An Overview and Analysis of Statistics used to Rate FIRST Robotics Teams

William Gardner

cheer4ftcblog@gmail.com

V6 : 2016-05-14

Overview:

This paper presents and analyzes a wide range of statistical techniques that can be applied to FIRST Robotics Competition (FRC) and FIRST Tech Challenge (FTC) tournaments to rate the performance of teams and robots competing in the tournament.

The well-known Offensive Power Rating (OPR), Combined Contribution to Winning Margin (CCWM), and Defensive Power Rating (DPR) measures are discussed and analyzed.

New measures which incorporate knowledge of the opposing alliance members are discussed and analyzed. These include the Winning Margin Power Rating (WMPR), the Combined Power Rating (CPR), and the mixture-based Ether Power Rating (EPR).

New methods are introduced to simultaneously estimate separate offensive and defensive contributions of teams. These methods lead to new, related simultaneous metrics called sOPR, sDPR, sWMPR, and sCPR.

New MMSE estimation techniques are introduced. MMSE techniques reduce overfitting problems that occur when Least Squares (LS) parameter estimation techniques are used to estimate parameters on a relatively small data set. The performance of LS and MMSE techniques is compared over a range of scenarios.

All of the techniques are analyzed over a wide range of simulated and actual FRC tournament data, using results from the 2013, 2014, and 2015 FRC seasons.

Table of Contents

Introduction

FIRST is a non-profit organization that sponsors four levels of robotics competitions for youth: FIRST Robotics Competition (FRC), FIRST Tech Challenge (FTC), FIRST Lego League (FLL) and Junior FIRST Lego League (JrFLL).

In FRC and FTC, teams build robots and compete in matches where alliances of multiple robots face off against each other. In FRC the alliances have 3 teams and in FTC the alliances have 2 teams. The remainder of the discussion in this paper focuses on the analysis of FRC 3 v. 3 matches but the analysis can be easily modified to accommodate FTC 2 v. 2 matches.

During a tournament, over a season, and from season to season, teams play many matches. Many in the FIRST community are interested in computing statistics to measure the absolute and relative performance of different robots and different teams. This paper describes many different methods that can be used to model performance and estimate model parameters. The paper also evaluates the performance of different methods in different scenarios.

Notation and Conventions

Scalars are shown italicized like *R*. Vectors and matrices are shown in bold like **O**. Vector and matrix transpose is denoted with a trailing apostrophe " ' " like **O**'.

Measurements of the offensive contribution of a team are indicated by some form of *O* or **O**. Offensive contributions usually have an average, *Oave*, of ⅓ of the average match score in a tournament.

Measurements of the defensive contribution of a team are indicated by some form of *D* or **D**. Defensive contributions described in this paper *usually* have an average, *Dave*, of 0, and *usually* a positive contribution corresponds to a positive outcome for a team (i.e., a *D* of +5 means that a team's defensive contribution is 5 points better than average).

Measurements of the contribution of a team to its alliance's overall Winning margin (i.e., their alliance's score minus their opposing alliance's score) are indicated by some form of *W* or **W**. Winning margin contributions usually have an average, *Wave*, of 0.

Measurements of the Combined contribution of offense and defense of a team to its alliance's performance are indicated by some form of *C* or **C**. Combined contributions in this paper are usually a measure of **O** + **D**, and as a result Combined contributions usually have an average *Cave* = *Oave*. Combined contributions are often related to the Winning margin contributions by the relationship **W** = **C** - *Oave*.

Parameter estimates (scalar or vector) are underlined, like $\underline{\mathbf{O}}_{ave}$.

MMSE-based parameter estimates (scalar or vector) are "hatted," like $\widehat{O}$.

Linear Match Models

Full Model Equation for Offense

Each FRC match produces two outcomes: the Red alliance final score and the Blue alliance final score[1]. In most seasons, the alliance with the larger score is also declared the match winner.

The linear match model used in this paper is:

$$R = (O_i + O_j + O_k) - (D_l + D_m + D_n) + N_r$$

$$B = (O_l + O_m + O_n) - (D_i + D_j + D_k) + N_b$$

Teams i, j, and k are on the Red alliance and teams l, m, and n are on the Blue alliance. *Oi* is the offensive contribution of team i and *Di* is the defensive contribution of team i. In both cases, a positive value is beneficial to a team: a larger positive value of *Oi* for team i means that their alliance will score more points, and a larger positive value of *Di* means that their opposing alliance will score fewer points. *R* and *B* are the final Red and Blue alliance scores.

Nr and *Nb* are noise. These noise terms can model either match-to-match variations in noise or noise due to nonlinearities in actual match play. The match noise is modeled in this paper as having constant variance from match to match. This can be viewed either as variation due to individual match randomness or as constant variation produced by teams themselves (e.g., drivers not driving exactly the same way, etc.). It may be of interest to model the match noise as being produced by the teams and the team-based match noises NOT all having the same variability (e.g., team 1 always scores 1000 points or 0 points while team 2 always scores 10 points or 0 points). This paper does not address this question.

In vector-matrix form, the above equations for teams 1-6 can be written as

---

[1] Outcomes on parts of the matches may also be available, like Red and Blue autonomous scores, teleop scores, end-game scores, penalty scores, etc. Most of the analyses in this paper can also be applied individually to these scores as if they were final match scores, and can be used to evaluate the performance of teams in these subareas as well. The FTC android apps “Watch FTC Tournament” and “FTC Online” do exactly that for FTC tournaments.

$$\begin{bmatrix} R \\ B \end{bmatrix} = \left[\begin{smallmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{smallmatrix}\middle|\begin{smallmatrix} 0 & 0 & 0 & -1 & -1 & -1 \\ -1 & -1 & -1 & 0 & 0 & 0 \end{smallmatrix}\right] \begin{bmatrix} O_1 \\ O_2 \\ \ldots \\ O_6 \\ - \\ D_1 \\ D_2 \\ \ldots \\ D_6 \end{bmatrix} + \begin{bmatrix} N_r \\ N_b \end{bmatrix}$$

This form can be expanded for a full tournament with m matches and t teams. Define **Ar** as the m x t matrix with i,jth element equal to 1 if team j is on the Red alliance in match i and 0 otherwise, **Ab** as the m x t matrix with i,jth element equal to 1 if a team is on the Blue alliance in match i and 0 otherwise, **Mr** as the m-length vector of the Red alliance scores, and **Mb** as the m-length vector of the Blue alliance scores. Then,

$$\begin{bmatrix} \mathbf{M}_r \\ \mathbf{M}_b \end{bmatrix} = \left[\begin{matrix} \mathbf{A}_r \\ \mathbf{A}_b \end{matrix}\middle|\begin{matrix} -\mathbf{A}_b \\ -\mathbf{A}_r \end{matrix}\right] \begin{bmatrix} \mathbf{O} \\ \mathbf{D} \end{bmatrix} + \begin{bmatrix} \mathbf{N}_r \\ \mathbf{N}_b \end{bmatrix}$$

or

$$\mathbf{M}_o = \left[\mathbf{A}_o \middle| -\mathbf{A}_d\right] \begin{bmatrix} \mathbf{O} \\ \mathbf{D} \end{bmatrix} + \mathbf{N}_o$$

“Full Model Equation for Offensive Scores”

Mo is the full vector of offensive match scores, **Ao** is the matrix describing which teams played offense producing a given match score (with an i,jth element of 1 if team j was on offense leading to match score i), **Ad** is the matrix describing which teams played defense producing a given match score (with an i,jth element of 1 if team j was on defense leading to match score i), **O** is the vector of team offensive contributions, **D** is the vector of team defensive contributions, and **No** is the full vector of offensive match noise values. **Mo** and **No** are 2m-length vectors, **Ao** and **Ad** are 2m x t matrices, **O** and **D** are t-length vectors.

Given a model with values of **O** and **D**, exactly the same match outcomes are produced with the related model **O**+*K* and **D**+*K*, where *K* is an arbitrary scalar constant. As a matter of convention, this paper normalizes models like this or estimates of the parameters **O** and **D** by computing the mean of all of the elements of **D** and then subtracting that mean from all elements of **O** and **D**. This produces the interpretation of elements of **O** as the expected contribution a team makes to its alliance's match scores when playing against average defense, and the interpretation of elements of **D** as the expected contribution a team makes against its opposing alliance's match scores, relative to the average defense. This is discussed further in later sections.

Full Model Equation for Winning Margin

Similarly, a linear model for the winning margin *R-B* can be formed by taking the difference of the previous equations as

$$R - B = ((O_i + D_i) + \dots + (O_k + D_k)) - ((O_l + D_l) + \dots + (O_n + D_n)) + N_r - N_b$$

If a team's contribution to their winning margin is defined as $Wi = (Oi - Oave) + Di$, then the expression can be similarly rewritten as:

$$\begin{bmatrix} R - B \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} W_i \\ W_j \\ \dots \\ W_n \end{bmatrix} + \begin{bmatrix} Nr - Nb \end{bmatrix}$$

The *Oave* terms are included to make **W** have zero mean, and they cancel each other out in the above equation. For a full tournament, the matrix-vector equivalent equation becomes

$$\mathbf{M}_r - \mathbf{M}_b = \left(\mathbf{A}_r - \mathbf{A}_b\right)\mathbf{W} + \left(\mathbf{N}_r - \mathbf{N}_b\right)$$

or

$$\mathbf{M}_w = \mathbf{A}_w \mathbf{W} + \mathbf{N}_w$$

"Full Model Equation for the Winning Margin"

Mw = **Mr** - **Mb** is the vector of winning margins with elements that are positive if Red won a match and negative if Blue won. **Aw** = **Ar** - **Ab** is a matrix describing which teams played in a given match (with i,jth element of 1 if team j was on Red in match i, -1 if team j was on Blue in match i, and 0 otherwise), **W** is the vector of winning margin contributions, and **Nw** = **Nr** - **Nb** is the vector of winning margin noise values. If **Nr** and **Nb** are independent and identically distributed, elements of **Nw** have twice the variance of elements of **No**.

