Robust and hybrid OPR for FRC

Mike Paulonis[1]

Mentor, FRC Team 4020 – Cyber Tribe

June 8, 2020

Abstract

Two alternate forms of offensive power rating (OPR) were developed and studied using 2020 FIRST Robotics Competition season data. Hybrid OPR combines per-robot data from the field management system (FMS) with component OPRs for scoring that is not uniquely identified by FMS to create an OPR where some components are known without error. Robust OPR identifies and removes outlier alliance performances from the set of all performances used to compute OPR in the traditional way. The hybrid and robust techniques can also be applied together.

The hybrid, robust, and hybrid robust OPRs were each better at predicting playoff alliance scores and scoring margin across 2020 events compared to standard OPR. The hybrid robust approach reduced the sum of squared errors (SSE) of playoff alliance score predictions by almost 7% compared to standard OPR and reduced the SSE of playoff match scoring margin prediction by almost 5%. Average improvement was similar for district and regional events.

The study was motivated by the substantial disagreement between standard OPR and scouting point contribution estimates for team 4020 at the 2020 Palmetto regional. The standard OPR for 4020 was 47.9 at Palmetto compared to a scouting estimate of 67.0 points per match and a hybrid robust OPR of 69.2. The much improved agreement with scouting came from clearly identifiable sources of poor endgame standard component OPR which was replaced with accurate FMS-supplied endgame contribution and an outlier match where the best team at the event scored 0 points in an alliance with 4020. Overall, hybrid robust OPR reduced the SSE of OPR vs team 4020 scouting estimates by 65% across all teams at 2020 Palmetto.

Splitting out fouls to compute playoff scores by using no-foul OPRs for a given alliance plus fouls-committed component OPRs for the opponent alliance were found to have a detrimental effect on playoff score prediction regardless of OPR technique for the 2020 season data.

---

[1] e-mail: mpaulonis@gmail.com. Team 4020, Cyber Tribe, is from Dobyns-Bennett High School in Kingsport, Tennessee, USA.

1 Introduction

This paper is motivated by the OPR results at the 2020 Palmetto Regional. Team 4020 was one of the participants. 4020 had a few issues during Friday quals matches which resulted in a 4th place qualification rank, but we were fortunately part of a strong #1 alliance with 694 and 1758 which won the event.

Good scouting (in our opinion ☺) led 694 to select us as their alliance partner. However, if they had been relying much on OPR, they almost certainly would have been looking elsewhere. We were ranked 9th in OPR at 2020scmb with an OPR of 47.9. Our own scouting showed us contributing an average 67 points per match and being ranked 2nd in point contribution at Palmetto. That is a huge difference numerically as well as in rank. I wanted to see if there was a way that OPR could be improved such that it was in better agreement with scouting data.

1.1 Robust OPR

We knew that our OPR was likely to suffer a large negative impact as a result of our first qualifying match. In 2020scmb_qm9 we played with 694 and 1708. 694 had a problem with their CAN bus and did not move for the entire match. Given that 694 turned out to have the strongest robot at Palmetto and they contributed 0 points to our first match we expected a negative impact to our OPR (and that of 1708 as well) through no fault of our own. My first question was if it could be possible to identify "outlier" alliance performances such as this in qualifications and if removing them before computing OPR would result in an improved "robust" OPR. Improved in this case would be better agreement with our scouting data, but perhaps more importantly in the big picture, improvement in predicting playoff alliance scoring. OPR already suffers from a low match-to-team ratio, so removing match information seems like a step in the wrong direction. However, if the right, misleading, information is removed, the net result might be more useful OPRs.

1.2 Hybrid OPR

The second question was motivated by the observation that the Field Management System (FMS) for Infinite Recharge contained robot-specific data for autonomous initiation line exiting as well as endgame position (hang, park, or none). Along with the FMS data on generator switch level, the entire endgame contribution of individual robots was known exactly. The same thing is true for a portion of autonomous contribution. The question, although it really doesn't seem like there should be any question, is if creating a "hybrid" OPR combining robot-specific contribution known exactly from FMS with estimated component OPRs for the other scoring modes not identifiable in a per-robot way from FMS would result in improved agreement with scouting data and playoff scoring.

1.3 Accounting for foul points in OPR

The last question was motivated by the difference in the OPR and our scouting data for team 2614. There is no question that 2614 was an outstanding performer at 2020scmb, but our scouting data showed them with an average contribution of 61.1 points per match while they had a 75.1 OPR. We know that our scouting data is less than perfect, but an average 14-point difference per match is well beyond the amount of error we would expect. Looking at their component OPRs suggested that the difference was due to foul points received. The foul points component OPR for 2614 was 14.2, which was the 3rd highest at 2020scmb. The high foul points component OPR was backed up by data showing that 2614's qualification alliances were the recipients of 183 foul points in 9 matches for an average of 20.3 points per match. This is much larger than the average 8.2 foul points per Palmetto alliance qualification match from TBA. The OPR-related question then becomes checking if playoff scoring predictions are more accurate by combining foul-points-free OPRs and adding *opponent* fouls-committed component OPRs as compared to simply summing standard OPRs which include fouls drawn.

2 Study of hybrid OPR

Given the questions at hand, the most likely way to improve overall OPR is to replace estimation with precisely known values where possible, which I have called hybrid OPR. This is also easy to carry out. For this reason, it is first to be studied.

2.1 Match data

I acquired match data for use in the study from The Blue Alliance API. The API /event/{event_key}/matches endpoint returns all the match scoring detail needed to compute OPR, component OPRs, and FMS per-robot contributions necessary for hybrid OPR as well as the other questions to be investigated. This study was conducted for all the FRC events completed in the 2020 season, of which there were 52.

2.2 Computing hybrid OPR

Computing hybrid OPR starts with computing component OPRs. This is conceptually the same as computing standard OPR, except overall alliance score in the OPR equation is replaced by the alliance score for just a portion of the game. The OPR equation is:

$$\boldsymbol{Ax} = \boldsymbol{b} \tag{1}$$

where $\boldsymbol{A}$ is a matrix with dimension [*matches*2, teams*] and represents team participation in matches, $\boldsymbol{x}$ is a column vector of length *teams* representing OPR, and $\boldsymbol{b}$ is a column vector of length *matches*2* of the scores you want to estimate with OPR. For example, if you want to estimate the component OPR of Power Cell points in autonomous ($cOPR_{apc}$) and in the first qualifying match, teams 1, 3, and 6 were an alliance and scored 24 auto cell points, while teams 2, 4, and 5 were an alliance and scored 10 auto cell points, the upper left part of equation (1) would be:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & \cdots \\ 0 & 1 & 0 & 1 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} cOPR_{apc,1} \\ cOPR_{apc,2} \\ \vdots \end{bmatrix} = \begin{bmatrix} 24 \\ 10 \\ \vdots \end{bmatrix} \tag{2}$$

note that in this way of structuring the OPR equations, there are two rows for each match. There are other ways of forming the equation system, but this is easy to understand and will be helpful later when computing robust OPR.

There are many ways to solve the system represented by (1). For my study, I programmed in Python and I chose to use LinearRegression from the Python Scikit-Learn package. Numpy linalg.solve would be another straightforward option. Both are plenty fast for the size of the system of equations in this study. There was no need to worry about OPR solution strategies that stabilize OPR as matches take place throughout an event. Techniques like ixOPR, or MMSE OPR are nice within an event, but since this analysis is post-event with all qualification matches complete, standard least-squares solutions are fine.

As described in the introduction, the FMS captures some robot-specific scoring for the 2020 game. From the TBA API dataset referenced earlier, initLineRobot1, initLineRobot2, and initLineRobot3 indicate directly if each of those robots exited the initiation line during auto or if they did not. An exit is worth 5 points, is known exactly per robot, and does not have to be estimated with a component OPR. Thus:

$$FMS_{autoExit,i} = \begin{cases} 0, & \text{if initLineRoboti=None} \\ 5, & \text{if initLineRoboti=Exited} \end{cases} \tag{3}$$

Similarly, endgameRobot1, endgameRobot2, and endgameRobot3 indicate directly what the final status was for those robots at the end of the match. Options are Hang, Park, or None. A hang is worth 25 points and a park is worth 5 points, so that part of endgame does not have to be estimated. The API dataset also has endgameRungIsLevel, with possible values IsLevel and NotLevel. A level rung is worth 15 points if any robots are hanging. From an OPR point of view, the most reasonable way to handle level rung points is to split them equally between all hanging robots. The hang or park status along with the equally divided share of a level rung completely defines the endgame contribution of a specific robot so an endgame component OPR is not needed.

$$H_i = \begin{cases} 0, & \text{if endgameRoboti} \neq \text{Hang} \\ 1, & \text{if endgameRoboti} = \text{Hang} \end{cases} \tag{4}$$

$$H = \sum_{i=1}^{3} H_i \tag{5}$$

$$FMS_{endgame,i} = \begin{cases} 0, & \text{if endgameRoboti = None} \\ 5, & \text{if endgameRoboti = Park} \\ 25, & \text{if endgameRoboti = Hang} \\ & \quad \text{and endgameRungIsLevel = NotLevel} \\ 25 + \dfrac{15}{H}, & \text{if endgameRoboti = Hang} \\ & \quad \text{and endgameRungIsLevel = IsLevel} \end{cases} \tag{6}$$

Component OPR is still needed for three parts of scoring that FMS cannot ascribe to individual robots -- autonomous cell points, teleop points minus endgame points, and foul points. Those values are available from the API as autoCellPoints, teleopPoints, endgamePoints, and foulPoints.

