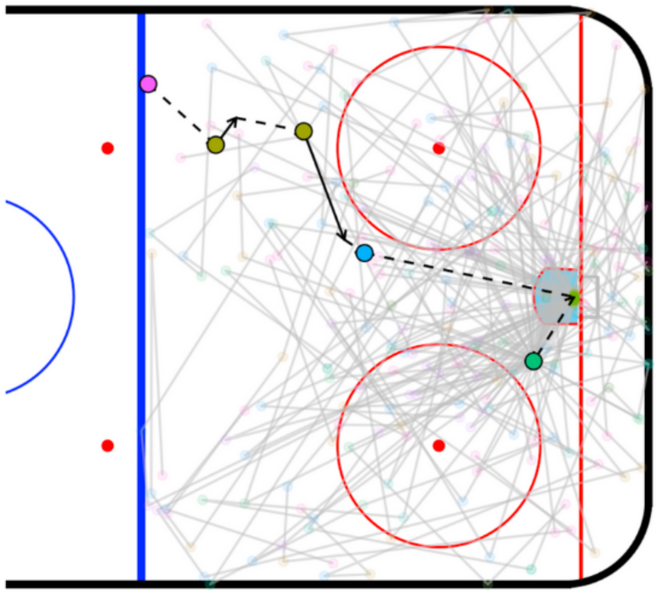


## xPV Before Shot

*Event n*

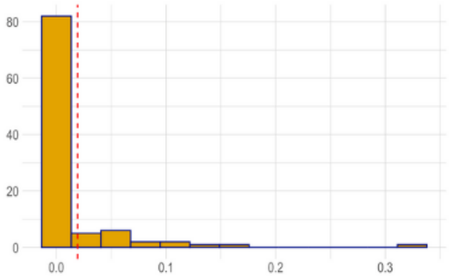


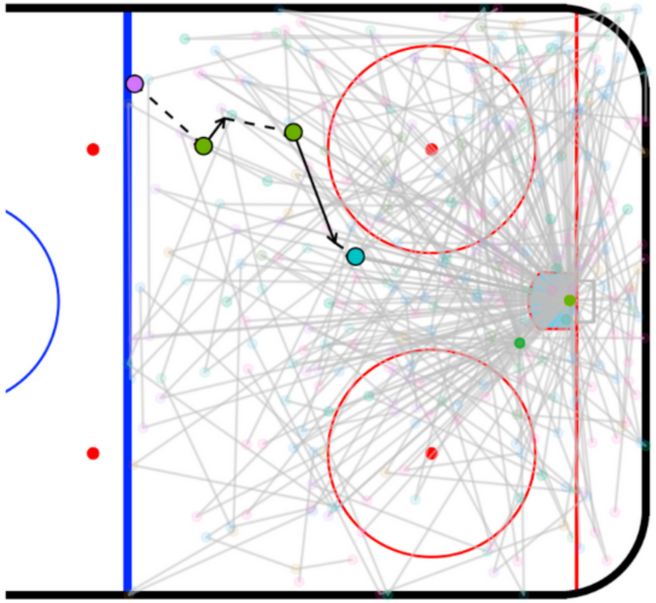

xG From Shot = 0.078



## xPV After Shot

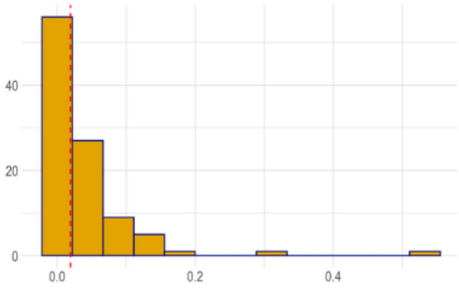
*Event n+1*



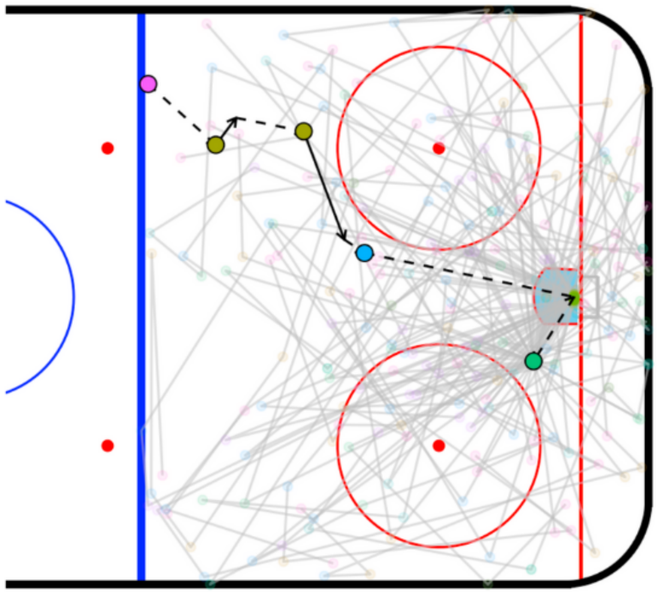


xPV Before Shot

*Event n*