Mw and **Nw** are m-length vectors, half the size of the corresponding **Mo** and **No** vectors. **Aw** is an m x t matrix, half the size of the corresponding **Ao** and **Ad** matrices.

Measures of Model Performance

FIRST teams are interested in statistics that can be used for the following purposes:

- Help scout upcoming matches during a tournament.
- Help select alliance members for elimination rounds.
- Compare teams across different tournaments and different seasons.
- Help scout teams prior to a tournament.
- Predict match outcomes prior to a match.

To achieve these purposes, the FIRST community has developed a number of statistical estimation methods to estimate future match outcomes **Mo** and **Mw** and to estimate teams' offensive, defensive, and winning margin contributions, **O**, **D**, and **W**.

Some ways of measuring the performance of a model are:

1. Compute the difference between the actual offensive match scores and the match scores predicted by a given model. Usually this is measured with unweighted squared error as: $Eo = (\mathbf{Mo} - \underline{\mathbf{Mo}})' (\mathbf{Mo} - \underline{\mathbf{Mo}})$. **Mo** is the measured vector of offensive match scores and $\underline{\mathbf{Mo}}$ is the predicted match scores from the model, and **X**' is the transpose of **X**.

2. Compute the difference between the actual winning margins and the winning margins predicted by a given model, as: $Ew = (\mathbf{Mw} - \underline{\mathbf{Mw}})' (\mathbf{Mw} - \underline{\mathbf{Mw}})$.

3. Compute the probability that a given model correctly predicts the outcome of a match (i.e., whether Red scored more than Blue or vice versa). This is linearly related to the following expression: $Es = (\mathrm{sign}(\mathbf{Mw}) - \mathrm{sign}(\underline{\mathbf{Mw}}))' (\mathrm{sign}(\mathbf{Mw}) - \mathrm{sign}(\underline{\mathbf{Mw}}))$, where sign(**X**) has elements that are +1 if Xi is positive and -1 if Xi is negative.

4. For simulated tournaments where the underlying model parameters are known, compute the difference between the estimated model parameters and the actual underlying model parameters. For example, given a vector of parameters **O** and an estimate of the model parameters $\underline{\mathbf{O}}$, measure the error $Ep = (\mathbf{O} - \underline{\mathbf{O}})' (\mathbf{O} - \underline{\mathbf{O}})$.

The results section of this paper usually normalizes these measures by the variance of the original vector being predicted, and converts the result into a percentage.

Estimating Model Parameters Using Averages

Offense - Average of Match Scores

One simple way to measure a team's offensive power is to simply average their alliance's match scores. In matrix form, this can be written as:

$$\underline{\mathbf{O}}_{ave} = \left(diag(\mathbf{A}_o'\mathbf{A}_o)\right)^{-1}\mathbf{A}_o'\mathbf{M}_o$$

Ao' Mo is just the vector of the sum of the match scores for each team, and the diagonal matrix in front of it has i,ith elements equal to 1/# of matches played by team i.

Note that this measure is a vector with ith element equal to the average match score of the ith team, as opposed to *Oave* which is the scalar average of *all* match scores for *all* teams.

Limiting Behavior

If the data were truly generated using the model equation for offense and if a team plays an equal number of matches with and against the other teams, as the number of matches played and the number of teams participating get large, this metric approaches:

$$\underline{\mathbf{O}}_{ave} \to \mathbf{O} + 2O_{ave} - 3D_{ave} = \mathbf{O} + 2O_{ave}$$

The final equation holds because the elements of **D** have been normalized to have a mean of zero as previously discussed.[2]

An estimate using the average of match scores that approaches a team's true underlying contribution is then:

$$\underline{\mathbf{O}}_{a} \equiv \underline{\mathbf{O}}_{ave} - 2O_{ave} \to \mathbf{O}$$

Defense - Average of Opponent's Match Scores

One simple way to measure a team's defensive power is to simply average their opponent's match scores. In matrix form, this can be written as:

---

[2] If the number of matches get large but the number of teams remains fixed, then the *Oave* scalar term in the equation is replaced by a vector with the ith element equal to the average of the elements of **O** of all of the teams *except* team i, which can usually be well approximated by *Oave* unless a team is a large outlier. As an example, in a tournament with 50 teams where team 1 always scores 100 points and the other 49 teams always score 0 points, the true *Oave* of all teams is 2 points, but the average of the elements of **O** for all teams except team 1 is 0 points and the comparable average for all of the other teams is 100/49.

$$\underline{\mathbf{D}}_{ave} = \left(diag(\mathbf{A}_d'\mathbf{A}_d)\right)^{-1}\mathbf{A}_d'\mathbf{M}_o$$

This measure is large for a team playing poor defense and small for a team playing better defense.

As in the calculation of the average of a team's match scores, **Ad' Mo** is just the vector of the sum of the opponent's match scores for each team, and the diagonal matrix in front of it has i,ith elements equal to 1/# of matches played by team i.

Limiting Behavior

If the data were truly generated using the model equation for offense and if a team plays an equal number of matches with and against the other teams, as the number of matches played and teams participating get large, this metric approaches:

$$\underline{\mathbf{D}}_{ave} \to 3O_{ave} - \mathbf{D} - 2D_{ave} = 3O_{ave} - \mathbf{D}$$

Since 3*Oave* is just the average *match* score, this metric approaches the average *match* score minus a team's defensive contribution[3].

An estimate using the average of opponent's match scores that approaches a team's true underlying defensive contribution is then:

$$\underline{\mathbf{D}}_a \equiv 3O_{ave} - \underline{\mathbf{D}}_{ave} \to \mathbf{D}$$

Winning Margin - Average of Match Scores minus Opponent's Match Scores

One simple way to estimate a team's contribution to their alliance's winning margins is to average the winning margins in the matches a team plays in, or

$$\underline{\mathbf{W}}_{ave} = \underline{\mathbf{O}}_{ave} - \underline{\mathbf{D}}_{ave}$$

Limiting Behavior

Under the same assumptions of limiting behavior as in previous sections,

$$\underline{\mathbf{W}}_{ave} \to (\mathbf{O} + 2O_{ave}) - (3O_{ave} - \mathbf{D}) = (\mathbf{O} - O_{ave}) + \mathbf{D} = \mathbf{W}$$

[3] Again, technically the scalar *Oave* should instead be a vector with the ith element equal to the average offensive contribution of all teams *except* team i, but this is usually well approximated by *Oave* unless a team is a large outlier as discussed in the previous footnote.

Estimating Model Parameters Using Least Squares Methods

This section describes the different methods that are used to estimate different sets of model parameters. Their strengths and weaknesses are evaluated and discussed in following sections.

Offense

Offensive Power Rating (OPR)

A team's average match score can be artificially increased if that team happens to play with stronger teams than average or artificially decreased if that team happens to play with weaker teams than average.

One way to correct for alliance partner strength is by finding the values of **O** in the model for **Mo** that minimize *Eo,* which is equivalent minimizing **No' No**. *This assumes that the defensive contribution **D** is either 0, much smaller than **No**, or simply incorporated in **No** as additional noise.* This is like using the simplified model for **Mo** with **D**=0 shown below:

$$\mathbf{M}_o = \mathbf{A}_o\mathbf{O} + \mathbf{N}_o$$

"Partial Model Equation for Offense"

Finding the value of **O** that minimizes *Eo* in a model where **D**=0 is a simple least squares problem, and the solution is

$$\underline{\mathbf{O}}_{opr} = \left(\mathbf{A}_o'\mathbf{A}_o\right)^{-1}\mathbf{A}_o'\mathbf{M}_o$$

Winning Margins and Combined Contributions

Combined Contribution to Winning Margin (CCWM)

The CCWM measure is computed by essentially finding the set of winning margin contributions that predict an alliance's winning margin using knowledge of the alliance's partners *but not knowledge of the alliance's opponents*. This is like using the simplified model for OPR to predict winning margins rather than offensive scores, or

$$\mathbf{M}_{w'} = \mathbf{A}_o\mathbf{W} + \mathbf{N}_o$$

"Partial Model Equation for the Winning Margin"

And where **Mw'** has the form[4]:

$$\mathbf{M}_{w'} = \begin{bmatrix} \mathbf{M}_r - \mathbf{M}_b \\ \mathbf{M}_b - \mathbf{M}_r \end{bmatrix} = \begin{bmatrix} \mathbf{M}_w \\ -\mathbf{M}_w \end{bmatrix}$$

Mw' is a 2m-length vector like **Mo**, and unlike **Mw** which is an m-length vector. Note also that **Ao** is used in the partial model equation for the winning margin, not **Aw**! So there is no knowledge of opposing alliance members used in the calculation.

The least squares solution is

$$\underline{\mathbf{W}}_{ccwm} = \left(\mathbf{A}_o^{'}\mathbf{A}_o\right)^{-1}\mathbf{A}_o^{'}\mathbf{M}_{w'}$$

This is identical to the form of the OPR solution except that the **M** vector used is different.

Winning Margin Power Rating (WMPR)

Similar to the computation for OPR, the winning margin contributions can be estimated by finding the values of **W** in the full model for **Mw** that minimizes *Ew*, which is equivalent to minimizing **Nw' Nw**. This leads to the least squares solution of

$$\underline{\mathbf{W}}_{wmpr} = \left(\mathbf{A}_w'\mathbf{A}_w\right)^{+}\mathbf{A}_w'\mathbf{M}_w$$

This is identical to the form of the OPR equation except that the pseudo-inverse of **Aw' Aw** is used instead of the normal inverse. This is necessary because **Aw' Aw** is not invertible, because any constant mean term *K* can be added to all of the elements of **W** simultaneously to produce the same set of predicted winning margins. *Or, the set of WMPR values for a given tournament always has an arbitrary mean offset.* While the mean term of all of the WMPRs can be arbitrarily chosen, the relative WMPRs between the teams in a tournament is the same regardless of what value is chosen for their mean.

The pseudo-inverse solution shown above finds the minimum-norm solution which results in **W** having a mean of 0.

Combined Power Rating (CPR)

For comparing across tournaments, it can be useful to normalize the **W** values so they have a mean equal to *Oave*. This provides for comparisons on the same scale as the offensive measures, but the elements include offensive and defensive components.

