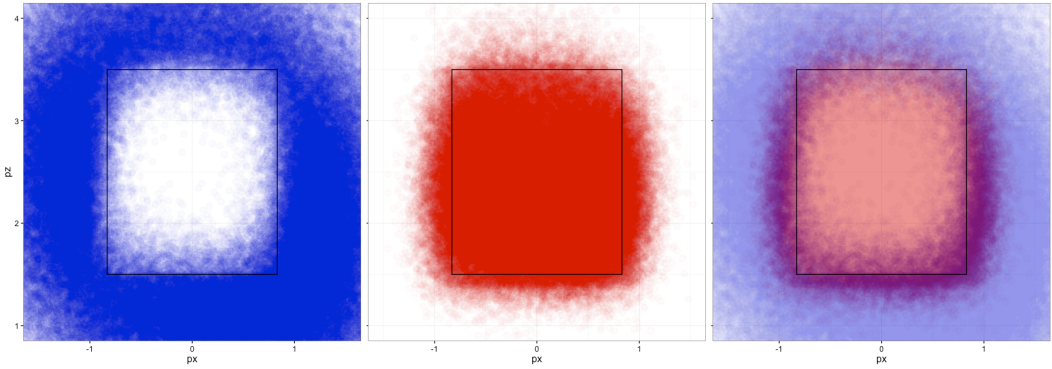
Measuring Umpire Consistency

David J. Hunter

January 11, 2018


pz
4
3
2
1
-1
0
1
px
-1
0
1
px
-1
0
1
px

Assumptions/Motivations

- Consistency is more important than conformity.
- A consistent zone need not be rectangular, but should be convex.
- Consistency *within* a game is important.
- Different zones for LH and RH batters are OK.
- One egregiously bad call is worse/as bad as several marginally bad calls.

Assumptions/Motivations

- Consistency is more important than conformity.
- A consistent zone need not be rectangular, but should be convex.
- Consistency *within* a game is important.
- Different zones for LH and RH batters are OK.
- One egregiously bad call is worse/as bad as several marginally bad calls.

Assumptions/Motivations

- Consistency is more important than conformity.
- A consistent zone need not be rectangular, but should be convex.
- Consistency *within* a game is important.
- Different zones for LH and RH batters are OK.
- One egregiously bad call is worse/as bad as several marginally bad calls.

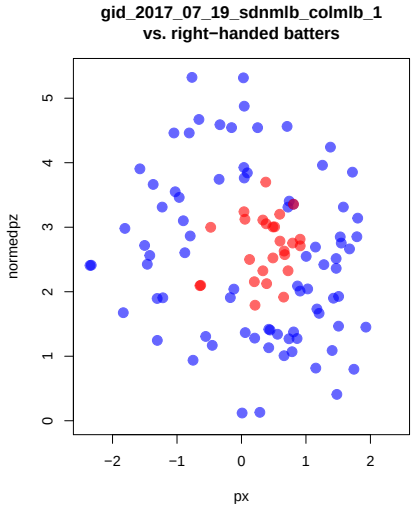
Assumptions/Motivations

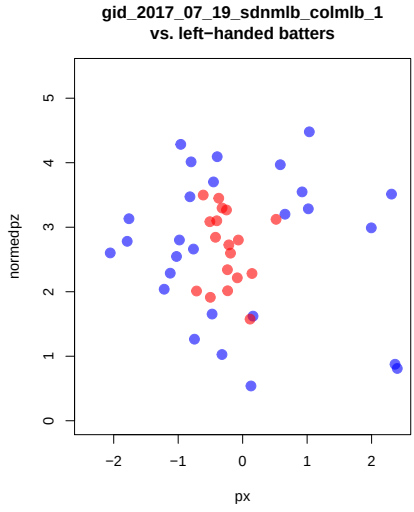
- Consistency is more important than conformity.
- A consistent zone need not be rectangular, but should be convex.
- Consistency *within* a game is important.
- Different zones for LH and RH batters are OK.
- One egregiously bad call is worse/as bad as several marginally bad calls.

Assumptions/Motivations

- Consistency is more important than conformity.
- A consistent zone need not be rectangular, but should be convex.
- Consistency *within* a game is important.
- Different zones for LH and RH batters are OK.
- One egregiously bad call is worse/as bad as several marginally bad calls.