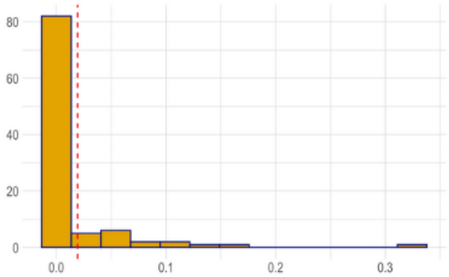



xG From Shot = 0.078

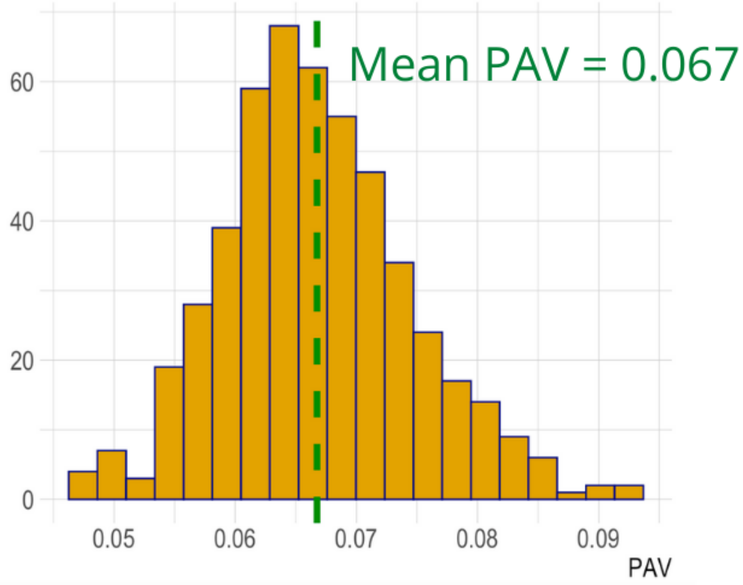


xPV After Shot

*Event n+1*




# Possession Added Value (PAV)




=

xG + xPV(n+1)*(1-xG) - xPV(n)

Probability of scoring

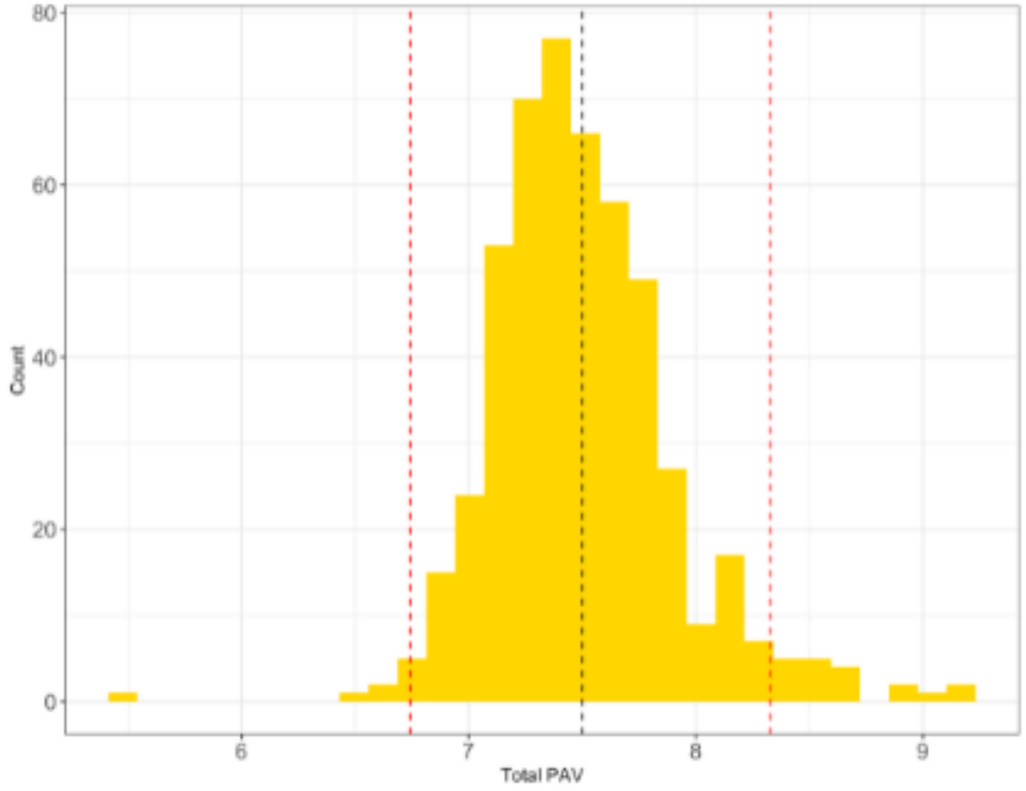
Value after taking action (given we don't score)