---

[4] **M_(w')** is using the "prime" symbol ‘ to show it is different vector from **Mw**. This is different from the transpose of **Mw**, denoted by **M'w.** Sorry for the confusing notation.

The CPR is essentially this:

$$\underline{\mathbf{C}}_{cpr} = \underline{\mathbf{W}}_{wmpr} + O_{ave}$$

This is a **C**-type vector because its mean is *Oave*, but it incorporates contributions of both offense and defense. For example, in a tournament where the average offensive match score divided by 3 was 50, a team with a CPR of 60 would contribute 10 more points than the average team to their alliance's winning margin, through a combination of above average offense and defense. In a tournament with little-to-no defense, this would be similar to a team having an OPR of 60.

Defense

Defensive Power Rating (DPR)

The DPR estimate is computed as:

$$\underline{\mathbf{D}}_{dpr} = (\mathbf{A}_d'\mathbf{A}_d)^{-1}\,\mathbf{A}_d'\mathbf{M}_o$$

Another equivalent equation for the computation is:

$$\underline{\mathbf{D}}_{dpr} = \underline{\mathbf{O}}_{opr} - \underline{\mathbf{W}}_{ccwm} = (\mathbf{A}_o'\mathbf{A}_o)^{-1}\mathbf{A}_o'\,(\mathbf{M}_o - \mathbf{M}_{w'})$$

In both equations, the matrix being inverted is a matrix with i,jth element equal to the number of times teams i and j played on the same alliance. The diagonal elements are the number of matches played by each team, and the off diagonal elements count the number of partnered matches between the teams. Note that this is same same regardless of whether **A'o Ao** or **A'd Ad** is used, since the number of times teams play offense together is the same as the number of times teams play defense together. In other words, **A'o Ao** = **A'd Ad** = **Ar' Ar** + **Ab' Ab**.

In both equations, the vector term to the right of the matrix inverse is the sum of the match scores of each team's opponents in the tournament. Again, they are two different expressions for the same result.

This is essentially using the OPR/ CCWM formulation to try to predict the scores of the opposing alliance using only knowledge about a given team's alliance partners and *no knowledge of the teams actually playing on the opposing alliance!* As a result, DPR often incorporates more knowledge about the average strength of a team's opposing alliances than it does incorporate a team's defensive strength or weakness. In a game with little-to-no defense, the DPR can be used purely as a measure of how lucky or unlucky a given team was by having weaker or stronger opponents than average.

Also, the DPR estimates have an opposite sign from the definition of defensive contribution used in this paper, meaning that a *small* DPR is good, as it signifies a team's opposing alliances had lower average scores.

As a result, this is *not* an estimate of the modeled defensive contribution vector **D**. The limiting behavior of DPR is

$$\underline{\mathbf{D}}_{dpr} \rightarrow O_{ave} - \mathbf{D}$$

and an estimate using the average of opponent's match scores that approaches a team's true underlying defensive contribution is then

$$\underline{\mathbf{D}}_{dprb} \equiv O_{ave} - \underline{\mathbf{D}}_{dpr} \rightarrow \mathbf{D}$$

The DPRb estimates are thus based on the DPR values but converge to the actual underlying values of **D** under limiting assumptions.

Simultaneous Offense and Defense

Simultaneous Offensive/ Defensive Power Rating (sOPR and sDPR)

Following the derivations for OPR and WMPR as in previous sections, the full offensive and defensive contributions of teams, **O** and **D**, can be jointly and simultaneously estimated by minimizing *Eo,* which is equivalent to minimizing **No' No** in the full model for **Mo**. This results in the least squares solution of

$$\begin{bmatrix} \underline{\mathbf{O}}_{sopr} \\ \underline{\mathbf{D}}_{sdpr} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_o'\mathbf{A}_o & -\mathbf{A}_o'\mathbf{A}_d \\ -\mathbf{A}_d'\mathbf{A}_o & \mathbf{A}_d'\mathbf{A}_d \end{bmatrix}^+ \begin{bmatrix} \mathbf{A}_o' \\ \mathbf{A}_d' \end{bmatrix} \mathbf{M}_o$$

As with the WMPR measure, a constant *K* could be added to all values of **O** and **D** and still yield the same match outcome, so the pseudo-inverse must be used.

For comparison with other metrics and across tournaments, it is suggested that the solution provided by the equation above be normalized so that the mean of the **D** values is zero. This allows for the interpretation that a team's sOPR value is their offensive contribution against an average defensive alliance, and that a team's sDPR value is their defensive contribution relative to the defense of the average team.

This formulation is the first time that the matrix products **Ao' Ad** appear. The i,jth elements of this matrix product is the number of times teams i and j played on *opposing* alliances. The diagonal elements of this matrix product are zero, as it is impossible for a team to play simultaneously on both the Red and Blue alliances.

Compare the equation above with an equivalent equation for computing both the standard OPR and DPR shown below, and it becomes apparent that the main difference with the new formulation is the inclusion of the **Ao' Ad** terms. Note that the pseudo-inverse of a full rank square matrix is just the regular inverse.

$$\begin{bmatrix} \underline{\mathbf{O}}_{opr} \\ \underline{\mathbf{D}}_{dpr} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_o'\mathbf{A}_o & 0 \\ 0 & \mathbf{A}_d'\mathbf{A}_d \end{bmatrix}^+ \begin{bmatrix} \mathbf{A}_o' \\ \mathbf{A}_d' \end{bmatrix} \mathbf{M}_o$$

Simultaneous Winning Margin Power Rating (sWMPR) and Combined Power Rating (sCPR)

The Simultaneous Winning Margin Power Rating (sWMPR) can be computed from the sOPR and sDPR as

$$\underline{\mathbf{W}}_{swmpr} = \left(\underline{\mathbf{O}}_{sopr} - Oave\right) + \underline{\mathbf{D}}_{sdpr}$$

The Simultaneous Combined Power Rating can be computed from the sOPR and sDPR as

$$\underline{\mathbf{C}}_{scpr} = \underline{\mathbf{O}}_{sopr} + \underline{\mathbf{D}}_{sdpr}$$

This is a measure of a team's overall contribution, normalized so the average element is equal to *Oave*. As with the CPR, this allows for better comparisons across different tournaments.

Mixtures

Ether's Power Rating (EPR)

Contribution estimates can be computed by minimizing mixtures of the error measures used in previous sections. For example, Ether has proposed minimizing

$$E_E(\alpha) = \alpha E_o + (1 - \alpha) E_w$$

with $\alpha = 0.5$. This tries to find the **C**-like vector that minimizes the sum of the prediction errors for both the offensive match score prediction *and* the winning margin prediction at the same time. Any combination of these two errors could also be used. α=1 corresponds to the normal OPR solution and α=0 corresponds to the normal WMPR solution.

With α = 0.5, this is equivalent to the model

$$\begin{bmatrix}\mathbf{M}_o \\ \mathbf{M}_w\end{bmatrix} = \begin{bmatrix}\mathbf{A}_o \\ \mathbf{A}_w\end{bmatrix}\mathbf{C} + \begin{bmatrix}\mathbf{N}_o \\ \mathbf{N}_w\end{bmatrix}$$

or

$$\mathbf{M}_e = \mathbf{A}_e\mathbf{C} + \mathbf{N}_e$$

The least squares solution is

$$\underline{\mathbf{C}}_{epr} = \left(\mathbf{A}_e^{'}\mathbf{A}_e\right)^{-1}\mathbf{A}_e^{'}\mathbf{M}_e$$

Note that the least squares solution usually is the maximum likelihood solution if the underlying noise is independent and identically distributed Gaussian/normal noise. This is NOT the case in the EPR solution as **No** and **Nw** are not independent but are instead directly related.

Estimating Model Parameters Using Minimum Mean Squared Error (MMSE) Methods

Problems with Least Squares Estimation Methods

The application of least squares estimation methods to FIRST tournaments suffers from the following problems:

1. Insufficient Data/ Overfitting. In full tournaments, usually the number of total matches/ the number of teams playing is somewhere around 2. For example, if there are 54 teams playing in a tournament with 108 matches total, then OPR, DPR, and CCWM measures are each estimated using 216 match results to estimate 54 parameters, or only 4 data points per parameter. For WMPR, CPR, sOPR, and sDPR measures, only 2 data points per parameter are used.

   In these cases, the estimated parameters are very sensitive to the noise in the match results, and the estimated parameters are often poor estimates of the underlying parameters and poor at predicting match results not used in the parameter estimation process (i.e., the “training”).

   If parameters are estimated using partial tournament results to try to predict later qualification matches, the performance is even worse.

   It is not even possible to compute any of the least squares estimates until a sufficient number of matches has been played, as the resulting matrix to be inverted is not full rank.

2. No incorporation of apriori knowledge of expected parameter ranges. The least squares approach assumes nothing about what the parameters should be. For example, least squares solutions weight the likelihood of a given OPR value being -1000, 50, or 100000 the same. If the range of the underlying **O** and **D** values is roughly known, this information can be incorporated in the parameter estimation techniques to improve the estimation.

MMSE Estimation of Parameters

In the previous sections, the unknown parameters were assumed to be constants that could take any value. If the unknown parameters are instead modeled as random variables themselves coming from a particular random distribution, improved estimation can be achieved using MMSE parameter estimation techniques. MMSE techniques essentially minimize the squared error in the prediction of the match scores *plus* the squared error in the prediction of the parameters themselves. If the match noise and the parameter distributions are both modeled as normally distributed, this becomes maximum likelihood estimation of the model parameters. Wikipedia has a good overview of MMSE Estimation.