Putting all this together, hybrid OPR for Infinite Recharge can be defined as:

$$OPR_h = FMS_{autoExit} + OPR_{autoCell} + OPR_{teleop-endgame} + FMS_{endgame} + OPR_{foul} \tag{7}$$

For comparing against 4020 scouting data, hybrid OPR without the foul contribution is needed. Our scouting over the years has included watching for which robots commit fouls. Our scouts have not recorded which robots have been fouled. We may want to change that in the future, but for now we can only use the data we collected this season.

$$OPR_{h,nf} = FMS_{autoExit} + OPR_{autoCell} + OPR_{teleop-endgame} + FMS_{endgame} \tag{8}$$

Of course, this definition of hybrid OPR is season-specific and the formula would have to be revised for new games. This is not a particular problem, as anyone who computes component OPR already needs to revise their code each season for new scoring modes. The hybrid OPR calculation would be just one more thing to revise.

2.3 Evaluating hybrid OPR

OPR and hybrid OPR were computed for all 52 events completed in the 2020 season. The OPRs used "training" data from qualification matches and then were used to predict alliance playoff scores which were not part of the OPR computation. The predicted alliance score is simply the sum of OPRs for the teams in the alliance.

$$\hat{S}_a = OPR_{a1} + OPR_{a2} + OPR_{a3} \tag{9}$$

A weighted sum of squared error (SSE) was used to evaluate playoff prediction performance. Weighting was needed since some alliances participate in more playoff matches than others. Without weighting, prediction errors related to an alliance that made it to finals would have much more influence on the SSE metric than errors related to an alliance that was out after quarterfinals. Weighting was simply the inverse of the number of matches played by an alliance in the playoffs.

$$w_a = \frac{1}{m_a} \tag{10}$$

No accommodation was made to the weighting if an alliance had to employ a backup robot. The score prediction would use the OPR for the backup robot, but from a weighting perspective, the match would still be counted in the sum for the given alliance. The impact of this simplification is expected to be very small.

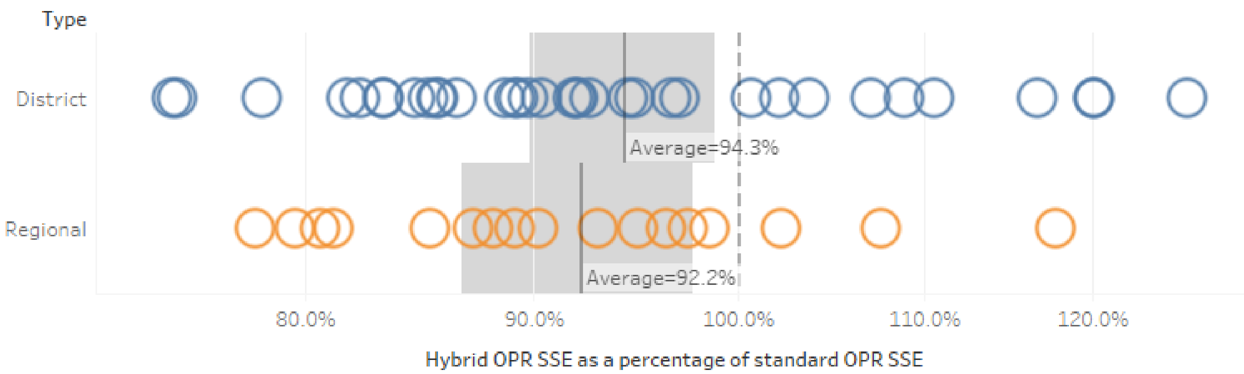
The prediction error metric becomes:

$$SSE = \sum_{i=1}^{2p} w_a \left(\hat{S}_{a,i} - S_a\right)^2 \tag{11}$$

where p is the number of playoff matches such that there are two data points for each match, one for each alliance.

The SSE for playoff predictions was computed for each of the 52 events and for both standard and hybrid OPR. A comparison of hybrid OPR prediction SSE as a percentage of standard OPR prediction SSE is shown in Figure 1.

Figure 1: Hybrid OPR SSE as a percentage of standard OPR SSE by event and type

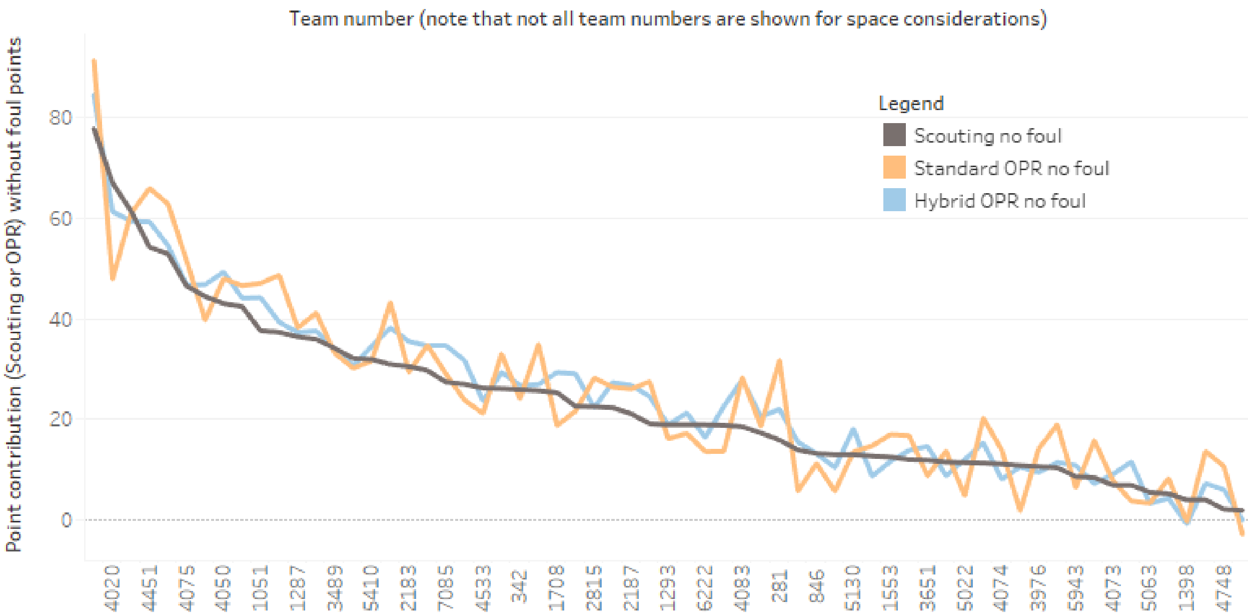


Each circle is one event. Hybrid OPR results in playoff alliance score predictions that are from 6-8% better than standard OPR on average. Regionals benefit slightly more than district events. This is not unexpected and is likely due to districts having more matches and fewer teams so standard OPR estimates should be more representative of team performance than at regional events. The gray shaded areas are the 95% confidence intervals of the averages. Since neither includes 100%, there is greater than 95% statistical confidence that hybrid OPR is a better predictor of playoff scores than standard OPR. If all events were combined rather than split into district and regional, the confidence interval would be even narrower and the statistical significance greater due to larger sample size.

It is no great surprise that hybrid OPR turns out to be a better scoring estimator than standard OPR since it incorporates some precisely known information. However, it is good to confirm with data that the precisely known information improves average playoff score predictions in actual events. Future years may have more or less robot-specific information in the FMS, so hybrid OPR may be more or less advantageous in the future.

Another way to assess OPR is to compare it to scouting data. Figure 2 shows a comparison of the point contribution from scouting data from Team 4020 at the 2020 Palmetto Regional with standard OPR and hybrid OPR. In all cases, the values do not include points for fouls. 4020 scouting captures information about fouls that robots commit rather than fouls robots receive. Since match scoring in 2020 and many previous years includes fouls that robots receive, standard OPR and 4020 scouting point contribution are not directly compatible. Thus, in the figure, scouting point contribution does not include fouls, and OPRs are computed on alliance scores without foul points. Equation 8 shows that computation for hybrid OPR.

Figure 2: Comparison of standard and hybrid OPR to scouting data for 2020 Palmetto



The teams are sorted by scouting point contribution order. The hybrid OPR values in blue are much closer to the scouting points than the standard OPR values in orange. Table 1 shows an aggregate numeric comparison via sum of squared errors and root mean squared error.