Value of space and time before action

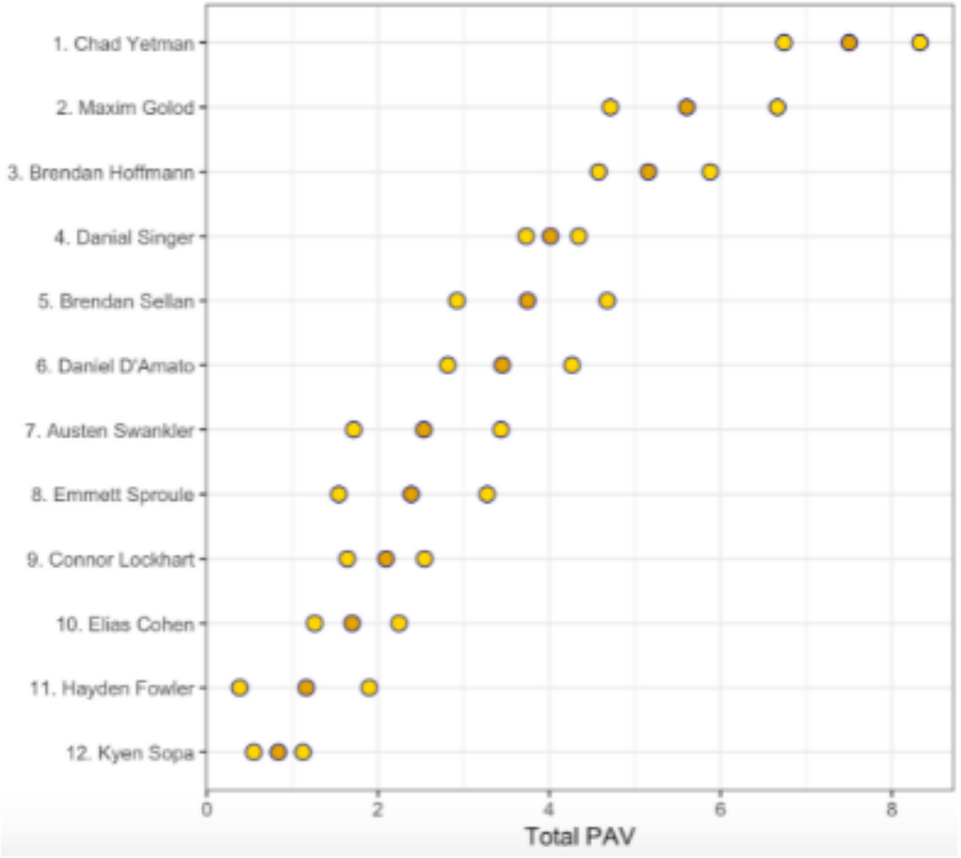
*No shot?* xPV(n+1) - xPV(n)