In particular, if the underlying elements of **O** have mean *Oave*= ⅓ of the average match score and variance σ_o^2, the underlying elements of **D** have mean *Dave*=0 and variance σ_d^2, and the elements of the match score noise are zero mean and have variance σ_n^2, then the various parameter estimation solutions using MMSE techniques become

$$\underline{\widehat{\mathbf{O}}}_{opr} = \left(\mathbf{A}_o'\mathbf{A}_o + \frac{\sigma_n^2}{\sigma_o^2}\mathbf{I}\right)^{-1}\mathbf{A}_o'\left(\mathbf{M}_o - 3Oave\right) + Oave$$

$$\underline{\widehat{\mathbf{W}}}_{ccwm} = \left(\mathbf{A}_o'\mathbf{A}_o + \frac{2\sigma_n^2}{\sigma_o^2 + \sigma_d^2}\mathbf{I}\right)^{-1}\mathbf{A}_o'\mathbf{M}_{w'}$$

$$\underline{\widehat{\mathbf{W}}}_{wmpr} = \left(\mathbf{A}_w^{'}\mathbf{A}_w + \frac{2\sigma_n^2}{\sigma_o^2 + \sigma_d^2}\mathbf{I}\right)^{-1}\mathbf{A}_w^{'}\mathbf{M}_w$$

$$\underline{\widehat{\mathbf{C}}}_{cpr} = \underline{\widehat{\mathbf{W}}}_{wmpr} + O_{ave}$$

$$\begin{bmatrix}\underline{\widehat{\mathbf{O}}}_{sopr}\\ \underline{\widehat{\mathbf{D}}}_{sdpr}\end{bmatrix} = \left(\begin{bmatrix}\mathbf{A}_o'\mathbf{A}_o + \frac{\sigma_n^2}{\sigma_o^2}\mathbf{I} & -\mathbf{A}_o'\mathbf{A}_d\\ -\mathbf{A}_d'\mathbf{A}_o & \mathbf{A}_d'\mathbf{A}_d + \frac{\sigma_n^2}{\sigma_d^2}\mathbf{I}\end{bmatrix}\right)^{-1}\begin{bmatrix}\mathbf{A}_o'\\ -\mathbf{A}_d'\end{bmatrix}\left(\mathbf{M}_o - 3Oave\right) + \begin{bmatrix}Oave\\ 0\end{bmatrix}$$

The MMSE solutions become the Least Squares solutions as the apriori variances of the **O** and **D** grow large, which is another way of saying that the Least Squares solutions are just the MMSE solutions where little-to-no apriori information is known about the values of the **O** and **D** values.

The MMSE solutions also become the Least Squares solutions as the number of matches grows large, as the **A' A** matrices can be roughly approximated by *K* **I**, where *K* is the number of matches a given team plays in the tournament. As the tournament size gets large, the first term in each matrix inverse grows, causing the equations to become closer and closer to the LS equations.

Note that σ_o^2 and σ_d^2 are the apriori estimates of the offensive and defensive variances *per team*, *not per match*. Note also that σ_n^2 is the variance of the noise component *on each individual* *match score, R or B*. As a result, the elements of **Mw** have a variance of $2\sigma_n^2$ leading to the factors of 2 that appear in the equations for the MMSE estimates of CCWM and WMPR.

The MMSE estimate for EPRs is complicated by the fact that the covariance matrix for the expected noise is not diagonal (i.e., the noise elements are not independent). Its derivation is discussed in Appendix B.

Discussion of MMSE Estimation

MMSE techniques have the following advantages:

1. They can be used even when the number of underlying equations is less than the number of unknowns! This means that initial parameters can be estimated in a tournament even when only a few matches have been played. Initial estimates start at *Oave* and 0 for **O** and **D** respectively and progressively incorporate match information as matches are played.

2. They are always full rank and thus remove the need for pseudo inverses in any computations.

3. They address overfitting problems that occur when parameters are estimated with few data points.

MMSE techniques have the following disadvantages:

1. They require apriori expectations of the different variances. Different choices for these variances result in different parameter estimates. The relative values of the variances will vary from season to season as the particular scoring statistics of that season's game change from those in previous years.

A suggested method for selecting the relative variance values is to compute the MMSE solutions for different sets of variance values and select the MMSE solution that best predicts match results not used in the training process. This is discussed in detail in later sections.

Care should be used before deciding that σ_d^2 =0, even in a game where there is "no defense." Teams can adversely impact the scoring of the opposing alliance in different ways, including direct robot-to-robot defensive tactics, robots being effective at grabbing important scoring elements that both teams need (e.g., in FRC Recycle Rush, robots winning the initial "can race"), robots being effective at quickly grabbing quantities of low-scoring elements that both teams need (e.g., in FTC Cascade Effect, robots quickly grabbing most large balls or grabbing most large balls from the opposing side of the field, making it harder for the opposing alliance to find large balls to score), or teams creating hazards or scoring negatives for opposing alliances (e.g., litter in FRC Recycle Rush). Perhaps the only time that σ_d^2 =0 is a valid assumption is in games where there is no possible interaction between the two alliances in any way. If there's any interaction at all, σ_d^2 might not be zero. However, σ_d^2 is often much less than σ_o^2 and the resulting elements of **D** will then often have much less variance than the elements of **O**.

Advanced MMSE Estimation

The MMSE estimation techniques are improvements over least squares techniques because they incorporate apriori knowledge about the expected ranges of the **O** and **D** vectors. However, the derivations above assume that the apriori knowledge is the same for all teams.

Further improvements to the parameter estimations could be made if knowledge about non-equal expectations is also incorporated. This is a way to incorporate scouting into the calculation of the parameters. A few examples might be:

1. If a team shows up with a dead robot or no robot, their individual expected apriori *Oi* could be set to 0 instead of *Oave* and the expected variance of their *Oi* and *Di* values could be set to 0.

2. If a team is consistently superior from tournament to tournament and season to season, their apriori expected *Oi* could be set to be above *Oave*. Similarly, if a team is consistently inferior, their apriori expected *Oi* could be set to be below *Oave*.

3. If a team is known to play strong defense, their apriori expected *Di* could be set above 0.

4. If a team is known to attempt a scoring technique that can score a large amount of points but that is inconsistent, their apriori expected σ_o^2 could be set to a value greater than that of other teams, signifying that they are expected to have a greater variance in scoring than the average team.

5. Before championships, the **O** and **D** estimates of all teams could be computed from regional tournaments and those estimates could be used as the apriori estimates of the **O** and **D** values for the teams at championships. This allows for match prediction to occur at championships even before any matches are played. The equations for this are presented in Appendix D.

These adjustments are similar to normal scouting techniques in that they attempt to incorporate additional knowledge that may exist into ratings of teams.

These possibilities are not discussed further in this paper. Appendix D shows the MMSE equations for OPR where *different* probabilistic estimates are known for the mean and variance of each team's OPR value.

Evaluation of Methods

This section presents the performance results for the metrics when used in both simulated tournaments and actual tournaments.

Simulated Tournaments

Simulation Environment

The baseline tournament structure used for the initial simulations was the 2014 casa FRC tournament. This tournament had 54 teams and 108 matches. Each team played once every 9 matches, and each team played 12 matches overall. There are 2 overall matches per team in the tournament.

This structure was used to simulate tournaments with more played. For larger tournaments, the 108 match schedule was repeated but with teams randomly permuted so that the same alliances did not occur every 108 matches.

The OPR was computed for the actual tournament results. It was found that the mean OPR was around 25 and the standard deviation of the OPR values was around 10. As a baseline for the simulated tournament results, tournaments were simulated with the actual underlying O values taken from a normal distribution with mean 25 and standard deviation of 10. Defensive and random match noise contributions were then added at varying levels to study the effects on the parameter estimation algorithms.

Metrics of performance were computed using the full dataset in three ways:

- **"Training Set"**: The full set of data (i.e., the Training Set) was used to compute the model parameters, the model parameters were used to predict the match scores and winning margins, and then the metrics were computed comparing the actual match data with the predicted match data. These results can be computed for simulated or actual tournaments.

  Measurements are reported as *Eo*/ (**Mo**-*Oave*)'(**Mo**-*Oave*) or *Ew/* (**Mw**' **Mw**) in percent, which shows the percentage of the variance of the **Mo** or **Mw** vector that is not predicted by the estimate. 0% corresponds to perfect prediction and 100% corresponds to no prediction.

- **"Testing Set"**: For actual tournament match data, one match was removed from the training set, the model parameters were computed based on all other matches, the computed model parameters were used to predict the scores and winning margin for the single removed match, and then the metrics were computed comparing the actual match data with the predicted match data for the one removed match. This was repeated, each time removing a different match, and the results were computed totalling or averaging the results from each individual test across all matches.

For simulated tournaments, the model parameters were computed based on a first set of simulated matches, and then the metrics were computed comparing a second set of simulated match outcomes with the predicted match outcomes using the parameters computed from the first set of simulated matches.

Measurements are reported as *Eo*/ (**Mo**-*Oave*)'(**Mo**-*Oave*) or *Ew/* (**Mw**' **Mw**) in percent, which shows the percentage of the variance of the **Mo** or **Mw** vector that is not predicted by the estimate. 0% corresponds to perfect prediction and 100% corresponds to no prediction. In overfit cases, it is possible for the measurements to be more than 100% on testing data, signifying that the "prediction" based on a small training set is worse than performing no prediction at all on testing data.

- **"Parameter Estimation"**: The full set of data was used to compute the model parameters and the estimated model parameters were compared with the actual underlying model parameters. This is only possible with simulated tournaments, as for real tournaments the model is only an approximation and there are no underlying model parameters generating the data.

  Measurements are reported as *Ep*/ (**P**- *Pave*)' (**P**-*Pave*) in percent, which shows the percentage of the variance of the parameter vector that is not predicted by the estimate. 0% corresponds to perfect prediction and 100% corresponds to no prediction.

Training Set results can look artificially good when there are relatively few match data points available for estimating the parameters (i.e., when "Overfitting" occurs).

Testing Set results cannot look artificially good due to overfitting, as the Testing set data is not used in the parameter estimation. Testing Set results are more appropriate for comparing how the models actually do in predicting unknown matches.

With a very large number of matches, the Training Set results should approach the Testing Set results. However, most FRC and FTC matches have between 1 and 3 overall matches per team in a given tournament, which introduces overfitting problems in the model parameter estimation procedures.

Very Large Tournaments

As a starting point, simulations were run for a tournament like the 2014 casa tournament with 54 teams but with *10 times the number of matches* as the actual tournament (i.e., 1,080 total matches instead of 108 matches).