Table 1: Comparing error between 2020 Palmetto scouting and types of OPR

Error Computation	Sum of squared errors	RMS error
Standard OPR vs Scouting	2860	6.74
Hybrid OPR vs Scouting	912	3.80

Comparison to scouting is perhaps less valuable than predicting playoff scores, but it is still good confirmation of the value of hybrid OPR. In this case it is satisfying as it starts to fix some of the problem with OPR for 4020 at Palmetto. OPR with no foul for 4020 was 47.9. Hybrid OPR with no foul is 61.3, which is much closer to our scouting estimate without fouls, which was 67.0. The 4020 hybrid OPR without fouls is also ranked 2nd at Palmetto, which is the same as we ranked ourselves via scouting. By standard OPR without fouls, we were ranked 8th.

3 Study of robust OPR

As described in the introduction, removing a very limited number of outlier alliance qualification match performances before computing OPR results in something I have termed robust OPR. An outlier performance for an alliance could be caused by a catastrophic event such as having a robot overturned or having a robot not participate, if that type of event is not characteristic of that robot as evidenced by repeated occurrences throughout the event. An outlier performance could also be caused by a foul flood, especially when there is nothing about the alliance robots that makes them particularly amenable to drawing a huge number of foul points and that happening in multiple matches.

3.1 Computing robust OPR

The essence of computing robust OPR is discounting certain alliance performances such that predicted b_i are close to actual b_i. That is, solving (1) by:

$$\underset{\boldsymbol{x}}{arg\min} \sum_{i=1}^{2m} w_i |b_i - \boldsymbol{A}_i \boldsymbol{x}|^p \qquad (12)$$

where w_i is a weighting factor on a particular alliance performance. There are $2m$ alliance performances where m is the number of qualification matches. One way to solve (12) for an arbitrary p-norm and to adjust the importance of each alliance performance individually through w_i is to use iteratively reweighted least squares. I tried this technique, but I was not able to improve prediction of match scores that were not part of the OPR computation set, specifically playoff matches from the same event. The technique was too aggressive in de-weighting many qualification matches and I was not able to find a cross-validation technique that was effective in regulating the algorithm to minimize errors in playoff matches (test set data) rather than just qualification matches (training set data).

Thinking of the problem in a somewhat simpler way, I wondered if it would be possible to simply identify alliance performances that were very poorly predicted by the OPR solution and then to eliminate them from the training data set and re-compute OPR. The technique I chose for identifying the very poor predictions was boxplot outliers. Given a solution to (1) of:

$$\boldsymbol{A}\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{b}} \tag{13}$$

then boxplot outliers would be rows i of $\boldsymbol{A}$ and $\boldsymbol{b}$ where:

$$\hat{b}_i - b_i > Q3 + k_o IQR \tag{14}$$

$$or$$

$$\hat{b}_i - b_i < Q1 - k_o IQR \tag{15}$$

and $Q1$ is the first quartile or 25th percentile of the complete set of $\hat{b}_i - b_i$ for the event, $Q3$ is the third quartile or 75th percentile, and $IQR = Q3 - Q1$ is the interquartile range. k_o is a constant that traditionally has a value of 1.5. However, there is nothing magic about using $k_o = 1.5$. Values of 1.0 or 2.0 would be very aggressive or very conservative, respectively, in identifying outliers. However, values closer but not equal to 1.5 would result in minor changes to the fraction of outliers and could be reasonable. k_o will be treated as a constrained optimization parameter to select a fraction of outliers that minimizes the prediction error in playoff scores.

If there are no boxplot outliers from (14) or (15) after the first or any computation of OPR then the computed OPR becomes robust OPR. If there are one or more boxplot outliers, another iteration is conducted of computing OPR without the outliers from (14) and/or (15) followed by another search for boxplot outliers.

Because hybrid OPR has been shown to be a better average predictor of playoff alliance scores compared to standard OPR, hybrid OPR will be used as the reference OPR for identifying outliers. This choice minimizes the possibility that something about the OPR calculation contributes to an identified outlier and increases the likelihood that it is something about the alliance match performance that is extremely unusual.

In the end, the removal of boxplot outliers is the equivalent of a very limited number of w_i in (12) being set to 0, while all others remain at 1.

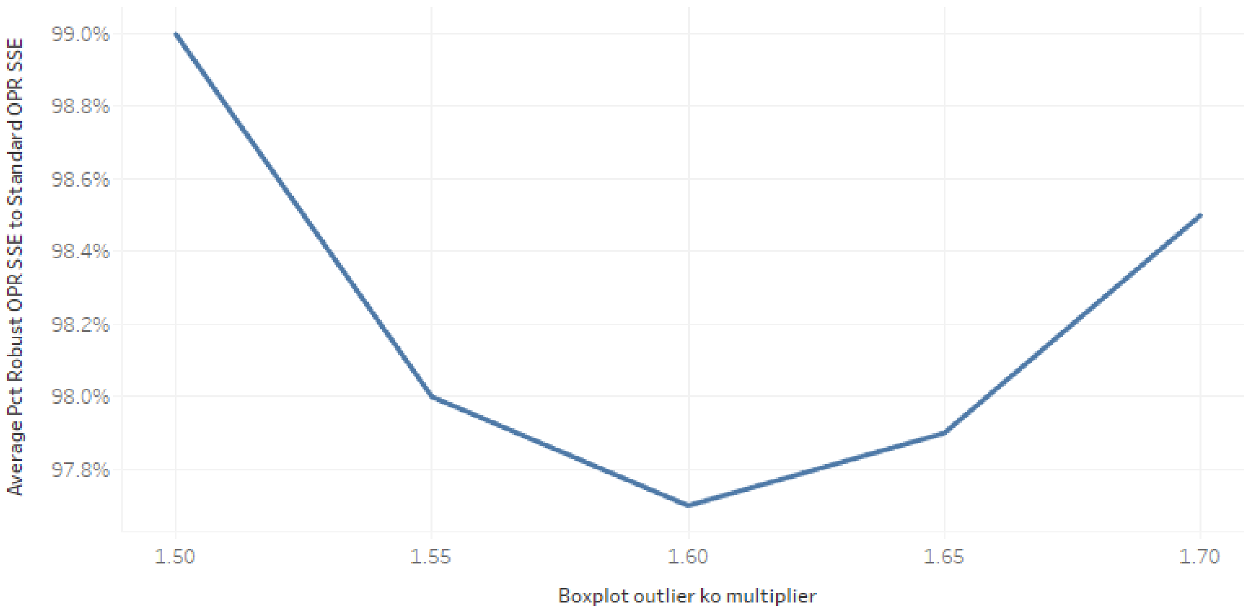
3.2 Evaluating robust OPR

The same weighted SSE technique applied to playoff alliance scores as described for hybrid OPR was used for evaluating robust OPR. Similarly, robust OPR was computed for all 52 completed FRC events in 2020.

Robust OPR was looked at in two ways. First is robust OPR computed with just outlier-filtered alliance scores. This would be compared to standard OPR and is the approach taken in this section. The second is hybrid robust OPR where hybrid OPR in (7) and (8) is computed with outlier-filtered alliance sub scores as well as robot-specific contributions determined from FMS data. This would be compared to hybrid OPR with all qualification performances included. This is the topic of the next section.

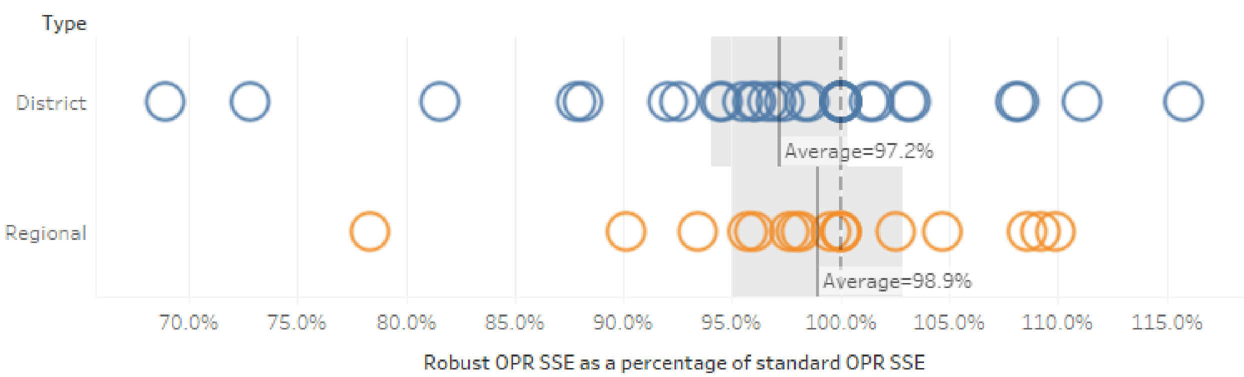
The k_o multiplier in (14) and (15) was determined experimentally by computing robust OPR and predicting playoff scores for all 2020 events. The k_o that minimized the average playoff score prediction SSE using robust OPR over all the events was 1.6. This is a slightly less aggressive identifier of outliers compared to the "standard" 1.5 typically used in boxplot computations. Requiring a more conservative test for identifying an outlier is consistent with the high information value of every alliance performance at an event, since there are so few "samples" from which to compute OPR.