# ESTIMATING UNCERTAINTY

**500 Samples of Total PAV for Chad Yetman with 93% Credible Interval (Red)**




**Top 12 Otters in Total PAV with 93% Credible Interval**



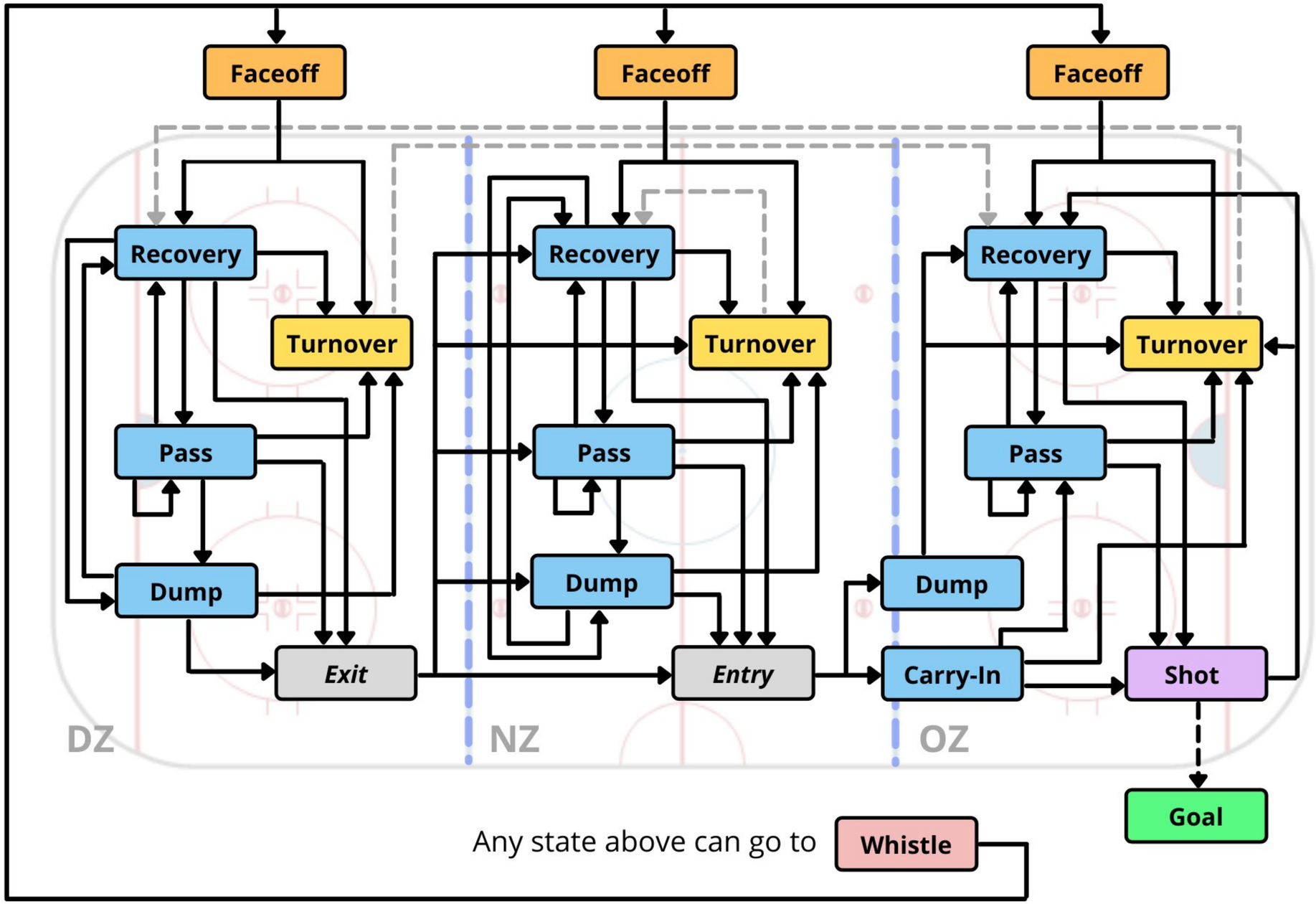
# Interpreting Downstream Metrics

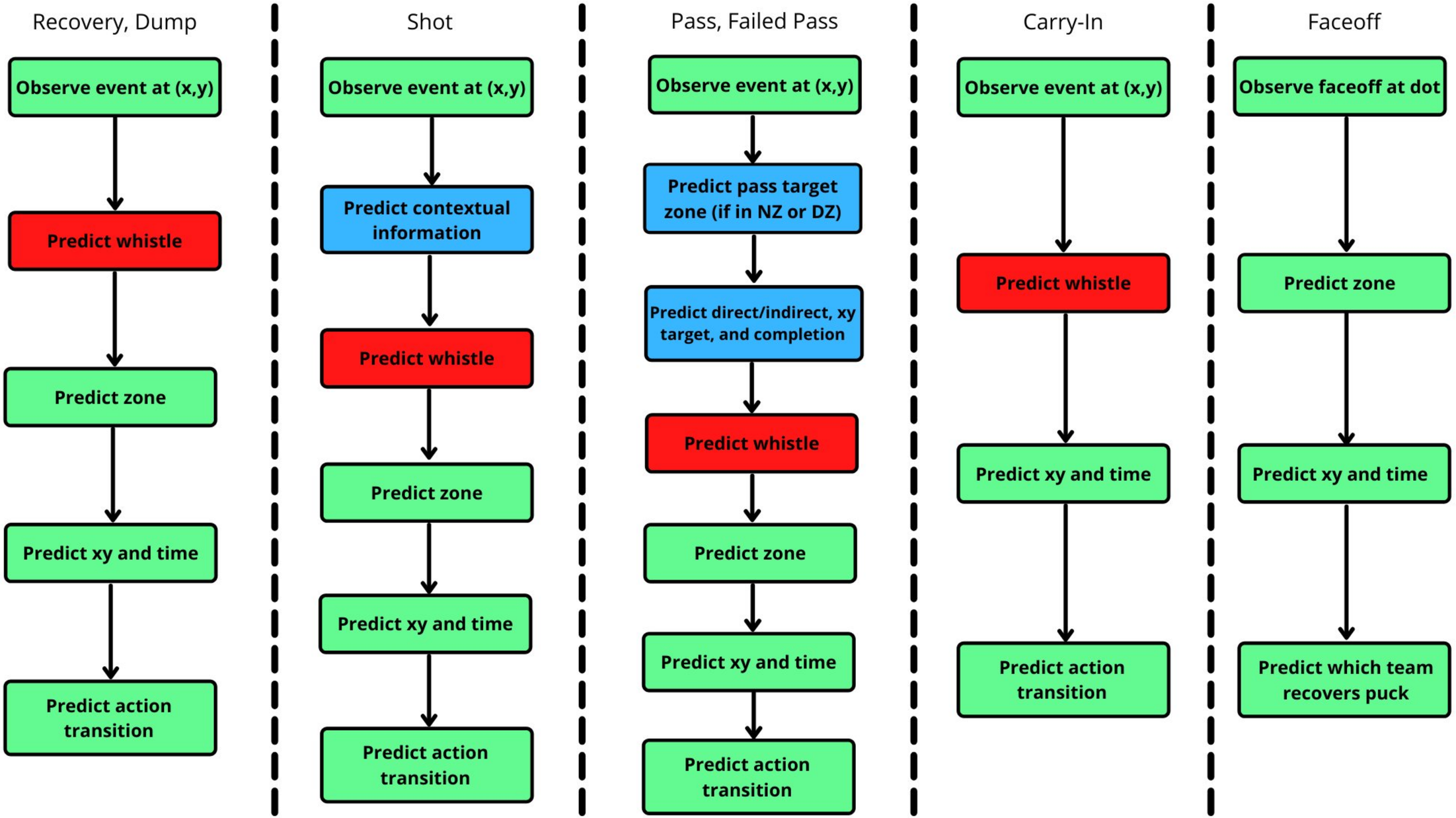
Need to be careful! The largest confounder is teammate and defender location.
We only capture this partially through time since entry and integrating over this uncertainty.