The results of a sample run are shown below:

VAR(O)=10;   VAR(D)=0*VAR(O);   VAR(N)=1*VAR(O);   Least Squares Methods used

```
Parameter Estimation Data
O: Percent of variance of offensive contributions O left
Oa   :   2.9
OPR  :   0.9
sOPR :   0.9

W: Percent of variance of winning margin contributions W left
Wa   :   4.2
OPR  :   0.9
sCPR :   1.3
CPR  :   1.3
CCWM :   2.2
EPR  :   1.0

Training Match Prediction Data
Mo: Percent of variance of offensive scores left
Oa   :  31.9:
Oa+Da:  32.5:
OPR  :  30.9:
O+DPR:  31.6:
sODPR:  30.4:

Mw: Percent of variance of winning margins left
Wa   :  32.7:
OPR  :  30.8:
sCPR :  30.4:
CPR  :  30.4:
CCWM :  31.6:
EPR  :  30.5:
```

Comments:

There is no defensive component to the true underlying scores (VAR(D)=0), so there is no benefit to adding defensive estimated parameters. Both OPR and sOPR have found the correct underlying O parameters to within around 1% of the true values, and all estimates of the underlying W parameters are within a few percent of the true values.

The actual match scores are of the form M=O1+O2+O3+N where the O and N values all have identical variance (VAR(N)=1*VAR(O)), so 75% of the overall match variance is due to offensive contributions and 25% of the overall match variance is due to the noise. Most estimates of the

O parameters successfully predict about 70% of the match variance and all of the estimates of the W parameters successfully predict about 70% of the variance in the match winning margins.

The CPR and sCPR values are identical. This is the case for LS estimates (where underlying parameters are identical) but not for the MMSE estimates.

In contrast, the results for a run with defensive contributions equal in size to offensive contributions (i.e, a LOT of defense, with VAR(D) = VAR(O) are shown below:

VAR(O)=10; VAR(D)=1*VAR(O); VAR(N)=1*VAR(O); Least Squares Methods used

```
Parameter Estimation Data
O: Percent of variance of offensive contributions O left
Oa   :   3.5
OPR  :   2.6
sOPR :   0.9

D: Percent of variance of defensive contributions D left
Da   :   4.6
DPRb :   3.0
sDPR :   1.0

W: Percent of variance of winning margin contributions W left
Wa   :   3.3
OPR  :  39.7
sCPR :   0.7
CPR  :   0.7
CCWM :   2.3
EPR  :  10.2

Match Prediction Data
Mo: Percent of variance of match scores left
Oa   :  57.3:
Oa+Da:  17.8:
OPR  :  56.9:
O+DPR:  17.0:
sODPR:  15.7:

Mw: Percent of variance of winning margins left
Wa   :  15.5:
OPR  :  46.2:
sCPR :  13.7:
CPR  :  13.7:
CCWM :  14.7:
EPR  :  21.4:
```

Comments:

Now, the defensive components are equal in size to the offensive components. Given the very large tournament size, the OPR estimate is still a good estimate for the underlying O values, though it's off by a bit more than the sOPR estimate (2.6% v. 0.9%) because the OPR calculation doesn't factor in the defensive contributions, resulting in a greater noise or deviation in the OPR calculation.

The different estimates of the underlying D values are also all reasonably good.

The winning margins are now equal part O and D, so the OPR can only account for about 50% of the variance in the winning margins while CCWM, CPR, and sCPR can all account for nearly all of the variance. EPR is a pseudo-combination of OPR and CPR/WMPR so its results are somewhere in between.

The actual match scores are now of the form M=(O1+O2+O3)-(D4+D5+D6)+N with all 7 components having equal variance. The OPR calculation can only account for the first 3 components, and so cannot account for 4/7=57% of the variation in match scores and winning margins, whereas CCWM, CPR, and sCPR can account for all but 1/7 = 14% of the variations.

The results for a run "in between" with defense being 10% of the variance of offense are shown below:

VAR(O)=10;   VAR(D)=0.1*VAR(O); VAR(N)=1*VAR(O);   Least Squares Methods used

```
Parameter Estimation Data
O: Percent of variance of offensive contributions O left
Oa   :   2.5
OPR  :   1.1
sOPR :   0.9

D: Percent of variance of defensive contributions D left
Da   :  28.2
DPRb :  27.5
sDPR :   8.9

W: Percent of variance of winning margin contributions W left
Wa   :   4.5
OPR  :   7.5
sCPR :   1.5
CPR  :   1.5
CCWM :   2.4
EPR  :   2.9

Match Prediction Data
Mo: Percent of variance of match scores left
Oa   :  28.7:
Oa+Da:  24.4:
OPR  :  27.9:
O+DPR:  23.5:
sODPR:  22.5:

Mw: Percent of variance of winning margins left
Wa   :  25.4:
OPR  :  28.4:
sCPR :  23.2:
CPR  :  23.2:
CCWM :  24.2:
EPR  :  24.4:
```

The Da and DPRb estimates suffer from the "signal" of the D values being small compared to the "noise" of both match noise and the large variations of the opponent's offensive scores.

A general conclusion drawn from these simulations is that the Least Squares parameter estimation techniques work well when the underlying data matches the model and when the number of data points is large.

Effects of Tournament Size

The effect of having fewer matches is studied in this section.

The following table shows the results of a run where a full simulated 10x tournament was created but the parameters were computed based only on the first 1x, 2x, 3x, 5x, and 10x matches.

VAR(O)=10; VAR(D)=0.0*VAR(O); VAR(N)=1*VAR(O); Least Squares Methods used

```
         1x      2x     3x     5x        10x
Parameter Estimation Data
Percent of variance of offensive contributions O left
OPR :    8.2    2.7    1.8    1.1       0.6
sOPR:    8.8    2.9    1.8    1.2       0.6

Percent of variance of winning margin contributions W left
OPR :    8.2    2.8    1.8    1.1       0.6
sCPR:   10.9    4.7    2.9    1.9       0.9
CPR :   10.9    4.7    2.9    1.9       0.9
CCWM:   24.2   10.8    7.1    4.7       2.2
EPR :    7.6    3.0    2.0    1.3       0.7

Training Match Prediction Data
Percent of variance of offensive scores left
OPR  :  14.5   15.3   15.7   15.9      16.8
sODPR:  10.1   12.9   14.5   15.2      16.5

Percent of variance of winning margin left
OPR :   15.1   14.8   15.0   14.8      15.9
sCPR:   10.9   12.7   13.9   14.1      15.6
CPR :   10.9   12.7   13.9   14.1      15.6
CCWM:   25.3   19.0   18.3   16.7      17.0
EPR :   12.1   13.2   14.2   14.3      15.7
```

Comments:

Even with the data exactly fitting the model, 1 tournament with 108 matches for 54 teams is only enough data to allow the OPRs to be estimated to within about 8-9% of their actual values[5]. The results above are for a single simulation run: the values vary substantially particularly for the small tournament sizes.

The WMPR, CPR, sOPR, sDPR, sCPR, and sWMPR metrics use half the relative amount of data as the OPR, CCWM, and DPR metrics do. Given this, they suffer somewhat more from

[5] The results vary substantially depending on the ratio of Var(N)/Var(O). If Var(N)/Var(O)=2, the percent of variance of offensive contributions O left is often more like 15-20%.

overfitting when the amount of data is small. This can be seen in the Parameter Estimation Data rows comparing OPR with sOPR and OPR with CPR and sCPR. When the tournament size is limited, there is greater variability in these estimates. This can also be seen in the Training Set data, where sOPR, CPR, and sCPR appear to do *better* on small tournaments, because they are fitting the noise in the training set rather than just the underlying model. This false improvement only appears on the Training Set data and does not appear on the Testing Set data.

Training Set vs. Testing Set Data for Least Squares and MMSE Estimates

The output results for a simulated tournament using LS and MMSE estimation techniques for the parameters are shown on the next page and explained on the following page.

```
2014: casa
Teams = 54, Matches = 108, Matches Per Team = 2.000
Simulated Data!! sig2D/sig2O = 0.100000     sig2N/sig2O = 1.000000

MMSE search parameters
VarD/VarO =  0.00 to  0.40, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Parameter Estimation Data
        LS     |  MMSE
O: Percent of variance of offensive contributions O left
Oa   :  19.0
OPR  :   9.5   |   7.9
sOPR :   9.0   |   7.6

D: Percent of variance of defensive contributions D left
Da   : 117.9
DPRb : 153.2   |  63.5
sDPR :  68.8   |  41.8

W: Percent of variance of winning margin contributions W left
Wa   :  30.2
OPR  :  12.1   |  11.6
sCPR :  13.3   |   8.0
CPR  :  13.3   |  10.7
CCWM :  25.0   |  15.0
EPR  :   9.2

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  26.5:  |  40.0
Oa+Da:  25.2:  |  41.6
OPR  :  18.0:  |  30.4:  30.4:   0.0        ( 0.00,  0.00)
O+DPR:  17.8:  |  33.6:  29.6:   2.7        ( 0.10,  2.00)
sODPR:   9.0:  |  31.0:  28.9:   4.9   4.9  ( 0.08,  1.00)

Mw: Percent of variance of winning margins left
Wa   :  27.1:  |  37.4
OPR  :  15.7:  |  25.7:  25.7:   0.0        ( 0.00,  0.00)
sCPR :   8.2:  |  27.4:  25.3:   1.8   1.8  ( 0.08,  1.00)
CPR  :   8.2:  |  27.4:  26.8:  -4.0  -4.0  ( 0.40,  1.00)
CCWM :  16.8:  |  30.1:  25.6:   0.5   0.5  ( 0.40,  2.00)
EPR  :  10.3:  |  24.4:
```

Explanation of Simulation Output:

The first three rows describe the simulation parameters.

The next three rows describe the MMSE search parameters used in the MMSE estimation procedure.

The next section describes the ability of the Least Squares (LS) and MMSE parameter estimation techniques to estimate the true underlying parameters. Results for **O**, **D** and **W** are shown.