Define established ball and strike zones





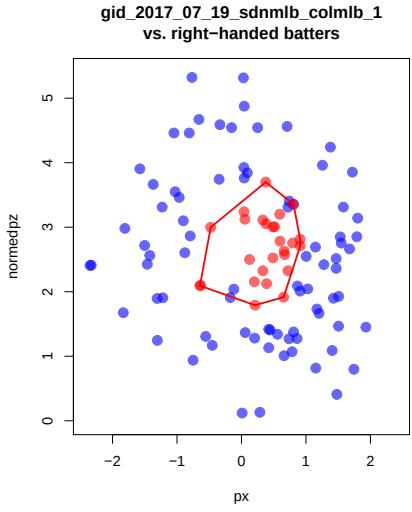
Established Strike Zone: Convex Hull

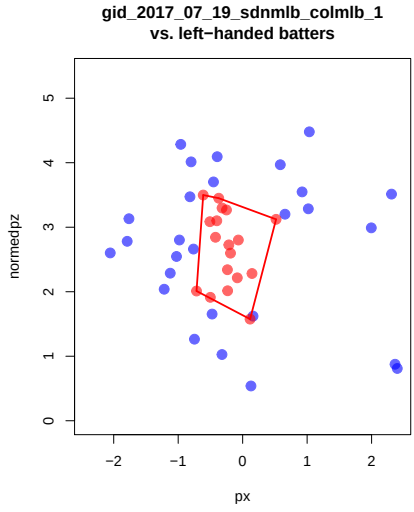
Let H_l denote the open half-plane bounded by the line l. The convex hull of a set of points P is the set

$$S = \bigcap_{\{H_l | H_l \cap P = \emptyset\}} H_l^c$$

When P contains the locations of all called strikes, S is the *established strike zone*.

Established Strike Zone: Convex Hull





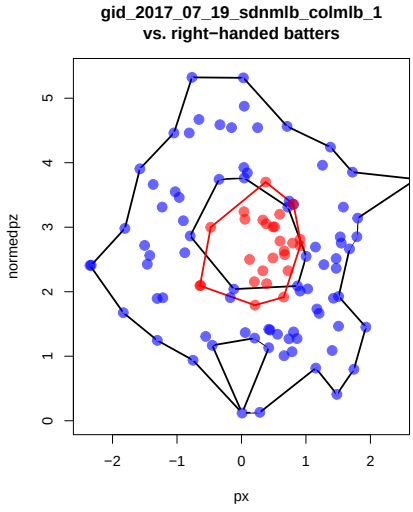
Established Ball Zone: α-Shape?

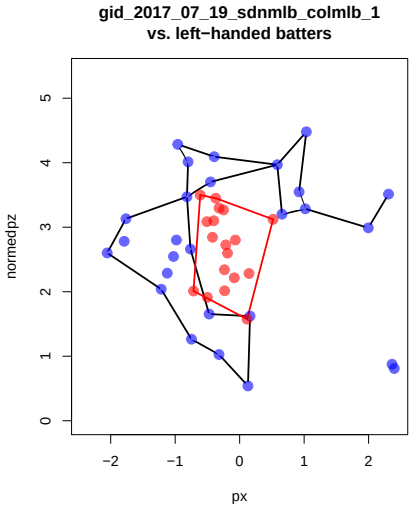
For $\alpha > 0$, let $B_{x,\alpha}$ denote the open ball in $\mathbb{R}^2$ of radius α centered at the point x. Given a set of points $P \subset \mathbb{R}^2$, two points $p_1, p_2 \in \mathbb{R}^2$ are *α-neighbors* if p_1 and p_2 lie on the boundary of some $B_{x,\alpha}$ such that $B_{x,\alpha} \cap P = \emptyset$.

The *α-shape* is the straight line graph formed by drawing line segments between α-neighbors.

- α-neighbors are close together but next to big empty regions.
- The α-shape is the outline of the points P.
- Does not need to be convex or even simply-connected.
- For large α, the α-shape is the boundary of the convex hull.

Established Ball Zone: α-Shape?