Figure 3: Experimental search for optimum boxplot outlier multiplier



Robust OPR is computed using $k_o = 1.6$ throughout the remainder of this paper. A comparison of robust OPR SSE as a percentage of standard OPR SSE is shown in Figure 4.

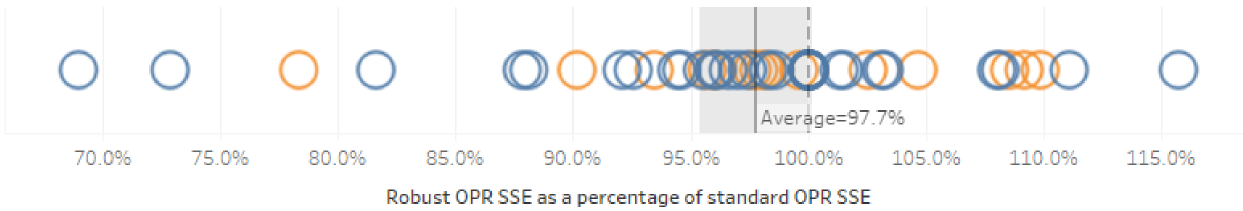
Figure 4: Robust OPR SSE as a percentage of standard OPR SSE by event and type



Each circle is one event. Robust OPR results in playoff alliance score predictions that are from 1-3% better than standard OPR on average. Districts benefit slightly more than regional events which is opposite of hybrid OPR. This is likely occurring because districts have more qualification matches per team than regionals so there are more chances for isolated outlier-level events to happen and subsequently be removed for robust OPR.

Figure 5 shows district and regional events combined into one analysis.

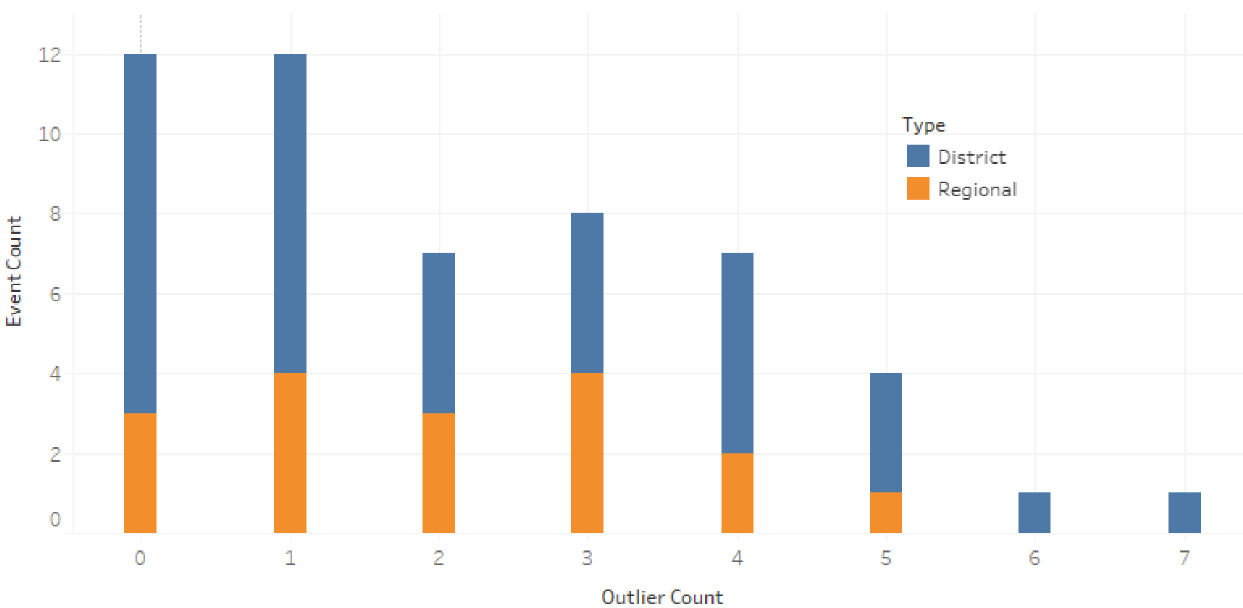
Figure 5: Robust OPR SSE as a percentage of standard OPR SSE by event



The gray shaded 95% confidence interval of the average improvement extends to 100.1%. Thus, there is approximately a 95% confidence that robust OPR is a better predictor of playoff scores than standard OPR.

It is not obvious in Figure 4 or Figure 5, but there are a number of events where the robust OPR is exactly 100% of standard OPR. These are cases where no outliers were found in alliance qualification performances so robust OPR is the same as standard OPR. Figure 6 shows the distribution of the outlier count per event.

Figure 6: Robust OPR outlier count distribution for all 2020 events

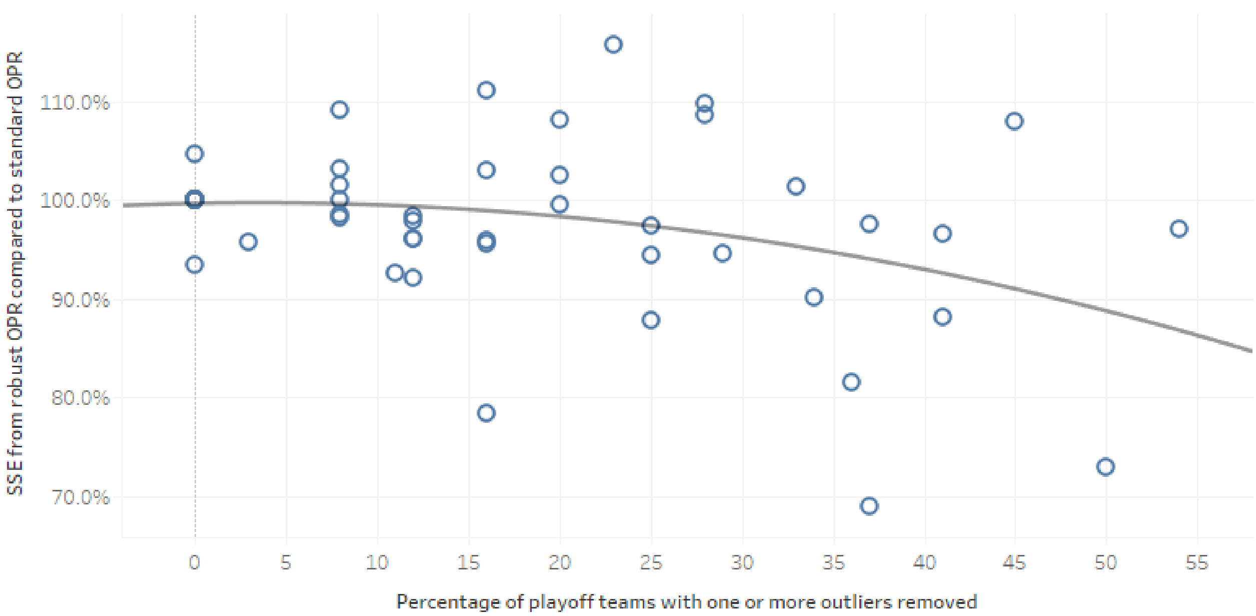


There were 3 regionals and 9 district events where no outlier alliance performances were detected. The maximum number of outliers was 7 at one district event. That event was

FIT Greenville from week 1. The error in predicted playoff scores was improved by 11.9% using robust OPR compared to standard at that event, so even with so many alliance performances removed from OPR calculation, the net impact on predictive power can still be positive.

Because there are many events with just a few outlier alliance performances, it is possible that teams involved in those outlier performances are not in the playoffs. In that case, playoff score predictions would only be indirectly influenced by outlier removal. One would hope that the more that playoff teams were directly involved in outlier qualification performances, the better that playoff alliance score predictions would be improved. Figure 7 shows this effect.

