

Hockey in Space!

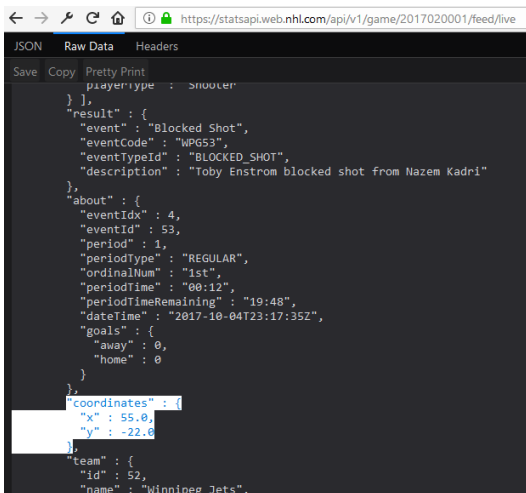
Characterizing team-wise differences in shot locations with
spatial point processes

Devan Becker
The University of Western Ontario
dbecker7@uwo.ca

September 16, 2018

Introduction

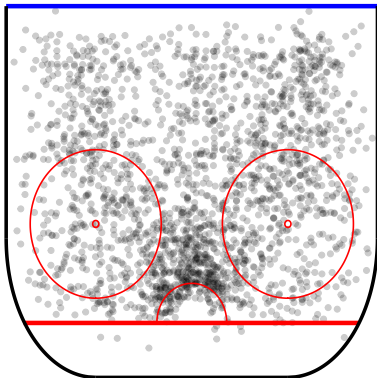
The Data



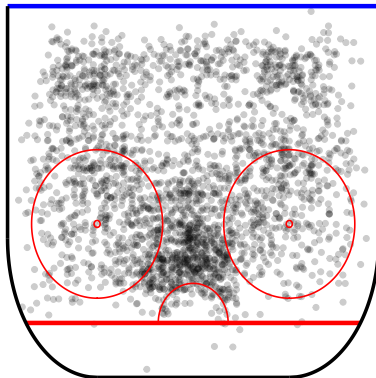
```
← → 🔧 ↻ 🏠 🔒 https://statsapi.web.nhl.com/api/v1/game/2017020001/feed/live
JSON Raw Data Headers
Save Copy Pretty Print
  ],
  "result": {
    "event": "Blocked Shot",
    "eventCode": "WPG53",
    "eventId": "BLOKED_SHOT",
    "description": "Toby Enstrom blocked shot from Nazem Kadri"
  },
  "about": {
    "eventIdx": 4,
    "eventId": 53,
    "period": 1,
    "periodType": "REGULAR",
    "ordinalNum": "1st",
    "periodTime": "00:12",
    "periodTimeRemaining": "19:48",
    "dateTime": "2017-10-04T23:17:35Z",
    "goals": {
      "away": 0,
      "home": 0
    }
  },
  "coordinates": {
    "x": 55.0,
    "y": -22.0
  },
  "team": {
    "id": 52,
    "name": "Winnipeg Jets",
```

2017-2018 Season Shot Locations

All Shots, Detroit



All Shots, Tampa Bay

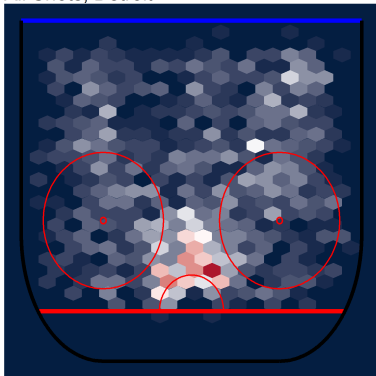


2017-2018 Season Shot Hexbins

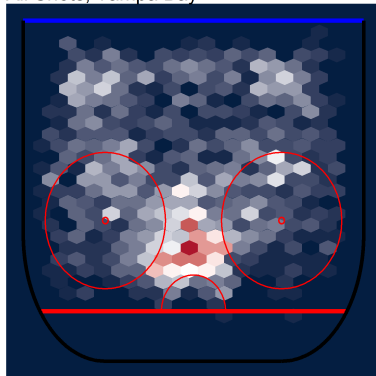
```
## Coordinate system already present. Adding new coordinate
```

```
## Coordinate system already present. Adding new coordinate
```