Established Ball Zone: α-Hull

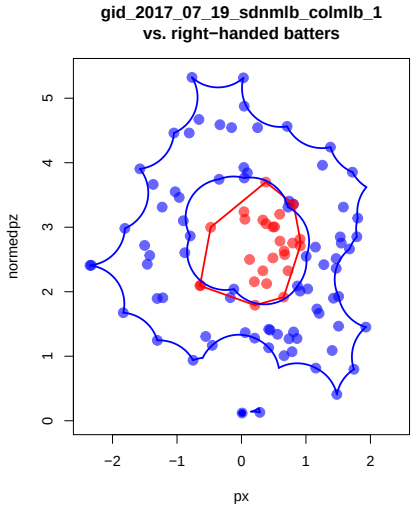
For $\alpha > 0$, let $B_{x,\alpha}$ denote the open ball in $\mathbb{R}^2$ of radius α centered at the point x. Given a set of points $P \subset \mathbb{R}^2$, the α-*hull* of P is the set

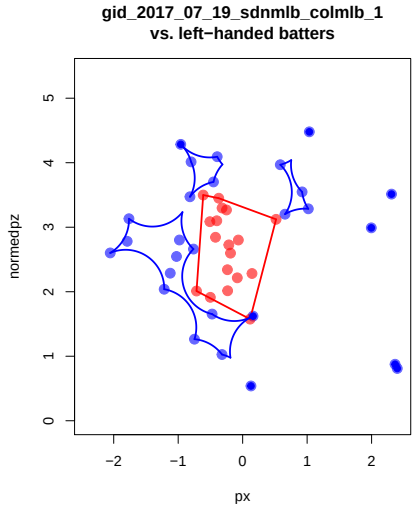
$$X = \bigcap_{\{B_{x,\alpha} | B_{x,\alpha} \cap P = \emptyset\}} B_{x,\alpha}^c$$

When P contains the locations of all called balls, X is the *established ball zone*.

- Approaches convex hull $S = \cap_{\{H_l | H_l \cap P = \emptyset\}} H_l^c$ as $\alpha \to \infty$.
- Intuitively: hole punch.
- May not be simply connected for smallish α.

Established Ball Zone: α-Hull





Inconsistency Index Metric

$S =$ established strike zone (convex hull of strikes)

$X =$ established ball zone (α-hull of balls)

$$\text{inconsistency} = \frac{\text{balls in } S + \text{strikes in } X}{\text{total calls}}$$

- This is a *per game* metric.
- Compute separately for right-hand batters and left-hand batters and add.

Inconsistency Index Metric

$S =$ established strike zone (convex hull of strikes)

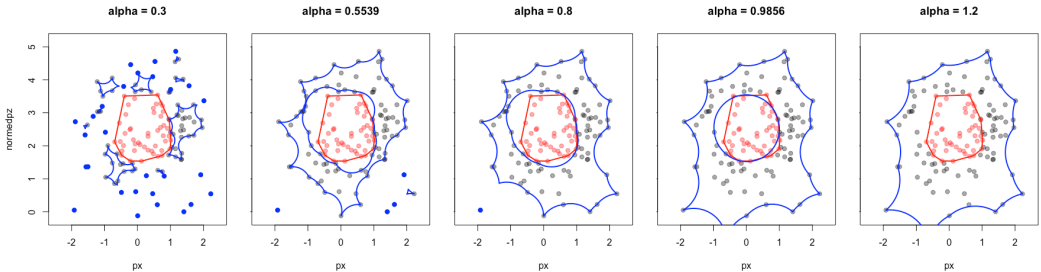
$X =$ established ball zone (α-hull of balls)

$$\text{inconsistency} = \frac{\text{balls in } S + \text{strikes in } X}{\text{total calls}}$$

- This is a *per game* metric.
- Compute separately for right-hand batters and left-hand batters and add.