The final section describes the ability of the LS and MMSE parameters to predict both the Training set match outcomes and the Testing set match outcomes. Results for **Mo** and **Mw** are shown.

%gn1 signifies the percentage reduction of the prediction residual compared to the LS OPR estimation that can be achieved: 100% signifies the same performance as the LS OPR parameters and 50% signifies that the prediction residual has half the variance of the LS OPR prediction residual.

%gn2 signifies the percentage reduction of the prediction residual compared to the MMSE OPR estimation that can be achieved: 100% signifies the same performance as the MMSE OPR parameters and 50% signifies that the prediction residual has half the variance of the MMSE OPR prediction residual.

The (VarD, VarN) numbers show the values of VarD/VarO and VarN/VarO in the MMSE search that produced the best predicted outcome on the Testing data.

Comments:

The results shown are fairly typical, but the results can vary substantially from run to run as the results depend on the randomly selected **O**, **D**, and **N** vectors.

The various parameters do an OK job of estimating the underlying **O** and **W** parameters. Interestingly, the LS EPR does the best job of predicting the winning margins of all of the LS-estimated parameters. It is speculated that this is because the EPR has the greatest number of data points per parameter value and thus suffers from the least amount of overfitting.

It is difficult to estimate the **D** parameters with high accuracy because the values are small, and the **O** and **N** values act as noise in the Da and DPRb estimates. The sDPR estimate is best, but is still not great.

As expected, the sODPR, sCPR, and CPR LS estimates do the best job of predicting the Training set data due to their excessive overfitting. But they are worse on the Testing set data, again due to the overfitting.

Gains can be achieved both by using MMSE estimation on the standard OPR values, and by using sODPR and sCPR estimates of match score and winning margin, respectively, as they can incorporate the defensive contributions.

Data for a simulated run with the match noise 3 times greater than the previous run is shown on the following page.

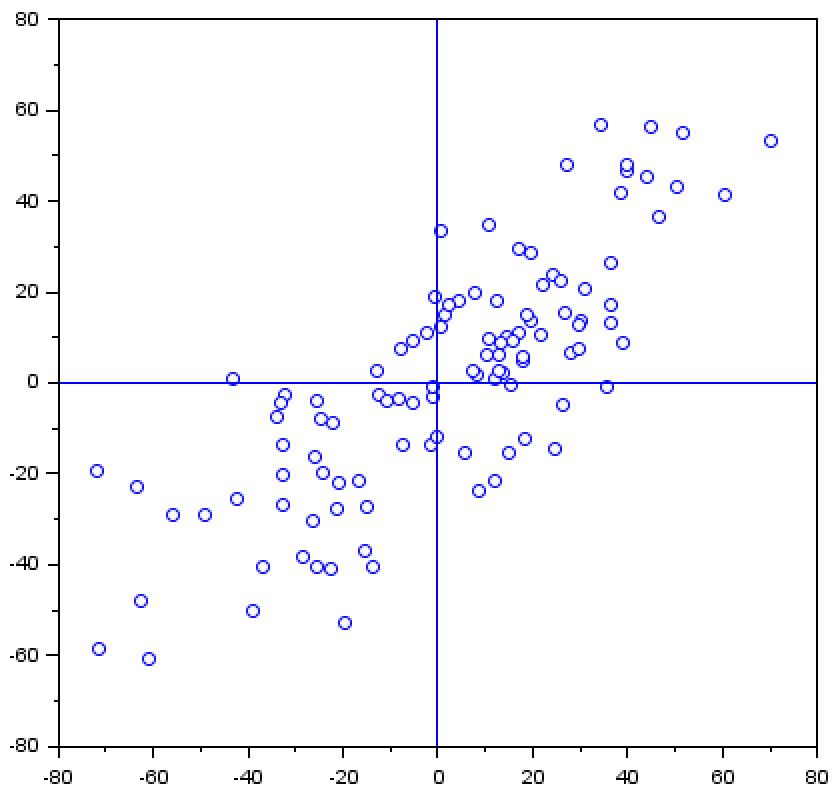
```
2014: casa
Teams = 54, Matches = 108, Matches Per Team = 2.000
Simulated Data!! sig2D/sig2O = 0.100000     sig2N/sig2O = 3.000000

MMSE search parameters
VarD/VarO =  0.00 to  0.40, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Parameter Estimation Data
        LS     |  MMSE
O: Percent of variance of offensive contributions O left
Oa   :  47.4
OPR  :  25.0   |  22.1
sOPR :  39.5   |  22.0

D: Percent of variance of defensive contributions D left
Da   : 283.1
DPRb : 375.3   | 138.2
sDPR : 428.8   |  75.2

W: Percent of variance of winning margin contributions W left
Wa   :  66.2
OPR  :  29.6   |  27.3
sCPR :  59.2   |  25.6
CPR  :  59.2   |  36.0
CCWM :  57.0   |  40.1
EPR  :  29.8

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  49.9:  |  68.6
Oa+Da:  44.8:  |  86.7
OPR  :  43.2:  |  56.9:  56.6:   0.5        ( 0.00,  1.00)
O+DPR:  36.0:  |  73.7:  62.1:  -9.2        ( 0.00,  5.00)
sODPR:  26.2:  |  67.0:  56.1:   1.2   0.7  ( 0.02,  1.00)

Mw: Percent of variance of winning margins left
Wa   :  51.2:  |  87.6
OPR  :  43.8:  |  55.0:  53.9:   2.0        ( 0.00,  2.00)
sCPR :  27.4:  |  63.6:  53.4:   2.9   1.0  ( 0.04,  2.00)
CPR  :  27.4:  |  63.6:  58.0:  -5.6  -7.7  ( 0.04,  2.00)
CCWM :  36.0:  |  69.2:  59.2:  -7.7  -9.8  ( 0.30,  4.00)
EPR  :  32.6:  |  54.7:
```

Now the overfitting is more extreme in the Parameter Estimation. The LS sODPR, sCPR and CPR values that best predict the Training set match data are far from the actual underlying values because they are predicting more noise contribution. This is exactly where MMSE techniques should be most helpful, and indeed they are.

With large noise and small defensive components, all of the estimates of D produce poor results.

Prediction of Match Winners

The figure below shows a plot of the actual winning margins vs. the predicted winning margins by the LS OPR estimation for a simulated 2014 casa tournament.



Better prediction results in match outcomes closer to the y=x line on the scatter plot.

A match winner is predicted correctly if the sign of the true winning margin is the same as the sign of the predicted winning margin. The probability of an error in predicting the match winner is thus the same as the probability of a point in the above plot being in the 2nd or 4th quadrant of the figure (NW or SE quadrants). While reducing the squared winning margin prediction error should reduce the error in match winner prediction on average, the small number of matches in a given tournament can cause a predictor with a smaller winning margin prediction error to have a slightly larger match winner prediction error.

In the simulations run for this paper, the general trend was for the match winner prediction error to follow the match winning margin prediction error, though from simulation to simulation and tournament to tournament there are cases where there were differences.

An example showing a slight difference is the actual (not-simulated) 2015 curie division results shown below.

```
2015: curie
Teams = 76, Matches = 127, Matches Per Team = 1.671

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
OPR  :  33.5:  |  68.5:  62.2:   9.1       ( 0.00,  3.00)
O+DPR:  36.5:  | 101.2:  69.5:  -1.5       ( 0.00,  5.00)
sODPR:  17.6:  | 110.3:  61.8:   9.7   0.6 ( 0.06,  3.00)

Mw: Percent of variance of winning margins left
OPR  :  32.4:  |  64.3:  56.7:  11.9       ( 0.00,  3.00)
sCPR :  15.3:  |  92.6:  54.4:  15.4   4.0 ( 0.10,  2.00)
CPR  :  15.3:  |  92.6:  59.8:   7.1  -5.4 ( 0.10,  2.00)
CCWM :  34.9:  |  81.7:  62.3:   3.2  -9.8 ( 0.10,  3.00)
EPR  :  21.8:  |  62.9:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  19.2:  |  24.0:  24.8
sCPR :  15.2:  |  32.8:  23.2
CPR  :  15.2:  |  32.8:  25.6
CCWM :  12.0:  |  24.8:  24.8
EPR  :  12.8:  |  25.6:
```

In this particular case, even though the variance of the winning margin prediction error for the OPR decreased from 64.3% with LS estimation to 56.7% with MMSE estimation, the probability of incorrectly predicting the match winner increased slightly (24.0% to 24.8%, or one match).

On LS and MMSE Parameters

The MMSE estimation procedure tends to cause the resulting parameters to have a smaller overall variance when compared to the comparable LS parameters. This happens because the MMSE parameters essentially pull the parameter estimates towards their means until enough data is available to create reliable estimates. With only 1-2.5 matches per team in most FRC tournaments, overfitting is often a problem and this MMSE scaling can be significant.

For example, results from a simulated 2014 casa tournament with Variance(O)=100 (or stdev(O)=10), Var(D)/Var(O)=0.0 and Var(N)/Var(O)=3 are shown on the figures on the next page (or maybe two pages from now, depending on how google docs formats the paper). The tournament had 54 teams, so each team played 1 time every 9 total matches. The tournament had 108 total matches, or 12 matches played by every team.

Each of the 4 plots on the left shows the estimated OPRs vs. the number of matches played per team (so X=1 means 9 total matches, X=2 means 18 total matches, etc.). The data points from 1-12 on the X axis correspond to 1 match per team, ... up to 12 matches per team (the whole tournament). The 13th point on the X axis is the *actual* underlying O values.

Plot 1 corresponds to the traditional Least Squares (LS) OPRs, which is also the MMSE solution where Var(N) is estimated to be equal to 0. Note that there are no OPR values until each team has played 4 matches, as that's the number of matches needed to make the matrix invertible. Plot 2 corresponds to the MMSE OPR estimates where Var(N) is estimated to be equal to Var(O) (“MMSE OPR(1)” or “MMSE 1”). As the actual Var(N)=3*Var(O), this is underestimating the noise in each match. Plot 3 corresponds to the MMSE OPR estimates where Var(N) is estimated to be equal to 3* Var(O), the "correct" value (“MMSE 3”). Plot 4 corresponds to the MMSE OPR estimates where Var(N) is estimated to be equal to 10* Var(O), greater than the actual noise (“MMSE 10”).