All Shots, Detroit

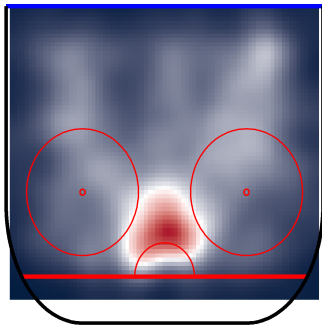


All Shots, Tampa Bay

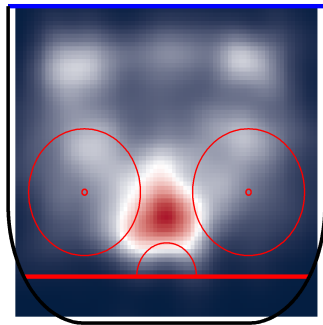


2017-2018 Season Shot Density

Shot Density, Detroit



Shot Density, Tampa Bay



Objective

Fit a parametric statistical model to determine:

- Where do different teams shoot from?
- Are the patterns consistent? (Variance!)
- ~~Which shots go in?~~
- What can goalies expect?

LGCP

Log-Gaussian Cox Processes (Yay math!)

$$\log(\Lambda(x, y)) = \mu + \beta C(x, y) + S(x, y)$$

The log of the **rate of points in a given location** is modelled as an intercept plus a **spatial covariate** plus a **random process**.

Log-Gaussian Cox Processes (Yay math!)

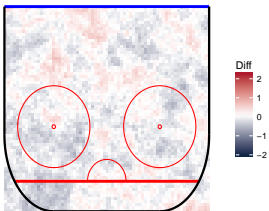
$$\log(\Lambda(x, y)) = \mu + \beta C(x, y) + S(x, y)$$

The log of the **rate of points in a given location** is modelled as an intercept plus a **spatial covariate** plus a **random process**.

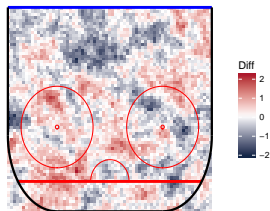
- “Random” doesn’t mean unstructured!
- The **random process** is a smooth function based on the normal distribution.

Variance and Clustering (Simulated)

Low Variance, Small Clusters



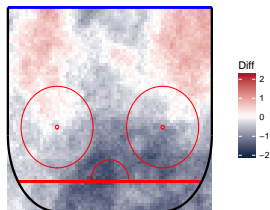
High Variance, Small Clusters



Low Variance, Large Clusters

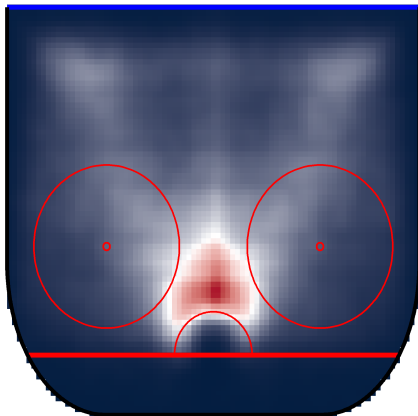


High Variance, Large Clusters



League Average as a Spatial Covariate

Density Estimate - 2017/2018 Season



League Average as a Spatial Covariate

Our interpretation becomes:

$$\log(\Lambda(x, y)) = \mu + C(x, y) + S(x, y)$$

The log of the **rate of shots** is modelled as an intercept plus the **league average** plus a **team specific deviation**.