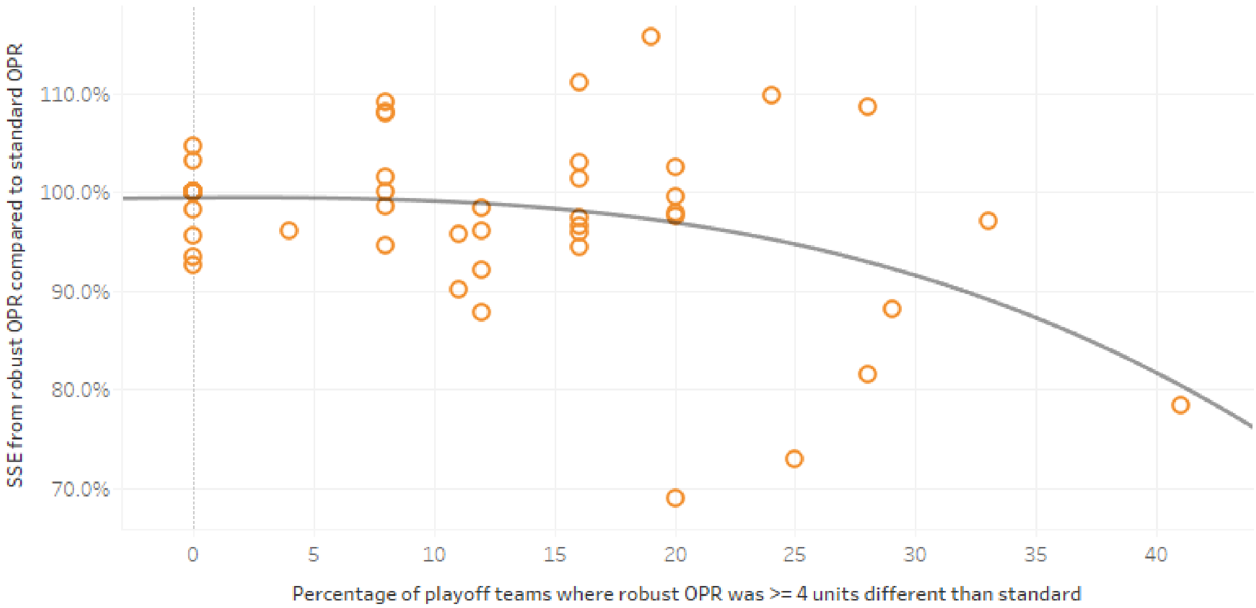
Figure 7: Effect of prediction accuracy by involvement of playoff teams in outlier matches



The more that teams involved in alliance performances which were identified and removed as outliers participated in the playoffs, the more that robust OPR outperformed standard OPR for playoff alliance score prediction.

Another way to look at this question is to see how playoff scoring accuracy is impacted by the percentage of teams whose OPR was significantly changed by using robust OPR compared to standard. For this analysis, a significant change was defined as 4 OPR units, which is about 15% of the average OPR for all the teams at all the events.

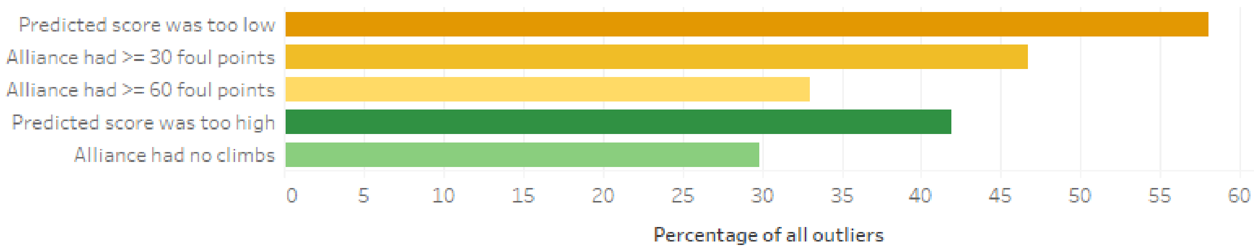
Figure 8: Effect of prediction accuracy by playoff team change in OPR



As would be hoped, in cases where more playoff teams had significantly different robust OPR compared to standard OPR, improvements in playoff alliance score predictions were seen.

It may be of interest to determine the nature of alliance performances that were identified as outliers. One thing to note is that for 2020 Palmetto, one of the two alliance performances identified as outliers was qualifying match 9 red alliance, mentioned in the introduction, which involved the highest-ranked robot 694 not moving at all. This is a promising sign for the robust OPR concept. Figure 9 shows some major scenarios involving outliers at all events.

Figure 9: Breakdown of alliance performances identified as outliers



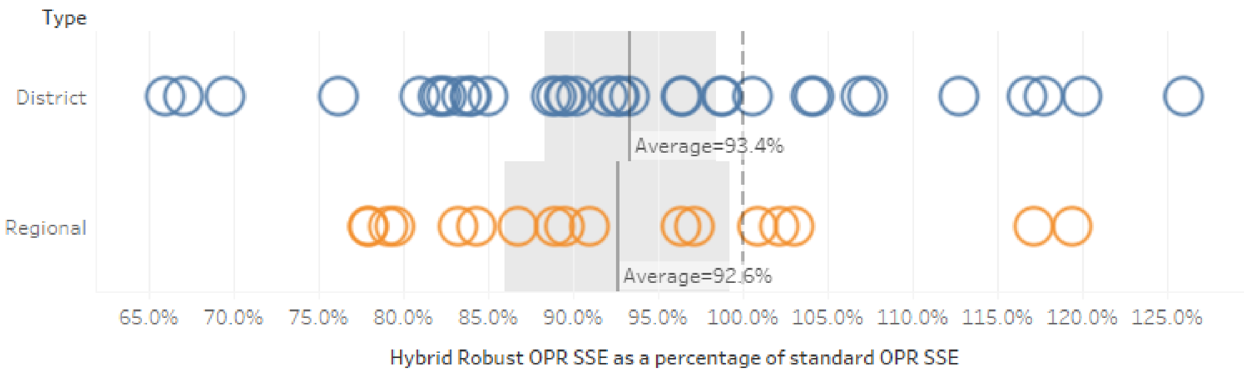
Of all outliers identified, 58% of the time the predicted score was too low. This suggests that some large, unexpected, positive event occurred for those alliances. A positive event that might be unexpected would be many foul points. Of all the outlier cases with low predictions, 80% (46.6/58) of those had 30 or more foul points. There were 57% (33/58) that had 60 or more foul points.

For the 42% of cases where the predicted score was too high, the idea is that there would be some event causing a large, unexpected, negative impact to the actual score. A major negative event that might be unexpected would be no climbs. Because climbing has the potential for so many points, if an alliance that had reliable climbing robots failed to get any climb points, it would be a major negative deviation. It turns out that of all the outlier cases with high predictions, 71% (30/42) of time there were no robots that climbed for the alliance.

3.3 Evaluating hybrid robust OPR

Hybrid robust OPR is the case where hybrid OPR in (7) and (8) is computed with outlier-filtered (robust) alliance sub scores as well as robot-specific contributions determined from FMS data. Figure 10 shows playoff score prediction performance with hybrid robust OPR.

Figure 10: Hybrid robust OPR SSE as a % of standard OPR SSE by event and type



Each circle is one event. Hybrid robust OPR results in playoff alliance score predictions that are approximately 7% better than standard OPR on average. District and regional predictions are improved by similar magnitudes. This makes sense based on previous results because the effects that caused regionals to be predicted better by hybrid OPR and districts to be predicted better by robust OPR are now combined. Considering all events together, Figure 11 shows that the 95% confidence interval on the average improvement of playoff predictions of hybrid robust OPR over standard is not anywhere close to including 100% which represents parity of the two metrics. The prediction improvement of hybrid robust OPR is extremely significant statistically.

Figure 11: Hybrid robust OPR SSE as a percentage of standard OPR SSE by event

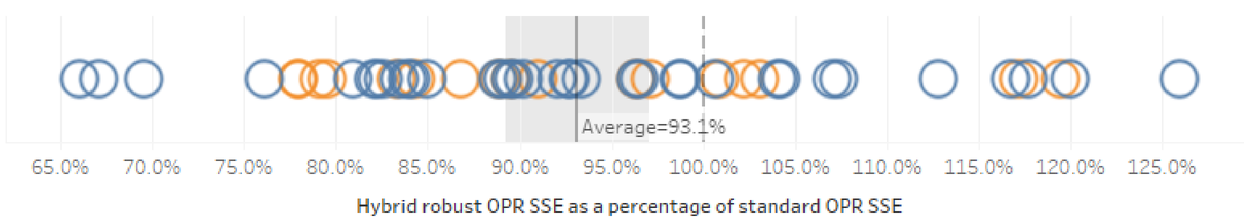


Table 2 shows the comparison of standard, robust, and hybrid robust OPR to average point contributions from 4020 scouting at the 2020 Palmetto regional. Recall from the previous hybrid OPR section that foul points are not included in any of these OPRs to be consistent with the available scouting data, which did not record fouls received, only fouls committed.

Table 2: Comparing error between 2020 Palmetto scouting and robust OPR forms

Error Computation	**Sum of squared errors**
Standard OPR vs Scouting	2860
Robust OPR vs Scouting	2585
Hybrid robust OPR vs Scouting	1007

Robust OPR improves agreement with scouting data slightly, but not nearly as much as hybrid robust OPR. Because the robust technique removes a very limited number of outlier performances (2 in the case of 2020 Palmetto), most OPRs are not significantly affected and the select few "fixes" result in a modest aggregate improvement in agreement with scouting. Hybrid OPR, on the other hand, improves accuracy for all teams, resulting in a much greater agreement with scouting of the entire event.

