OPR VERSUS TRUE AVERAGE SCORE: AN OBSERVATIONAL STUDY

Brennon Brimhall

WHAT'S OPR?

- Let $robot_1, robot_2\, , robot_3 \ldots robot_n$ be variables that represent the scoring output of robots at competition.
- Set up a system of equations, like so:
  - $robot_1 + robot_2 + robot_3 = combinedScore$
  - Solve the system for each robot's contributed score.
- The value of each variable is known as the OPR (Offensive Power Rating) for that robot.

HOW ACCURATE IS OPR?

- Due to the math used to determine OPR, we assume that OPR is an approximation of true average scoring power.
- However, no one really knows how good OPR is, in a quantifiable sense. A study has not ever been done before.
- Let's see!

METHODOLOGY

- We'll confine our population of interest to all robots who played in 2013.
  - Eliminates possible multimodal results (OPR may be more or less accurate in other years).
- We use Team 20's census (scouting) across their three official 2013 competitions as our data to compare OPR to.

METHODOLOGY

- We'll use the [Adambot's Automated Scouting Kit] to determine OPR.
  - A nice package of software that calculates OPR for us, by event.
  - Free and open-source.
- 196 Event/Team combinations were included in our sample.
  - We distinguished between teams' performances at multiple events; there are multiple OPRs and average scores in these cases.
  - List of teams used (listed on a per event/team combination basis; number listed twice means that that team has two event/team combinations associated with it):

LIST OF TEAMS USED

11	195	571	1027	1747	2408	3059	3623	4557
20	228	571	1071	1756	2415	3104	3627	4572
20	228	639	1071	1784	2457	3145	3634	4585
20	230	663	1075	1796	2468	3146	3654	4589
33	236	694	1100	1836	2486	3146	3718	4609
51	236	694	1100	1868	2543	3182	3719	4628
71	237	701	1124	1884	2590	3182	3930	4637
78	254	801	1218	1902	2604	3204	3958	4673
95	263	812	1289	1967	2614	3205	4055	4731
121	291	836	1289	1991	2621	3242	4067	4753
126	314	839	1334	2016	2648	3245	4097	4810
126	316	839	1448	2064	2704	3280	4134	4812
155	321	840	1493	2067	2705	3310	4254	
157	326	846	1519	2067	2709	3314	4265	
172	365	868	1595	2081	2785	3461	4410	
173	379	910	1610	2104	2791	3464	4450	
173	395	930	1660	2137	2836	3467	4466	
175	433	948	1665	2165	2877	3478	4470	
177	467	955	1687	2168	2959	3499	4471	
177	469	973	1699	2168	3003	3504	4499	
178	525	987	1714	2170	3008	3525	4501	
181	558	999	1735	2205	3044	3555	4537	
190	558	1011	1740	2370	3044	3612	4545	

BIAS

- Scouting is a form of survey. Since we're using Team 20's scouting data to get True Average Score, it's important to discuss the bias that may be present in that data.
- Due to the nature of FRC, we can rule out voluntary response bias and response bias.
  - All teams who participate are able to be observed by scouts for our census. E.g., you can't volunteer to play in a match.
  - Playing a match is a neutral situation. In 2013, the ranking system made it so alliances had a single objective: win. Assuming no teams threw matches, we can rule that matches did not influence match outcomes.

BIAS

- Potential undercoverage bias.
  - Out of 4116 team/event combinations for 2013, data was collected for only 196 (4.76%).
- Potential convenience bias.
  - The only surveyed events are the ones we attended. This may not be representative of the rest of the entire FRC population.
- Potential nonresponse bias.
  - When a team is broken/doesn’t show up, there is no way to accurately determine what their score should have been in that match.

HOW WE'LL ANSWER OUR QUESTION

- We're going to answer our question in two ways:
  - Look at the correlation of OPR to True Average Score (perform a least-squares regression; make sure the slope is close to 1 and the intercept is close to 0).
    - This takes care of OPR on average, over a large population.
  - Look at the distribution of error in OPR.
    - This takes care of determining the chance a given team will have an accurate OPR.