League Average as a Spatial Covariate

Our interpretation becomes:

$$\log(\Lambda(x, y)) = \mu + C(x, y) + S(x, y)$$

The log of the **rate of shots** is modelled as an intercept plus the **league average** plus a **team specific deviation**.

- $S(x, y)$ can be seen as the *intended* strategy difference.
- The variance and range illustrate the team's *consistency*.

Estimation of LGCP

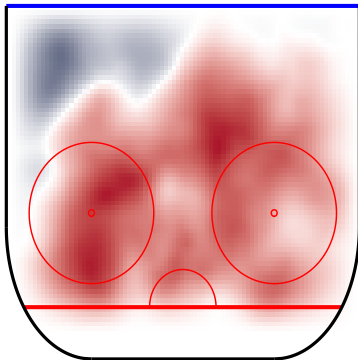
Estimation of LGCP

oh no

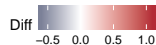
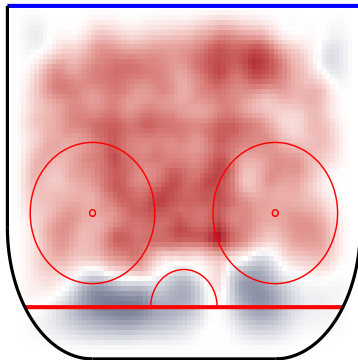
Results

Random Processes - $S(x,y)$

Diff. from League – Detroit

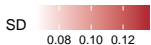
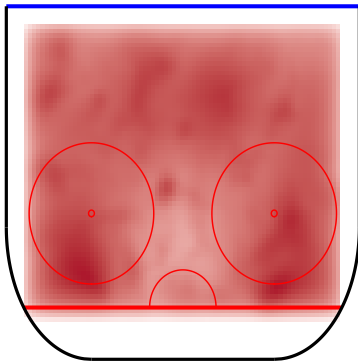


Diff. from League – Tampa Bay

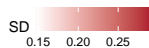
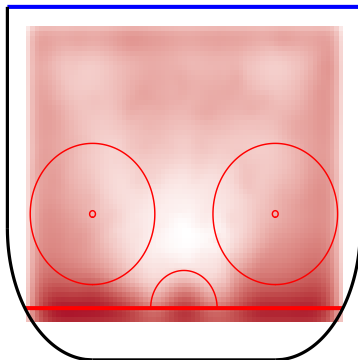


Variance at Every Location

SD – Detroit



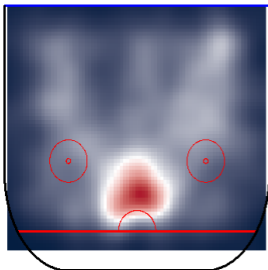
SD – Tampa Bay



All Results - Shiny App

Hockey in Space!

Shot Density, DET

Shot Density
0.5 1.0 1.5

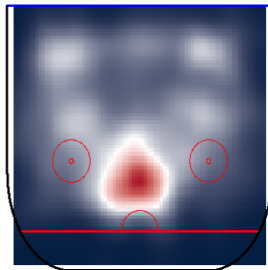
Team 1

DET

Data

Density

Shot Density, TBL

Shot Density
0.5 1.0 1.5 2.0

Team 2

TBL

Team Colours

WPG

Limitations

- Differences from league are often minor
- One dense cluster means small clusters estimated everywhere
- Needs a lot of data
 - Can't just look at one team's goals

Final Notes

Conclusions

- This method can indicate the manifested strategy
- Variance and cluster size indicate a teams consistency
 - Careful interpretation

Future Work

- Different play types (e.g. first shot after possession)
- Statistical comparison of teams
- Spatially varying range parameter
- All the things that the audience suggests

Acknowledgments

Thank you to my supervisors: Doug Woolford, Charmaine Dean,
and W. John Braun

Thanks for funding:



Canadian Statistical Sciences Institute
Institut canadien des sciences statistiques

Thanks for listening! Any further questions: dbecker7@uwo.ca