4 Study of split-foul OPR

In each of the last five competition seasons, a given alliance score is increased by fouls committed by an opponent alliance. As a result, a portion of standard OPR comes from this action of other robots. It is debatable to think of receiving foul points as an "offensive" action as suggested by the term offensive power rating. If you have watched many matches you have invariably seen a robot playing defense poorly or repeatedly violating some rule and just handing out free points to opponents like Mardi Gras trinkets. In this situation, these windfall points are not due to the action of the robots receiving them and it is unlikely to receive similar gifts in future rounds. When OPR is influenced by these kinds of points, it contributes to an inaccurately high value. On the other hand, there are some robots, by the nature of their design or their gameplay strategy, which are adept at drawing foul points. In this situation, there is a relatively high probability that foul points would be received with some consistency from round-to-round and that an OPR which includes those points would help make the OPR more appropriate for future score prediction.

In this study, we will examine whether playoff alliance score prediction is more accurate when foul points are accounted for in the standard way as a component of OPR (such that alliance score predictions from OPR are independent of the choice of opponent alliance), or whether splitting-out the foul points from OPR and predicting alliance scores by combining "no-foul" OPRs with fouls-committed component OPRs for the opponent alliance robots is a better approach.

4.1 Computing split-foul OPR

A definition of no-foul OPR for the hybrid case where some of the scoring contributions are known exactly from the FMS is already provided in (8). Similarly, in the case of standard OPR or robust OPR, the no-foul cases of these are computed using (1) where the $\boldsymbol{b}$ vector is comprised of the alliance score minus the alliance foul points received.

$$b_i = S_i - f_i \tag{16}$$

To compute the component OPR for fouls committed, the $\boldsymbol{b}$ vector is comprised of the negative value of foul points received by the opposing alliance in the match records. I define the values as negative so the fouls-committed component OPR is consistent with other components. A higher component OPR value is better and lower is worse.

$$b_i = -f_{i,opponents} \tag{17}$$

For instance, if a portion of a match record looks like this:

Table 3: Example match record excerpt for split-foul OPR calculation

Blue team1	Blue team2	Blue team3	Blue score	Blue foulPts	Red team1	Red team2	Red team3	Red score	Red foulPts
6	1	3	100	3	2	5	4	155	30

then the corresponding portion of the linear system to be solved for no-foul OPR would be:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & \cdots \\ 0 & 1 & 0 & 1 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} OPR_{nf,1} \\ OPR_{nf,2} \\ \vdots \end{bmatrix} = \begin{bmatrix} 97 \\ 125 \\ \vdots \end{bmatrix} \quad (18)$$

The corresponding portion of the linear system to be solved for fouls-committed component OPR would be:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & \cdots \\ 0 & 1 & 0 & 1 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} cOPR_{fc,1} \\ cOPR_{fc,2} \\ \vdots \end{bmatrix} = \begin{bmatrix} -30 \\ -3 \\ \vdots \end{bmatrix} \quad (19)$$

4.2 Evaluating split-foul OPR

The predicted alliance score using a split-foul OPR approach is the sum of no-foul OPRs for the teams in the alliance minus the sum of the fouls-committed component OPRs of the opponent alliance.

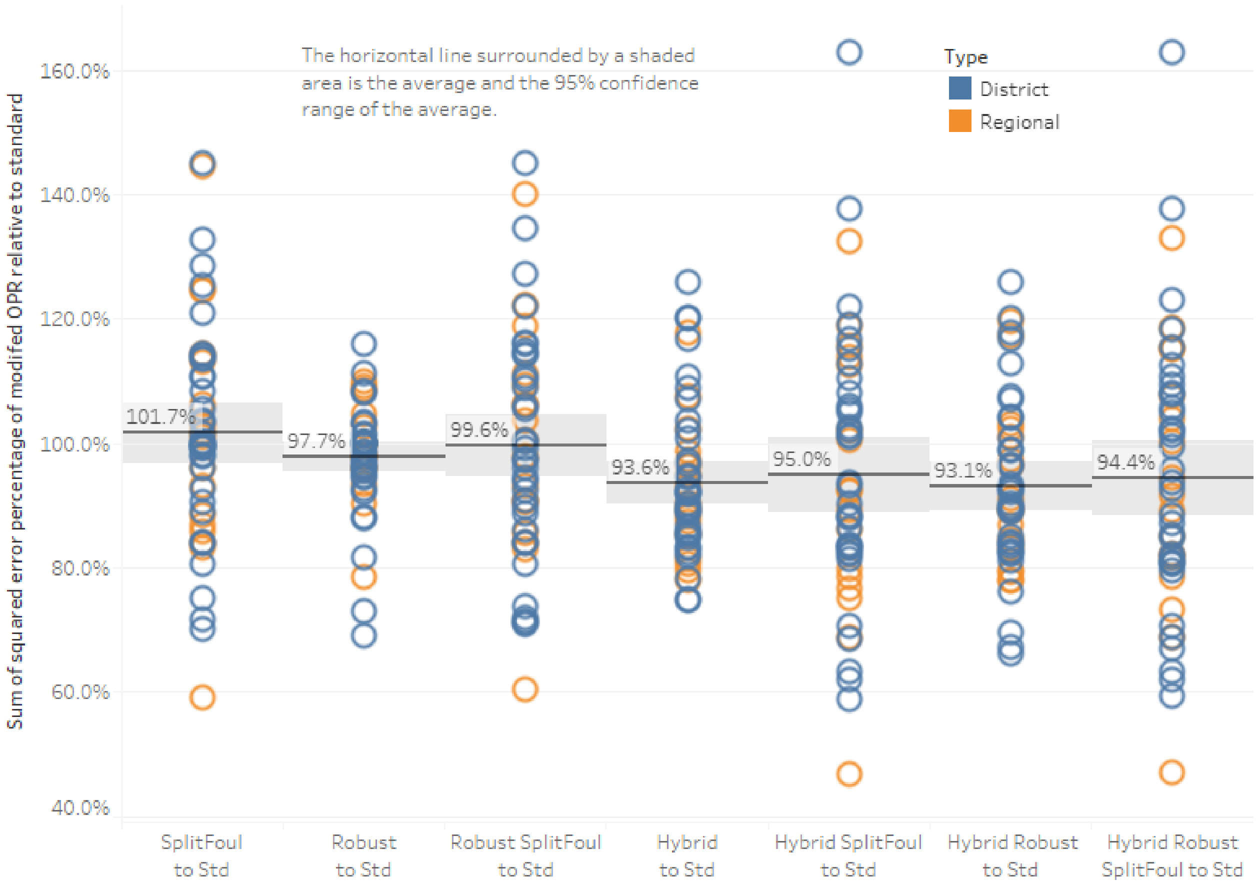
$$\hat{S}_a = OPR_{nf,a1} + OPR_{nf,a2} + OPR_{nf,a3} - cOPR_{fc,o1} - cOPR_{fc,o2} - cOPR_{fc,o3} \quad (20)$$

The predicted alliance score using a standard OPR approach is already provided in (9).

Computation of prediction error metrics of playoff matches with split-foul OPR is performed with the same weighting approach described in section 2.3.

Figure 12 shows how the sum of squared errors of playoff alliance score predictions compare between using score calculations with split-foul OPR (no-foul OPR minus opponent fouls-committed component OPR) and score calculations where OPR includes foul points received.

Figure 12: Split-foul OPR SSE comparisons to other forms of OPR



The first column of marks in Figure 12 shows how the split-foul prediction error compares to standard OPR. Split-foul results in predictions that are less accurate by 1.7% using an SSE metric. The 2nd and 3rd columns of marks show what happens if you use split-foul with robust OPR. As shown earlier, robust OPR with fouls included results in a 2.3% accuracy improvement on average compared to standard OPR. However, if the split-foul computation is used with the outlier-free (robust) match dataset, the accuracy becomes 1.9% worse as compared to robust OPR although it is still better than standard OPR. Note also how the error variation increases for individual events in the split-foul case. Score predictions are already noisy due to the error distributions of the estimated OPRs. Apparently, adding noise from three opponent fouls-committed component OPRs overwhelms the potentially lower noise of the no-foul OPRs such that the total score prediction variation increases.

The split-foul approach results in similar reductions in accuracy and increases in variation for the hybrid OPR and robust hybrid OPR techniques. Thus, for predicting playoff scores in the Infinite Recharge game or likely any minor modifications of it, OPRs with foul points included are the best bet. It should not be assumed that this conclusion would hold for other games where fouls are accrued differently or have different magnitudes relative to “offensive” scoring. A study like this one could be done for other competition seasons, but that is beyond the scope of this paper.

5 Study of other error metrics

Sum of squared error for alliance playoff score prediction is a good way to measure the performance of any form of OPR. However, there are other ways that may also be of interest. In this section, sum of squared error for playoff match score differential (or margin) as well as mean prediction error will be compared between forms of OPR.

5.1 Comparing forms of OPR by SSE of playoff match margin

Predicting the difference between the scores of the competing alliances in a match is another way that OPR could be used. Hybrid and robust OPR as described in this paper reduce the SSE of individual alliance score predictions. Do they also reduce the SSE of match margin prediction?