CORRELATION/SCATTERPLOT

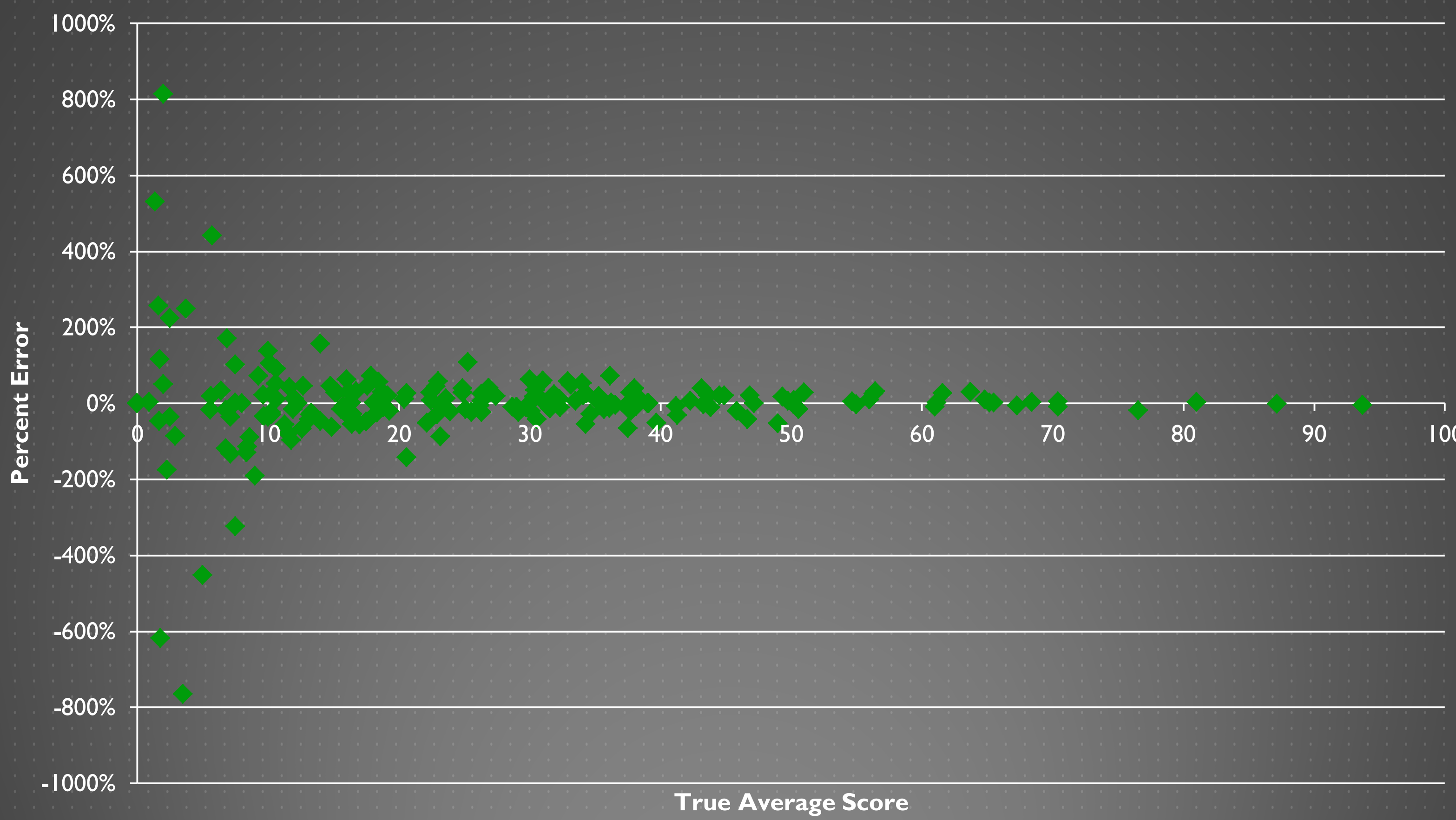
CORRELATION/SCATTERPLOT

- The best fit line is pretty darn close to the line we were expecting ($y = x$).
  - The slope (b_0) is only off by an error of 3.8%.
- The R^2 value is pretty good; our model has a medium level of strength.
  - However, we still want to evaluate the effects of this scatter with analysis of percent error.

PERCENT ERROR

- We'll also look at OPR's "percent error."
  - We want to know not only what happens on average, but what happens when we use OPR to predict the True Average Score for a single robot. How accurate will that be?
  - An error of 9 points is very different when a robot is scoring 90 points per match versus a robot scoring 10 points per match. To compensate, we'll use percent error to describe the discrepancy between OPR and True Average Score.
    - Calculated as follows: $\frac{OPR - True\ Average}{True\ Average}$
    - Out of necessity, we removed four team/event combinations because their robot had a True Average Score of zero, which would produce an undefined percent error.