Not causal and likely to be fooled!

# Full Ice Extension

- Where do possessions end?
- Need to model the zone location and actions available in each zone
- Outcome needs to be with respect to xG for and against
- Branching state (Need to incorporate value due to whistle location)
- Computational difficulties


Faceoff
Faceoff
Faceoff
Recovery
Turnover
Pass
Dump
Exit
Recovery
Turnover
Pass
Dump
Entry
Recovery
Turnover
Pass
Dump
Carry-In
Shot
DZ
NZ
OZ
Goal
Any state above can go to
Whistle


Recovery, Dump
Observe event at (x,y)
Predict whistle
Predict zone
Predict xy and time
Predict action transition
Shot
Observe event at (x,y)
Predict contextual information
Predict whistle
Predict zone
Predict xy and time
Predict action transition
Pass, Failed Pass
Observe event at (x,y)
Predict pass target zone (if in NZ or DZ)
Predict direct/indirect, xy target, and completion
Predict whistle
Predict zone
Predict xy and time
Predict action transition
Carry-In
Observe event at (x,y)
Predict whistle
Predict xy and time
Predict action transition
Faceoff
Observe faceoff at dot
Predict zone
Predict xy and time
Predict which team recovers puck

# Extensions

- Finish the full-ice, more work to be done on (sub)-model validation and optimization
- Could extend reasonably easily to models which condition on the location of all players on the ice
- This would open the door for interesting simulations to better understand value creating and the spatial properties and tactics of hockey
- Similarly, this would partially solve our selection bias issues. Allow us to evaluate player decision making directly
- But need to think carefully about how we attribute value to the players

# REFERENCES


Cervone, D., D'Amour, A., Bornn, L. & Goldsberry, K. (2016). **A Multiresolution Stochastic Process Model for Predicting Basketball Possession Outcomes**, Journal of the American Statistical Association, 111:514, 585-599, DOI: 10.1080/01621459.2016.1141685

Fernandez, J., Bornn, L., & Cervone, D. (2020). **A framework for the fine-grained evaluation of the instantaneous expected value of soccer possessions**. arXiv preprint arXiv:2011.09426.https://arxiv.org/abs/2011.09426

Singh, K. (2019). **Introducing Expected Threat (xT)**. Retrieved from https://karun.in/blog/expected-threat.html

Yu, D., Piggott, P., Jung, C. & Forstner, S. (2020). **A comprehensive analysis of pass difficulty, value and tendencies in hockey**. ISOLHAC. Retrieved from https://www.youtube.com/watch?v=Q-kWb6Vshmo&t=2313s

Chatel, T. (2020). **Introducing Offensive Sequences and The Hockey Decision Tree**. Retrieved from https://hockey-graphs.com/2020/03/26/introducing-offensive-sequences-and-the-hockey-decision-tree/

Lee, J.Y.L., Green, P.J., Ryan, L.M. (2017). **On "Poisson Trick" and its Extensions for Fitting Multinomial Regression Models**. arXiv preprint arXiv:1707.08538

Simpson, D.P., Rue, H., Martins, T.G., Riebler, A. & Sorbye, S.H. (2014). **Penalising model component complexity: A principled, practical approach to constructing priors**. arXiv preprint arXiv: 1403.4630

Rue, H., Martino, S. & Chopin, N. (2009). **Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations**. Journal of the Royal Statistical Society: Series B (Statistical Methodology),71: 319-392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