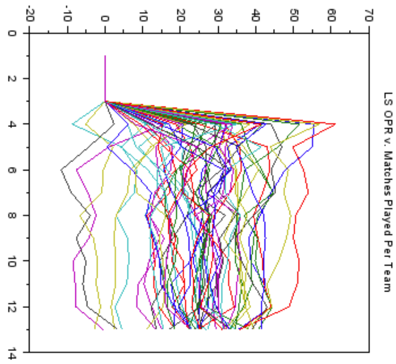
The plot on the right shows the percentage error each curve has in estimating the actual underlying O values.

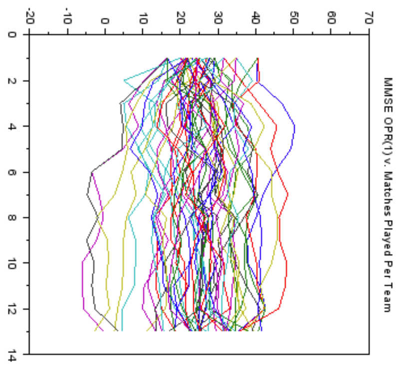
The LS OPR values start out with a high variance and then settle down a bit. Looking at the step from X=12 (the final OPRs) to X=13 (the "real" O values), the final OPRs have more variance than the real O values. This means that the final OPRs are still overestimating the variance of the abilities of the teams.

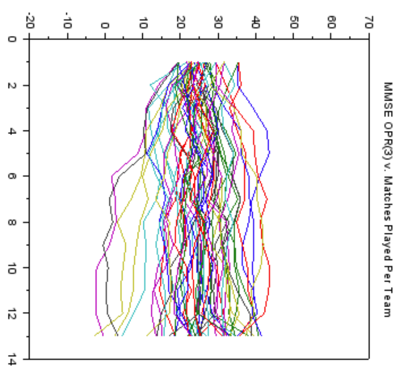
Look at the X=1 points for Plots 2-4. The MMSE estimates start conservatively with the OPRs bunched around the mean and then progressively expand out. Plot 4 shows the noise overestimated (the most conservative estimate), so the OPRs start out very tightly bunched and stay that way. Plot 2 starts out wider, and Plot 3 starts out in the middle.

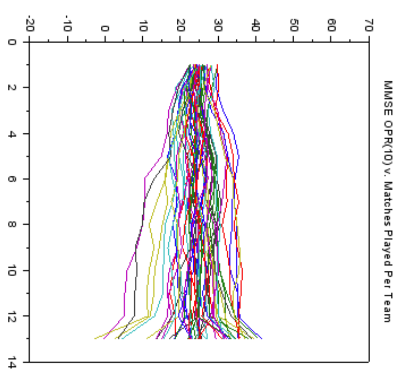
Look at the X=12 (the final estimates) vs X=13 (the "real" O values) points for Plots 2-4. Plot 2 looks like it's still overestimating the variance, Plot 3 has it about right, and Plot 4 has underestimated the true variance even at the end of the tournament (The Plot 4 OPRs expand out from X=12 to X=13). The variances of the OPRs computed by LS, MMSE 1, MMSE 3, and MMSE 10 were respectively 164, 138, 102, and 47. The MMSE 3 solution using the "right" Var(N) estimate is quite close to the true underlying variance of 100. Over multiple runs, the MMSE 3 solution is slightly biased under 100 on average, showing that more matches are needed for it to converge to the "right" variance. All of the techniques do eventually converge to the right solution and variance if the tournament is simulated to be much greater than 108 matches.

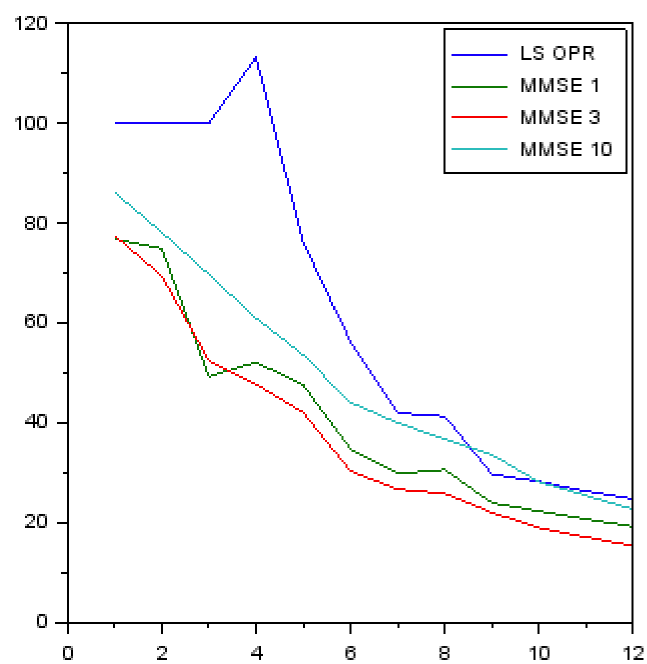
In Plot 5, the performances of the different techniques get close to each other as the tournament nears completion. They should all converge as the number of matches grows large as the LS and MMSE solutions will eventually converge to each other. But they are off by quite a bit early on. Even though the MMSE 1 solution with Var(N) underestimated at 1*Var(O) is underestimating the Var(N), it still gives pretty good results.


LS OPR v. Matches Played Per Team


MMSE OPR(1) v. Matches Played Per Team


MMSE OPR(3) v. Matches Played Per Team


MMSE OPR(10) v. Matches Played Per Team


LS OPR
MMSE 1
MMSE 3
MMSE 10

Actual Tournaments

The different estimation algorithms were run on actual tournament data. For MMSE estimation, different ranges of values for Var(D)/Var(O) and Var(N)/Var(O) were searched for each tournament. The baseline range was usually:

```
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000
```

If the optimal MMSE solution was found at the end of the range for a particular tournament (e.g., with Var(D)/Var(O)=0.10), then the simulation was sometimes rerun with a larger range.

2014 Tournaments

Initial runs were performed on a variety of 2014 tournaments of interest. Some in particular are:

- The 2014 championship divisions
- casa, which was used for the simulated tournaments in the previous section
- casb, which had the largest number of matches per team (2.2).
- gadu, which had the smallest number of matches per team (1.5) and thus the largest potential for overfitting

Results for a number of 2014 tournaments are provided in Appendix C.

As an example, the output summary for the 2014 gadu tournament is:

```
2014: gadu
Teams = 64, Matches = 96, Matches Per Team = 1.500

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  45.9:  |  79.7
Oa+Da:  49.5:  | 104.9
OPR  :  34.8:  |  78.6:  68.4:  12.9       ( 0.00,  3.00)
O+DPR:  31.0:  | 108.9:  78.1:   0.6       ( 0.00,  5.00)
sODPR:  14.4:  | 125.2:  68.0:  13.4   0.6 ( 0.06,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  63.0:  | 108.6
OPR  :  32.4:  |  71.0:  64.4:   9.4       ( 0.00,  3.00)
sCPR :  14.4:  | 118.6:  63.5:  10.5   1.3 ( 0.06,  2.00)
CPR  :  14.4:  | 118.6:  71.7:  -1.0 -11.4 ( 0.10,  2.00)
CCWM :  29.3:  |  96.1:  76.6:  -7.9 -19.0 ( 0.00,  3.00)
EPR  :  21.0:  |  70.5:
```

Comments on this tournament:

MMSE estimation of the OPR values results in a 12.9% reduction in the variance of the match score prediction error on the Testing set compared with LS estimation of the OPR values, and a 9.4% reduction in the variance of the match winning margin prediction errors.

The LS sODPR, sCPR, and CPR metrics do a better job of fitting the Training data, but as previously discussed, this is largely because they are overfitting the Training data. In this tournament in particular (which had the smallest matches per team, 1.5, of all 2014 tournaments), there are only 1.5 data points per parameter being estimated for these parameter sets, 4.5 data points per EPR parameter being estimated[6], and 3.0 data points per parameter being estimated for the rest of the parameter sets.

[6] Though it appears that EPR has more data per parameter being estimated, ⅓ of the data used by EPR (the winning margins) is just a linear combination of the rest of the data (the match outcomes).

As a result, the LS sODPR, sCPR, and CPR metrics do a very poor job of predicting the outcomes in the Testing data, with all 3 having results >100% signifying that they are worse than simply predicting match scores or match winning margins by their respective averages!

However, MMSE estimation corrects this problem. sODPR/sCPR with MMSE estimation results in a very slight improvement in predicting Testing set match scores and winning margins (0.6% and 1.3%, almost nothing) compared with the comparable MMSE-estimated OPR values.

In the best cases, the best predictors are able to predict between 30-40% of the variability in the match scores and winning margins in the Testing set, and 60-70% of the variability remains after prediction. This reinforces that the parameters estimated include some meaningful information about team strength, but also that there is great variability from match to match that cannot be predicted.

General Comments:

In the 2014 Championship divisions, estimating the OPR values using MMSE techniques results in around a 12-13% reduction in the prediction error variance for match scores and winning margins.

In the 2014 Championship divisions, there was little benefit to estimation techniques that incorporated defense. For example, in the Archimedes and Newton divisions there were 3.6% and 0.8% improvements in winning margin respectively, and in the Curie and Galileo divisions there were no improvements over the results achieved by MMSE estimation of the OPRs.

For the 2014 tournaments, the optimal MMSE value for Var(N)/Var(O) averaged around 3 or 4, signifying again that the unpredictable match variation was around the same as the predictable variation which is 3 times Var(O).

For the 2014 tournaments, the optimal MMSE value for Var(D)/Var(O) ranged between 0 and 0.2 and was usually below 0.1, signifying again that there was much more variability between the offense of the teams than the defense, or that the best teams were much better on offense than the worst but only slightly better on the defense than the worst, on average.

The best parameters were better at predicting winning margins than they were at predicting match scores. For example, for the four championship divisions, the percentage of match score variability that could not be predicted was 72.5%, 75.1%, 80.6%, and 78.5% vs. the percentage of winning margin variability that could not be predicted of 58.6%, 64.9%, 70.0% and 67.6%.