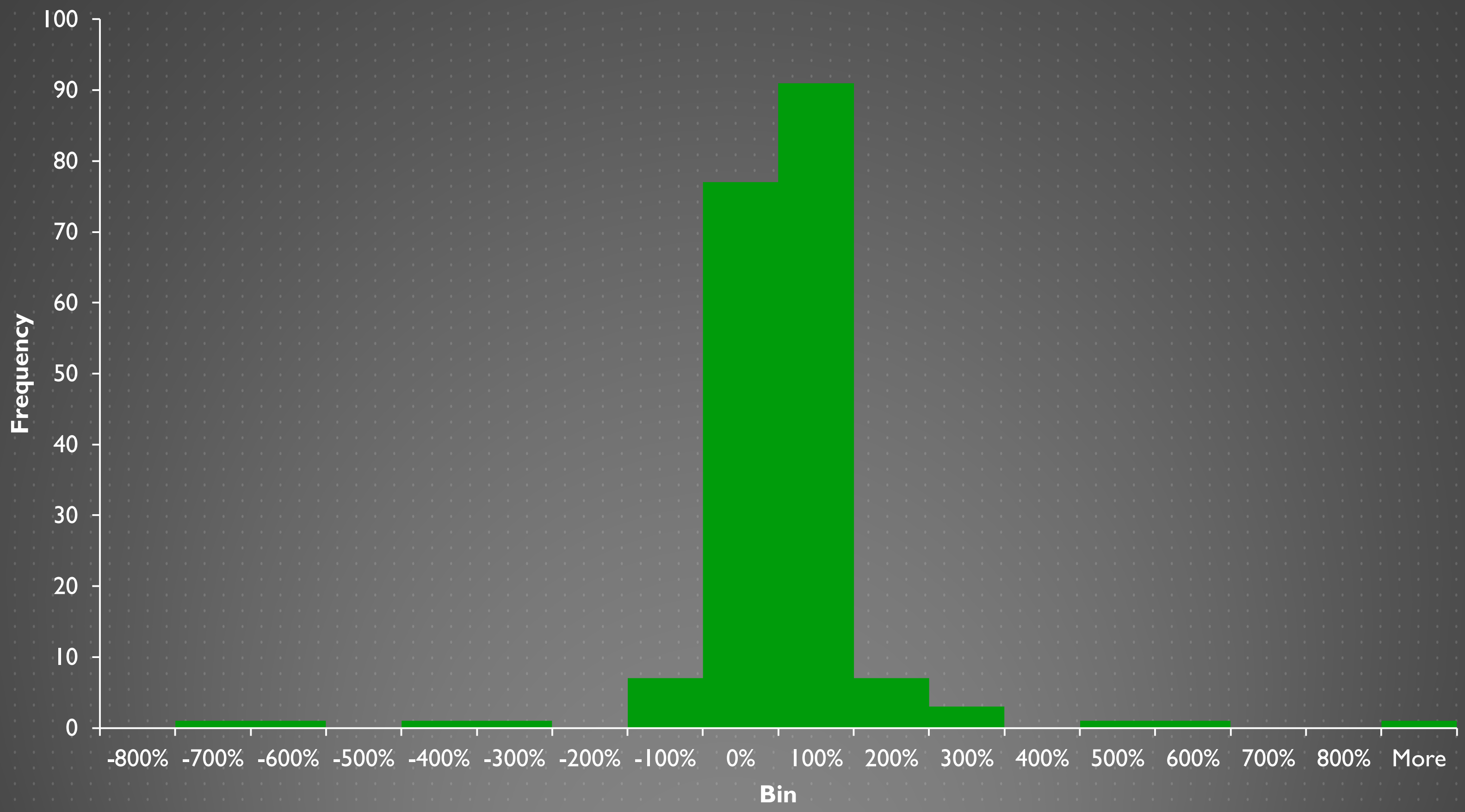
PERCENT ERROR SCATTERPLOT



PERCENT ERROR

- Based on our scatterplot, the absolute value of percent error decreases as True Average Score increases.
  - Many different reasons could explain this: teams who don’t score many points may contribute to OPR in other ways (synergy with alliance partners), have an easier schedule, or more variation in scoring output.
  - However, to validate any of the speculation above, an experiment would have to be performed.

PERCENT ERROR HISTOGRAM



PERCENT ERROR

- Based on our histogram (which looks approximately Normal), we'll assume that the normal model applies.
- $\mu(percent\ error) = 3.08\%$
- $\sigma(percent\ error) = 131.97\%$
- So, using a model of $N(3.08\%, 131.97\%)$, we can now calculate how likely OPR is going to be within a specified tolerance.

PERCENT ERROR

Percent Error Tolerance	Probability of being within Tolerance
$\pm$10%	6.039%
$\pm$20%	12.043%
$\pm$30%	17.978%
$\pm$40%	23.813%
$\pm$50%	29.515%
$\pm$60%	35.055%
$\pm$70%	40.409%
$\pm$80%	45.551%
$\pm$90%	50.463%
$\pm$100%	55.128%
$\pm$110%	59.533%
$\pm$120%	63.668%
$\pm$130%	67.529%
$\pm$140%	71.112%
$\pm$150%	74.418%
$\pm$160%	77.452%
$\pm$170%	80.220%
$\pm$180%	82.730%
$\pm$190%	84.995%
$\pm$200%	87.025%

CONCLUSION

- As an estimation of true scoring power at competition, **for a given team,** there's only around a 30% chance of OPR being within 50% of the True Average Score.
  - Depending on your tolerance, OPR is only going to be useful less than half the time.
    - After 50% error, OPR becomes much less useful at the competition to estimate teams' scoring output in real time.
    - Only 55% of all OPRs collected are within 100% of True Average Score. After this point, OPR is useless as an indicator of True Average Score on a per team basis.
  - OPR is a bad metric to be using at the competition with an expectation of fine granularity and high accuracy/precision.

CONCLUSION

- As an estimation of true scoring power **on average**, OPR can be very accurate.
  - It correlates and regresses nicely with True Average Score.
  - OPR and True Average Score's coefficient of correlation is within 3.8% of our expected value, with a negligible intercept constant (less than a point).
  - Our R^2 value is 80.3%, meaning that 80.3% of the variation in True Average Score is reflected in the variation of OPR.
  - If you're not needing high granularity, and you're not looking at a specific team, OPR can be useful.