# References Continued

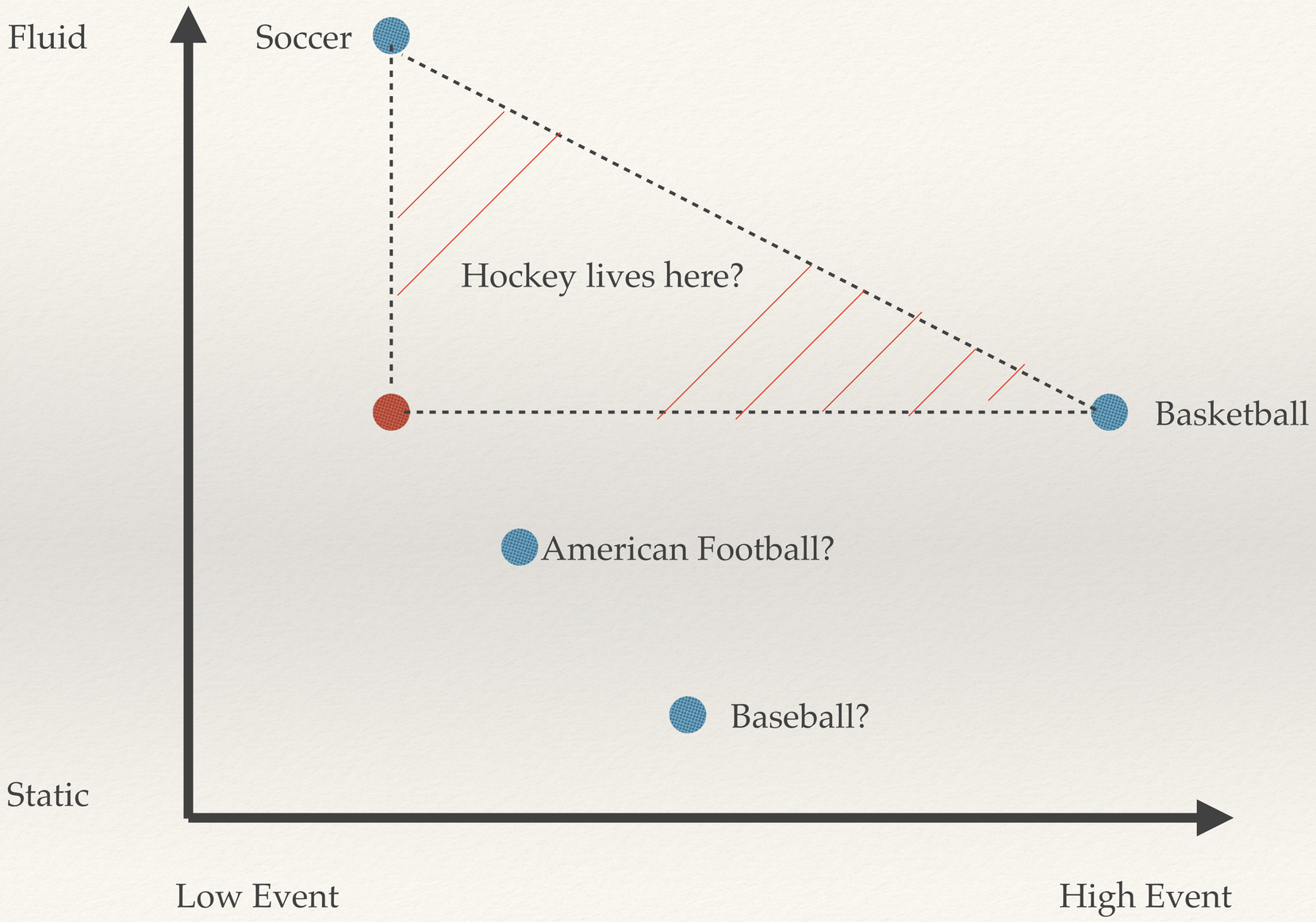
Krainski, E., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., … & Rue, H. (2018). *Advanced spatial modeling with stochastic partial differential equations using R and INLA*. Chapman and Hall/CRC. https://becarioprecario.bitbucket.io/spde-gitbook/

Gómez-Rubio, V. (2020). *Bayesian inference with INLA*. CRC Press. https://becarioprecario.bitbucket.io/inla-gitbook/index.html

Lindgren, Finn, Håvard Rue, and Johan Lindström. "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.4 (2011): 423-498.