Margin is the difference between alliance scores in a match. As a convention, define the margin to be the score of the red alliance minus the score of the blue alliance.

$$M_{r,b} = S_r - S_b \tag{21}$$

The predicted margin is the difference in the OPRs between the alliance teams.

$$\widehat{M}_{r,b} = OPR_{r1} + OPR_{r2} + OPR_{r3} - OPR_{b1} - OPR_{b2} - OPR_{b3} \tag{22}$$

Weighting for margin is slightly different than weighting for alliance scores. For margin, the weight is the inverse of the number of playoff matches played by unique alliance pairs.

$$w_{r,b} = \frac{1}{m_{r,b}} \tag{23}$$

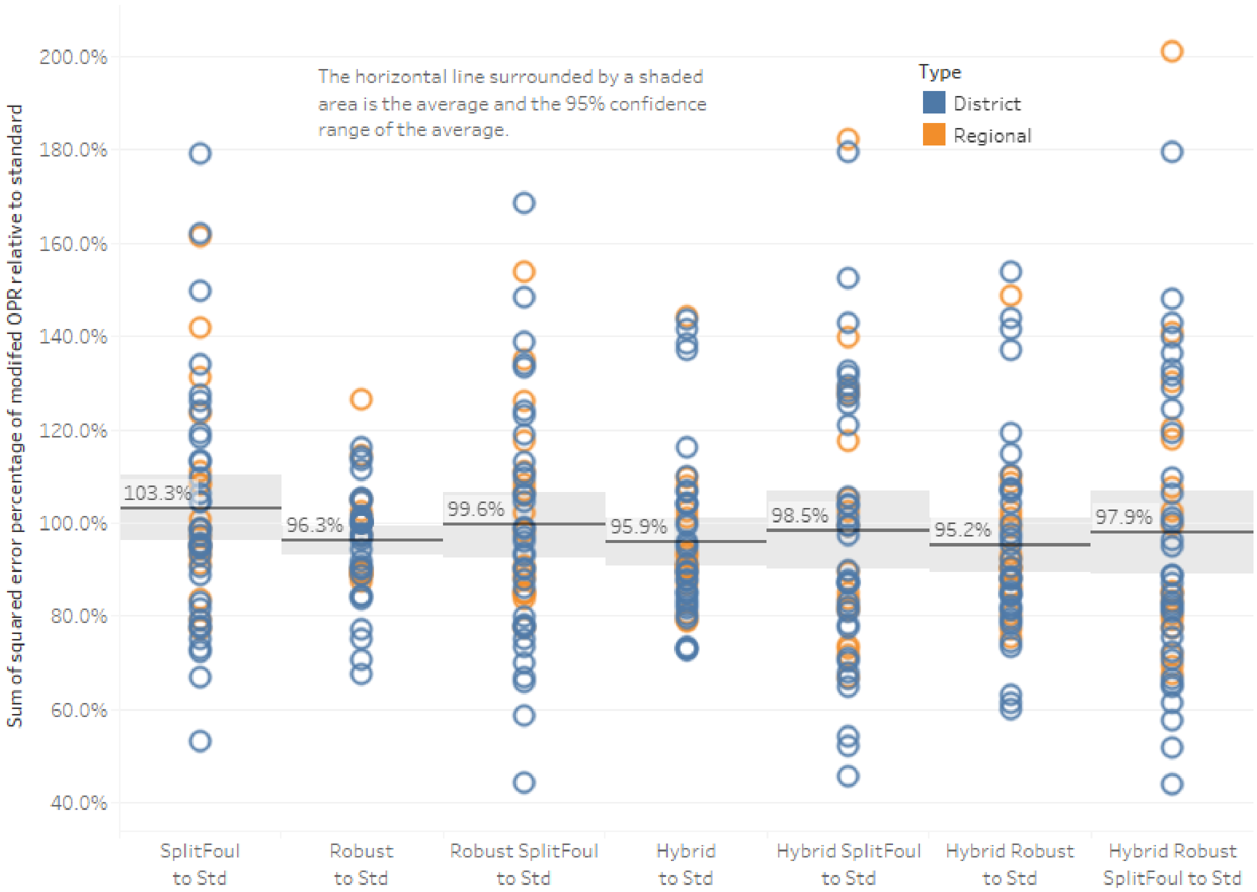
The error metric for margin becomes:

$$SSE = \sum_{i=1}^{p} w_{ri,bi}\left(\widehat{M}_{ri,bi} - M_{ri,bi}\right)^2 \tag{24}$$

where p is the number of playoff matches.

Figure 13 shows the comparison of all the alternate OPR forms previously discussed as measured by the SSE of playoff score margin prediction relative to the SSE of playoff score margin predicted by standard OPR.

Figure 13: Comparison of alternate OPR forms by SSE of playoff score margin predictions



Both robust and hybrid OPR yield about 4% improvement over standard OPR for playoff margin prediction error. By combining the two techniques, almost 5% improvement over standard OPR is achieved. The split-foul forms of OPR for margin prediction do not perform as well as the foul-included forms, whether using standard, robust, hybrid, or hybrid robust OPR. The prediction error is higher and the variability of prediction error by event is higher with split-foul.

5.2 Comparing forms of OPR by mean error of playoff alliance score

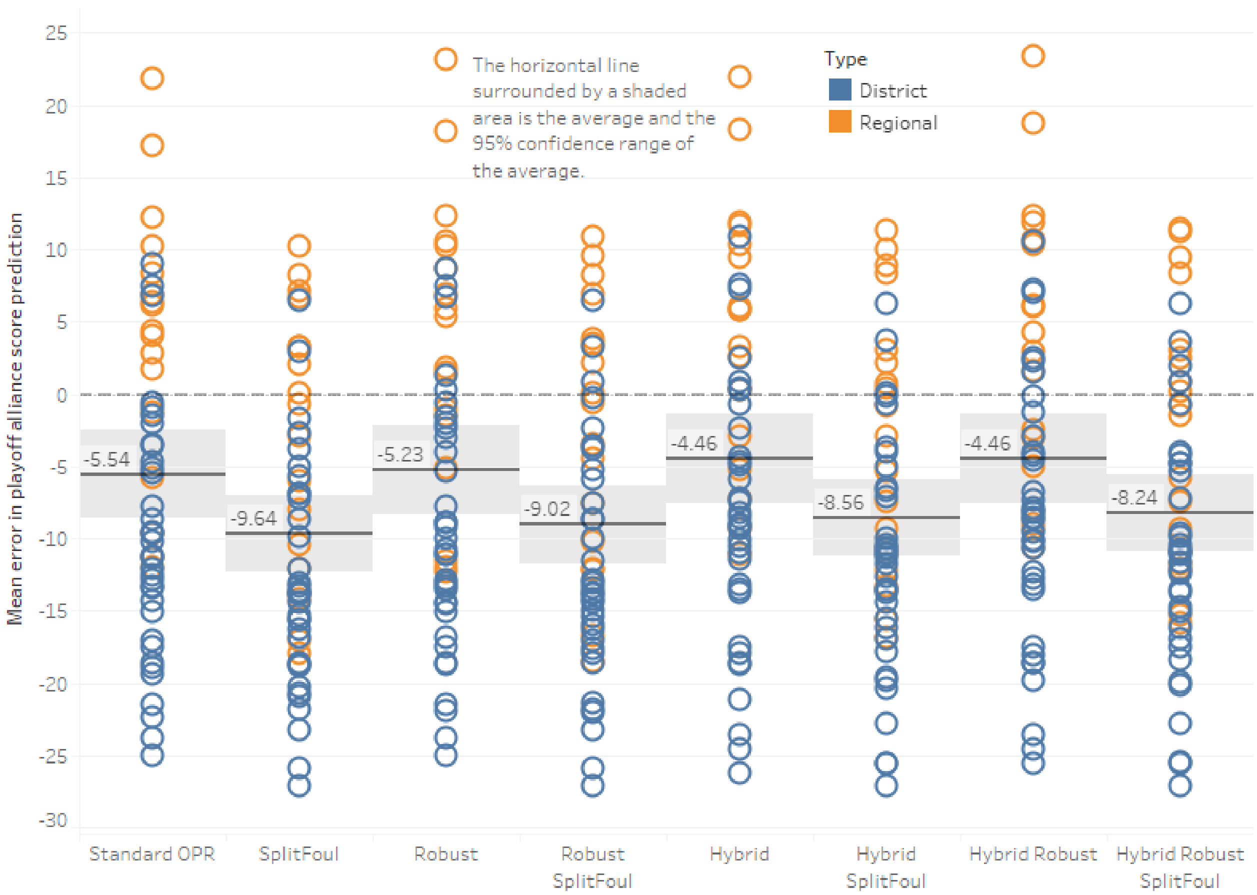
Examining the mean of playoff score prediction errors can help reveal if alternate forms of OPR are able to reduce bias in score predictions. That is, are scores being consistently predicted higher or lower than the actual scores? The mean alliance score prediction error for an event is:

$$\bar{E} = \frac{\sum_{i=1}^{2p}(\hat{S}_i - S_i)}{2p} \tag{25}$$

where p is the number of playoff matches. Each match has two alliance scores.