Choosing α

- α too small: Ball zone has lots of components
- α too large: Seems unfair to umpires
- α way too large: Ball zone becomes simply connected
- $\alpha = 0.5539$: $1/3$ rule-book zone width



Correlations

```
round(cor(umpMetrics),2)
```

```
           Games ZoneSize Accuracy WalkRate AveIncon
Games       1.00     0.08     0.05    -0.03    -0.02
ZoneSize    0.08     1.00    -0.52    -0.38     0.38
Accuracy    0.05    -0.52     1.00    -0.01    -0.65
WalkRate   -0.03    -0.38    -0.01     1.00     0.09
AveIncon   -0.02     0.38    -0.65     0.09     1.00
```

Correlations: observations

	Games	ZoneSize	Accuracy	WalkRate	AveIncon
Games	1.00	0.08	0.05	-0.03	-0.02
ZoneSize	0.08	1.00	-0.52	-0.38	0.38
Accuracy	0.05	-0.52	1.00	-0.01	-0.65
WalkRate	-0.03	-0.38	-0.01	1.00	0.09
AveIncon	-0.02	0.38	-0.65	0.09	1.00

- Walk rate uncorrelated to inconsistency, accuracy.
- Smaller zones tend to be more accurate and consistent.
- Inconsistency is moderately correlated with accuracy.

Correlations: observations

	Games	ZoneSize	Accuracy	WalkRate	AveIncon
Games	1.00	0.08	0.05	-0.03	-0.02
ZoneSize	0.08	1.00	-0.52	-0.38	0.38
Accuracy	0.05	-0.52	1.00	-0.01	-0.65
WalkRate	-0.03	-0.38	-0.01	1.00	0.09
AveIncon	-0.02	0.38	-0.65	0.09	1.00

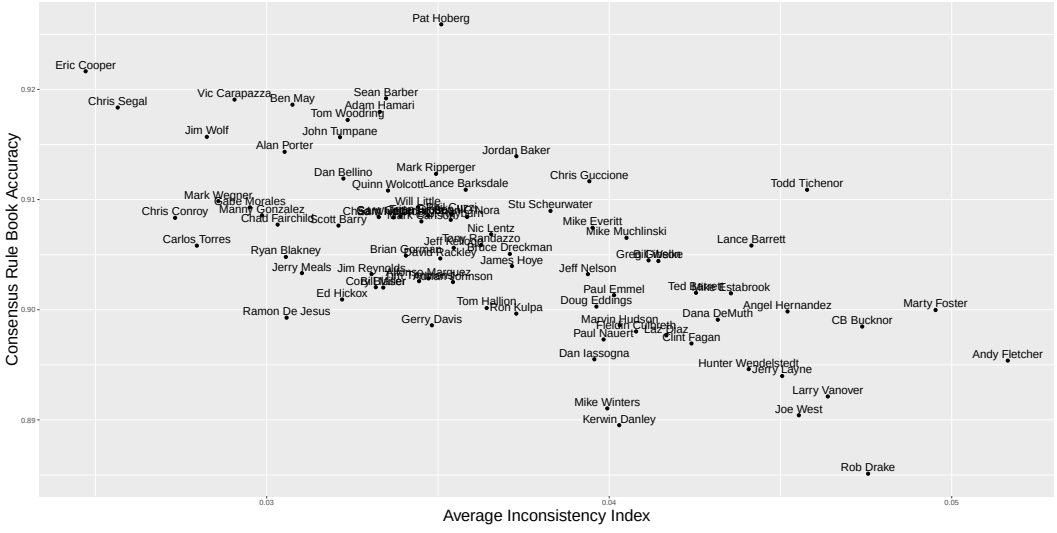
- Walk rate uncorrelated to inconsistency, accuracy.
- Smaller zones tend to be more accurate and consistent.
- Inconsistency is moderately correlated with accuracy.

Correlations: observations

	Games	ZoneSize	Accuracy	WalkRate	AveIncon
Games	1.00	0.08	0.05	-0.03	-0.02
ZoneSize	0.08	1.00	-0.52	-0.38	0.38
Accuracy	0.05	-0.52	1.00	-0.01	-0.65
WalkRate	-0.03	-0.38	-0.01	1.00	0.09
AveIncon	-0.02	0.38	-0.65	0.09	1.00

- Walk rate uncorrelated to inconsistency, accuracy.
- Smaller zones tend to be more accurate and consistent.
- Inconsistency is moderately correlated with accuracy.

Accuracy vs. Inconsistency ($r = -0.65$)



References

- Data from MLBAM.
- R packages: alphahull, pitchrx, baseballr
- Source code on GitHub: `djhunter/inconsistency`