# Extra Slides

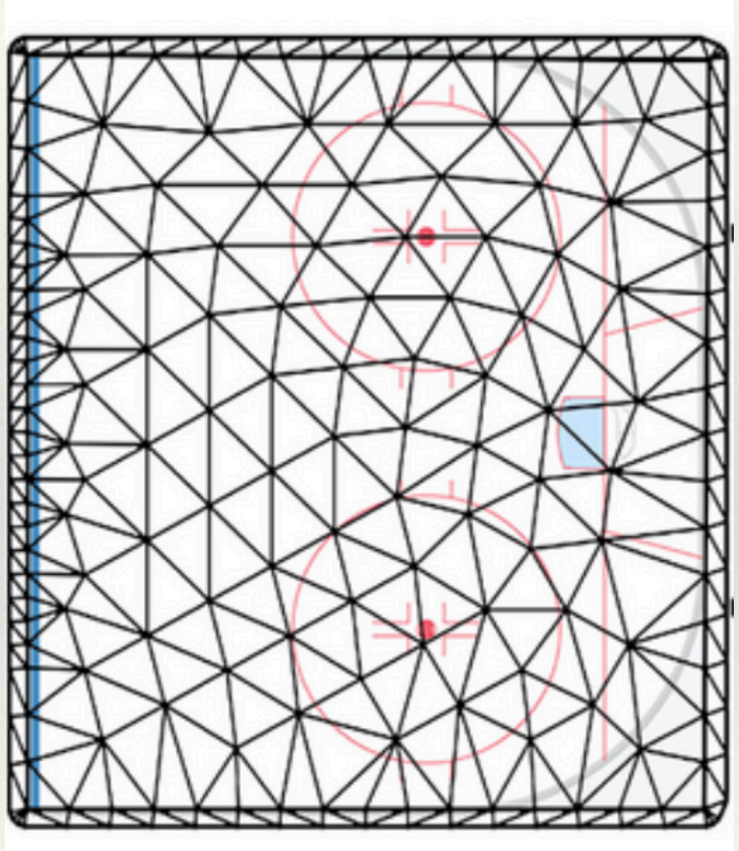
# Sport Specific Considerations



# INLA (Integrated Nested Laplace Approximation)

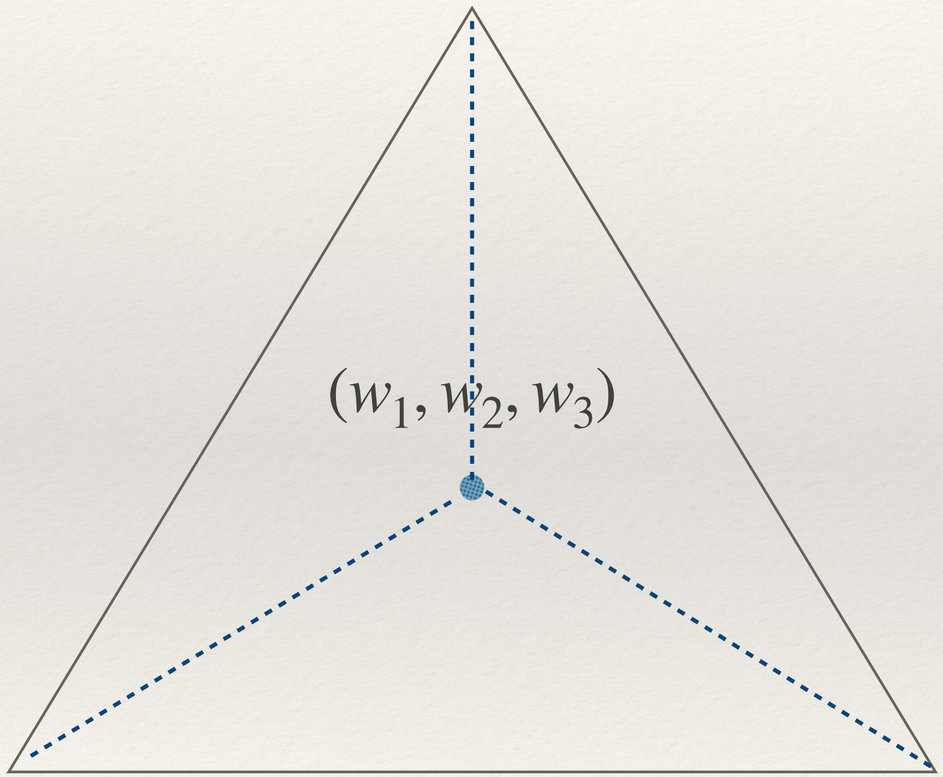
- Very fast way to do approximate Bayesian Inference (R-INLA package)
- Works for a restricted class of models called latent gaussian models (Includes GLMs, GLMMs, GAMs, and some space-time models like those used here)
- Broadly speaking works because the "difficult" latent part (spatial effects for example given hyper-parameters) is already gaussian. This allows then to target the marginal posteriors in a clever way, integrating out some of the difficultly leaving terms which can be well-approximated by (Laplace) approximations

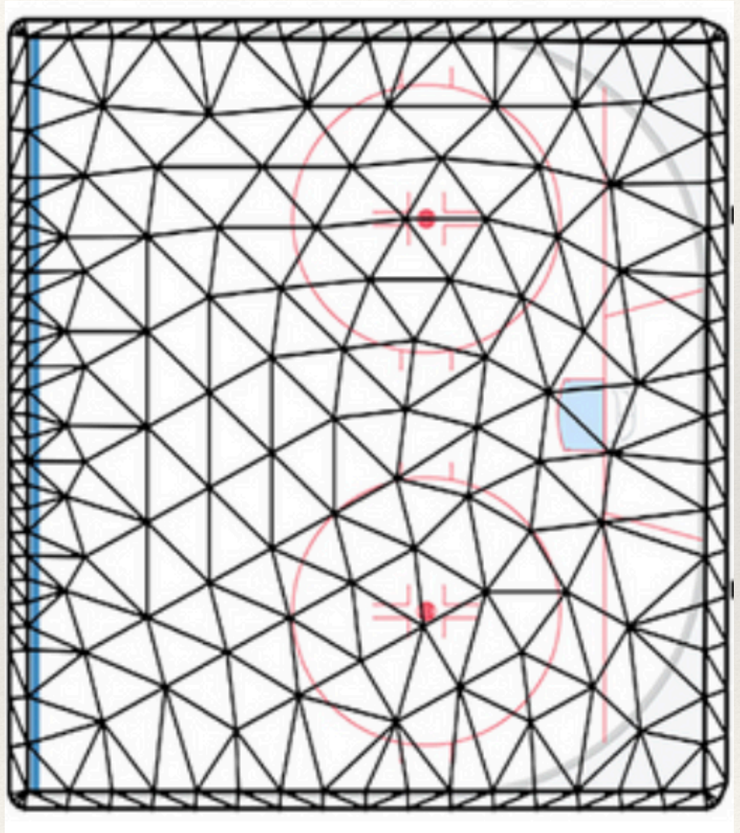
# MESH



# MESH

Each point on the ice weighted by the nodes in the triangle they live in using barycentric weights