Figure 14 shows the comparison of all the alternate OPR forms previously discussed as measured by the mean error of the playoff alliance score prediction relative to the mean error of playoff score predicted by standard OPR.

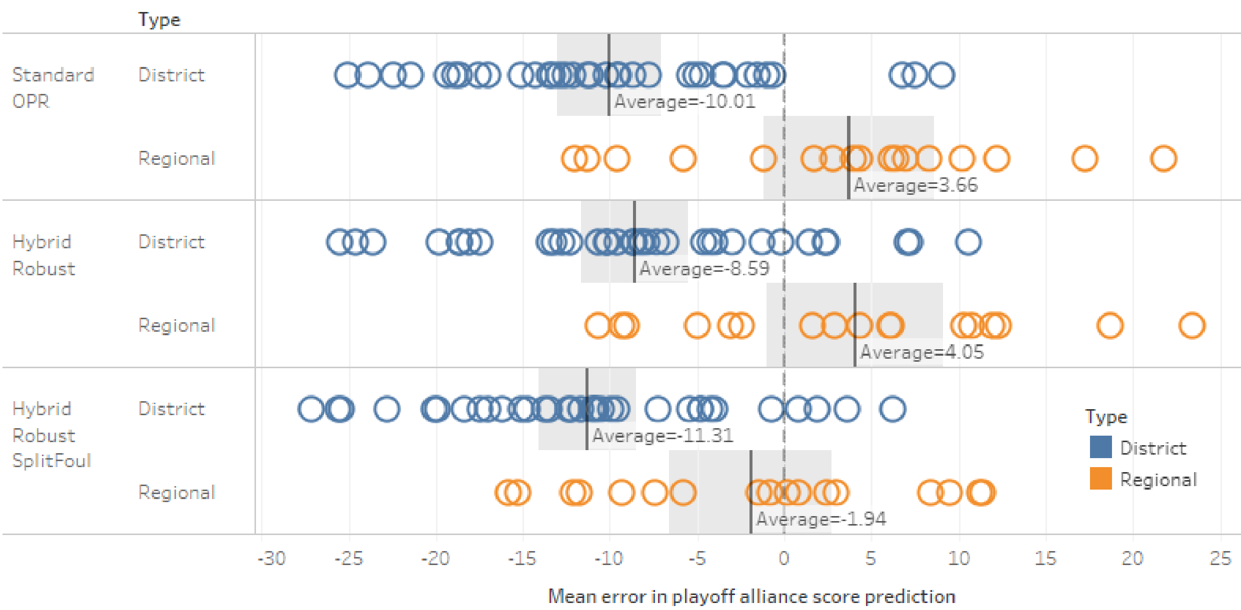
Figure 14: Comparison of alternate OPR forms by mean error of playoff score predictions



The first thing to note is that robust, hybrid, and robust hybrid OPR forms very slightly reduce the underprediction of playoff scores compared to standard OPR. These improvements are not statistically significant, though. Using split-foul OPRs causes playoff alliance scores to be even more underpredicted, regardless of form of OPR. Moving to robust, hybrid, and robust hybrid OPR split-foul forms also slightly reduces the split-foul underprediction as well.

Perhaps the most interesting observation is that across all the forms shown, the prediction error is much more negative for district events as compared to regional events. Figure 15 shows district and regional results separately for several forms of OPR, highlighting the difference in prediction error between the two types of events.

Figure 15: Mean error of playoff score prediction by event type for several OPR forms



At district events, actual playoff scores appear to get a boost that no form of OPR can predict and that is also not occurring at regionals. The difference persists across the various forms of OPR studied. Predicting scores by splitting out fouls shifts the grand mean from positive to negative for regionals, but it also makes districts more negative so the difference between the event types remains. In just looking for patterns that might cause this, nothing jumped out as a clear cause. It would probably take a modelling exercise to find the source of the difference.

6 Summary / Conclusions

Hybrid OPR, a technique combining FMS-supplied per-robot data with component OPRs for scoring that is not uniquely identified by FMS, results in a greater than 6% decrease in the sum of squared error for playoff alliance score predictions in all 2020 FRC events as compared to standard OPR. It also reduced SSE of playoff match scoring margin by over 4% and had a slightly beneficial effect on mean error of playoff alliance scores. For the 2020 Palmetto regional event, where my team, 4020, had scouting data, hybrid OPR reduced the sum of squared error between OPR and scouting estimated point contribution by a factor of 3 across all teams as compared to standard OPR. 4020's hybrid OPR at Palmetto was 61.3 versus standard OPR of 47.9. The difference was almost entirely due to replacing a highly inaccurate standard endgame component OPR of 8.15 with a known average endgame contribution of 21.39 based directly on FMS data. Hybrid OPR could deliver differing amounts of improvement in future FRC seasons depending on how much robot-specific data is available from the FMS for the future games. Assuming that the 2021 game is very similar to the 2020 game, hybrid OPR should be an effective improvement in 2021.

Robust OPR, a technique that identifies and removes a very limited number of outlier alliance performances before computing OPR in the standard way, reduced SSE of playoff alliance score predictions by 2.3% in the 2020 FRC events compared to standard OPR. It also reduced SSE of playoff match scoring margin by 3.7% and had a slightly beneficial effect on mean error of playoff alliance scores. Agreement with 2020 Palmetto scouting data was improved by a modest 10% across all teams compared to standard OPR. 4020's robust OPR at Palmetto was 59.5 versus standard OPR of 47.9 and a scouting estimate of 67.0. The improved agreement with scouting came because one of the two alliance performances at that event identified as outliers by the robust technique was qm9 red alliance. This performance involved 4020, 1708, and 694, where 694 with a standard OPR of 91.3 did not move in the match and contributed 0 points.

Combining the hybrid and robust techniques in OPR calculation is synergistic and results in greater improvement in playoff predictions than either technique separately. SSE of playoff alliance score predictions was reduced by almost 7% with hybrid robust OPR compared to standard OPR and the average improvement was nearly identical for district and regional events. SSE of playoff match scoring margin was reduced by almost 5%. Mean error of playoff alliance scores was still only slightly improved compared to standard OPR. The hybrid robust OPR for team 4020 at 2020 Palmetto regional was 69.2, compared to standard OPR of 47.9 and an estimated scouting contribution of 67.0. The hybrid technique fixed a major inaccuracy in the standard OPR endgame estimate and the robust technique fixed an outlier impacting OPR where the best team at the event scored 0 points in a qualification match involving 4020.

Splitting out foul points such that a playoff score prediction is the sum of no-foul OPRs for a given alliance and fouls-committed component OPRs for the opponent alliance was detrimental to playoff score prediction for the 2020 events. All forms of OPR studied had poorer playoff score predictions with split fouls compared to use of OPR with foul point contributions included. This study did not include seasons other than 2020, so the result that a split-foul approach results in poorer playoff score predictions should not be assumed to hold for other seasons or games. It is likely to apply to the 2021 season if game changes are minimal.

7 Supporting Content

Tableau Public workbooks with figures:

Robust Hybrid OPR FRC Main

Robust Hybrid OPR FRC Scouting

Robust Hybrid OPR FRC Outlier Opt

Results datasets in .csv format

- Performance of all forms of OPR on 2020 event playoff predictions by several error measures
- Playoff score predictions for 2020 events using all forms of OPR
- Qualification match performances identified as robust OPR outliers for 2020 events
- All forms of event OPRs and component OPRs

Python code is provided to compute the forms of OPR developed in the paper. The files read data for completed or underway events from The Blue Alliance API and generate an output file with the OPR results.

There are two files:

robustOpr.py

This file only computes standard and robust OPR and intended to be season-independent.

robustHybridOpr2020.py

This file is intended for use with match data from the 2020 season only because things like component OPR, no-foul OPR, and hybrid OPR are season-specific. Changes for the 2021 season will likely be minor. Modifications for future seasons are likely more extensive but expected to be straightforward.

Output includes:

- Standard OPR
- Robust OPR - OPR on a match dataset with outlier performances removed
- Hybrid OPR - OPR with some scoring components known exactly from FMS data
- Robust Hybrid OPR - combination of the two techniques above
- No-foul versions of all four forms above
- Component OPR for a select set of scoring components

- Robust component OPR - same as component, but with outlier performances removed

Both files require an FRC-standard event key to identify the event for which to compute OPRs. One easy way to find an event key is to search for a desired event on The Blue Alliance and the URL of the main event page includes the event key at the end. For example, the TBA URL for the 2020 Palmetto Regional is https://www.thebluealliance.com/event/2020scmb and the event key is 2020scmb.

Since the files acquire data from The Blue Alliance API, a TBA API key is required for access. The API key is free but must be requested from your TBA account.

A tba_key.json file template is provided for use. You can paste your TBA API key into this file for easy, repeated access by the Python scripts. Alternatively, the key can be optionally passed as an argument.

python robustOpr.py eventKey [tbaAPIKey]

python robustHybridOpr2020.py eventKey [tbaApiKey]

Output is provided as:

- robustOpr_*eventKey*.csv
- robustHybridOpr2020_*eventKey*.csv

8 License