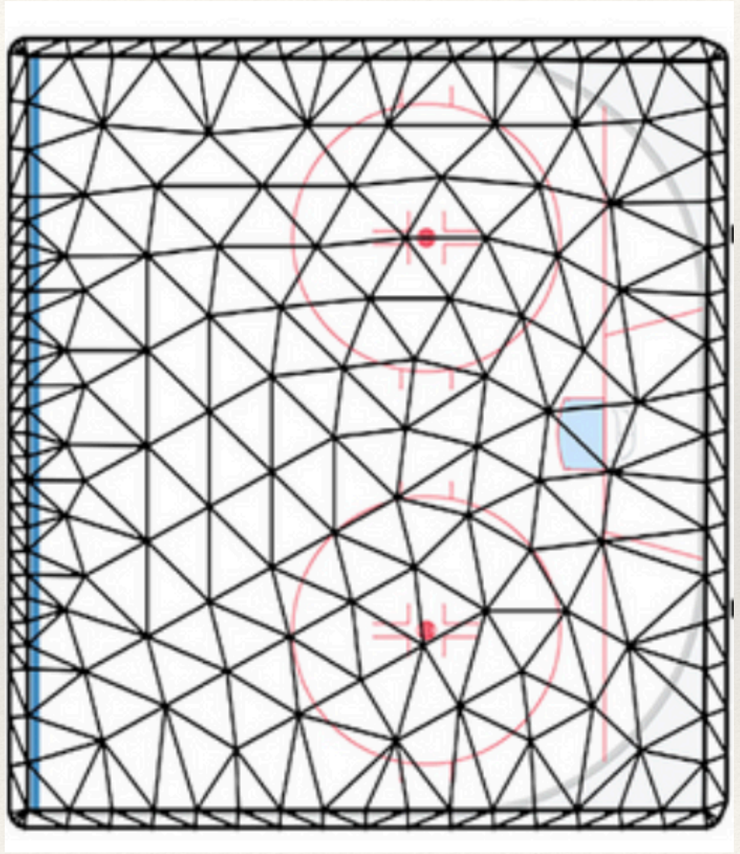


# MESH

In general want smaller triangles where we expect larger change in the underlying process AND we have enough data to estimate that change

We left really exploring best mesh design for these models for future work, likely lots of interesting things left to do.



# Matérn Covariance function

- Want to share information across space and time in proportion to how "close" they are
- Essentially we use a gaussian process (with matérn kernel/covariance function which defines what "close" means)
- We can think of our spatial model as a unique projection from a continuous SPDE onto our discrete mesh (Lindgen et al 2011)

# Penalized Complexity Priors

- There are two hyper-parameters for the matérn covariance function which can be re-expressed in terms of the Range (distance at which correlation falls below .1) and the spatial noise
- Priors are important in these kinds of (non-regular) models and even impact asymptotic and rates
- We use penalized complexity priors (Simpson et al 2017) (implemented directly in R-INLA!!!)

# Penalized Complexity Priors

- The idea is that we think about our model as existing somewhere on the plane of most complicated to a simpler nested model

$M_0$ $M_\infty$

Simple ⟷ Complex

Infinite Range + Zero Variance

Zero Range + Infinite Variance

Penalize the distance of our model $M$ from the base simple model $M_0$ using the KL divergence. (Information lost from using base model)

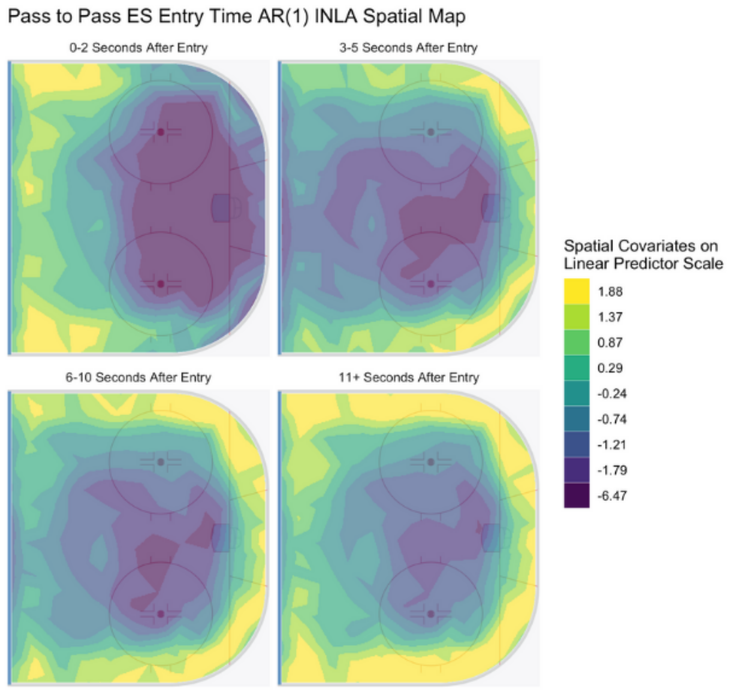
The hyper-parameters take the form of tail probabilities on the Range and variance

# GENERAL FINDINGS: WHAT ABOUT PASSES?

**Only 15% of players have a positive PAV from passing plays**

- High quantity of passes are along the boards
- Where passes are a lateral move in terms of value
- AND there's a possibility of a turnover

This leads our model to see the value of a pass as slightly negative on average.




**Question: Do too many passes hurt offence?**

**Unknown Factors**

- Player movement data and defensive structure.
- Is the player pressured to make a pass or not?
- No short-term value, but long-term benefits?