In the casb tournament with 2.2 matches per team, around 50% of the winning margin variability could be predicted whereas for other tournaments with 1.5-2 matches per team only around 30-40% of the winning margin variability could be predicted.

2013 and 2015 Tournaments

Results for 2013 and 2015 tournaments are also provided in Appendix C.

General Comments:

The overfitting problem is particularly severe in the 2013 Championship divisions where only 1.33 matches per team were played. In these cases, the LS Training set data is particularly good and the LS Testing set data is particularly bad for sODPR, sCPR, and CPR where overfitting is more problematic. The MMSE results are OK though, with consistent improvements of MMSE based estimation over LS based estimation on Testing set data.

There are not many tournaments where incorporating defensive measures provide benefits in 2013 or 2015 tournaments.

In MMSE estimates of winning margins in 2013 tournaments, the optimal Var(N)/Var(O) ratios are around 1-2, signifying that there was less variance in the overall relative match noise score in the 2013 game compared to the 2014 game.

Areas for Future Work

1. Advanced MMSE/ Scouting Techniques as described previously.
2. Improved MMSE searching. The simulations run here searched the MMSE space using a fixed grid which sometimes may not have had sufficient resolution and sometimes may not have been wide enough to capture the optimal value of Var(D)/Var(O) and Var(N)/Var(O). Improved linear programming techniques could probably be used to search for the optimal values in less time, possibly using Newton or pseudo-Newton methods.
3. MMSE calculations for EPR. While the author speculates that the ad-hoc nature of EPR will not produce superior results to the results shown here, MMSE estimation of the EPR parameters could be performed to prove this speculation correct or incorrect. The MMSE calculation of EPR is discussed in Appendix B.
4. Full comparisons. The parameters could be estimated on the complete set of tournament data to determine the overall benefits of MMSE estimation and sODPR/sCPR based prediction over standard OPR prediction. The match winner prediction could be computed for all tournaments for which data is available to determine the correlation between match winning margin prediction error and match winner prediction error.

Conclusions

New improved techniques for incorporating defense into FRC and FTC tournament statistics have been introduced.

New MMSE techniques for estimating model parameters have been introduced.

Most FRC tournaments do suffer from a small data size, causing Least Squares estimates to be overfit to the noisy tournament data which degrades their performance in predicting match outcomes not in the Training set. MMSE techniques appear to provide limited but significant and consistent improvements in match score and winning margin prediction compared to similar Least Squares techniques.

While incorporating defense into the statistics using MMSE estimation techniques does not result in any decrease in the statistical prediction performance, the advantages in doing so are usually quite small and may make it not worth the effort to do so unless a given FRC season is expected to have substantial defensive components. Occasionally incorporating defense can result in around an 8-12% further reduction in winning margin prediction error (e.g., 2014 casb, 2015 incmp, 2015 micmp tournaments), but this is rare.

MMSE based estimation of the sOPR, sDPR, and sCPR parameters results in the smallest squared prediction error for match scores and match winning margins across all of the studied parameters. MMSE based estimation of OPR parameters often produces results that are quite close.

Least Squares estimates of OPR, CCWM, and DPR using FRC tournament data probably overestimate the relative differences in ability of the teams. MMSE estimates probably underestimate the relative differences.

The small amount of data created in FRC tournaments results in noisy estimates of statistics. Testing set match outcomes from 2013-2015 often had very significant random components to them that could not be predicted by the best linear prediction methods, most likely due to purely random issues that occur in FRC matches.

Thanks and Credits

Thanks to the following folks (in pseudo-alphabetical order)!

AGPapa - for proposing that WMPR be normalized to have an average of *Oave* instead of 0, thus creating CPR, and for many thoughtful and insightful posts.

Ed Law - for conceiving of CCWM, for nice presentations and databases containing OPR and CCWM values, and for many thoughtful and insightful posts.

Ether - for conceiving of EPR, for many, many thoughtful and insightful posts, whitepapers, and personal communications, for providing extensive constructive comments on early drafts of this paper, and for providing the raw tournament data in various forms.

Karthik Kanagasabapthy - for initial work on OPR (which Ed Law says Karthik called Calculated Contribution).

Scott Weingart - for the initial (?) CD posting of OPR in April 2006.

Appendices

Appendix A - Pseudo Code for Parameter Estimation Techniques

```
/* Let
Ar = M_by_t binary matrix of red alliance teams
Ab = M_by_t binary matrix of blue alliance teams

Mr = M_by_1 column vector of red alliance match scores
Mb = M_by_1 column vector of blue alliance match scores

The various metrics can be computed using Least Squares methods as follows: */

nT = 3; // teams per alliance.  use 3 for FRC, 2 for FTC
Oave = mean([Mr; Mb]) /nT;

Ao = [Ar;Ab];      Ad = [Ab;Ar];      Aw = [Ar-Ab];
Mo = [Mr;Mb];                         Mw = [Mw-Mb];

Oa = inv(eye(Ao'*Ao).*(Ao'*Ao)) *Ao' *Mo - (nT-1)* Oave;
Da = nT*Oave - inv(eye(Ad'*Ad).*(Ao'*Ao)) *Ad' *Mo;
Wa = (Oa -Oave) + Da;

OPR = pinv(Ao)*Mo;

Mwprime = [Mw; -Mw];
CCWM = pinv(Ao)*Mwprime;

DPR = OPR - CCWM;
DPRb = Oave - DPR;

A = [Ar-Ab];
M = [Mr-Mb];
WMPR = pinv(Aw)*Mw;
CPR = WMPR + Oave;

Aepr = [Ao;Aw];
Mepr = [Mo;Mw];
EPR = pinv(Aepr)*Mepr;

As = [Ao,-Ad];
sODPR = pinv(As)*Mo + Oave/2;
sOPR = sODPR(1:t);
sDPR = sODPR(t+1:2*t);

sCPR = sOPR + sDPR;
sWMPR= sCPR - Oave;
```

```
/* And the various metrics can be computed using MMSE as follows: */

sig2O=1;

sig2N=2.0;  // pick your value relative to sig2O, or search a range.
            // 2.0 means you expect the match noise to have roughly
            // 2 times the variance as the variance of the OPRs.

sig2D=0.02; // pick your value relative to sig2O, or search a range.
            // 0.02 means you expect defense to be 2% of offense.


OPR = inv(Ao'*Ao + sig2N/sig2O * eye(t)) *Ao' *(Mo-(nT-1)*Oave) + Oave;

CCWM = inv(Ao'*Ao + 2*sig2N/(sig2O+sig2D) * eye(t)) *Ao' *Mwprime;

// if sig2N=0, use LS WMPR estimation instead
WMPR = inv(Aw'*Aw + 2*sig2N/(sig2O+sig2D) * eye(t)) *Aw' *Mw;
CPR = WMPR + Oave;

// if sig2N=0, use LS sOPDR estimation instead
// if sig2D=0, use MMSE OPR estimation instead
sODPR = inv( As'*As +[sig2N/sig2O*eye(t), zeros(t,t);
                      zeros(t,t),         sig2N/sig2D*eye(t)])
                                                   *A' *(M-(nT-1)*Oave);

sOPR = sODPR(1:t) + Oave;
sDPR = sODPR(t+1:2*t);

sCPR = sOPR + sDPR;
sWMPR= sCPR - Oave;

// all matrices are symmetric positive definite, so could use Cholesky
// decomposition to compute inverses.
```

Appendix B - MMSE Estimation of EPR Parameters

The MMSE estimate of the EPR parameters is

$$\underline{\widehat{\mathbf{C}}}_{epr} = \left(\begin{bmatrix} \mathbf{A}_o^{'} & \mathbf{A}_w^{'} \end{bmatrix} \mathbf{C}_N^{-1} \begin{bmatrix} \mathbf{A}_o \\ \mathbf{A}_w \end{bmatrix} + \frac{1}{\sigma_c^2} \mathbf{I} \right) \begin{bmatrix} \mathbf{A}_o \\ \mathbf{A}_w \end{bmatrix} \mathbf{C}_N^{-1} \left(\mathbf{M}_o - 3Oave \right) + Oave$$

where $\mathbf{C_N}$ is the covariance matrix of the noise vector and where σ_c^2 is the expected variance of the EPR values. The limiting value of σ_c^2 is uncertain, as the EPRs would ideally approach the values of **O** to minimize the top part of the EPR equation but they would ideally approach the values of **C** = **O** + **D** to minimize the bottom part of the EPR equation.

In all other cases except EPR, the MMSE solution has a similar form and $\mathbf{C_N}$ is diagonal which allows for the substantial simplifications of the resulting equations as shown in the main body of this paper. However, for EPR,

$$\mathbf{N} = \begin{bmatrix} \mathbf{N}_o \\ \mathbf{N}_w \end{bmatrix} = \begin{bmatrix} \mathbf{N}_r \\ \mathbf{N}_b \\ \mathbf{N}_r - \mathbf{N}_b \end{bmatrix}$$

and as a result, $\mathbf{C_N}$ is *not* diagonal and is instead

$$\mathbf{C}_N = \sigma_N^2 \begin{bmatrix} \mathbf{I} & 0 & \mathbf{I} \\ 0 & \mathbf{I} & -\mathbf{I} \\ \mathbf{I} & -\mathbf{I} & 2\mathbf{I} \end{bmatrix}$$

Because of these complications and because the EPR parameters were not found to be superior to other parameters (i.e., the WMPR, CPR, sOPR, sDPR, and sCPR values), MMSE estimation of EPRs was not performed as part of the simulations that were run for this paper.

Appendix C - Raw Output for Some Tournaments

2013 FRC Regional Tournaments

```
2013: mele
Teams = 38, Matches = 83, Matches Per Team = 2.184

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  37.7:  |  57.2
Oa+Da:  41.2:  |  72.6
OPR  :  35.0:  |  58.8:  56.0:   4.8       ( 0.00,  3.00)
O+DPR:  37.6:  |  81.9:  64.5:  -9.6       ( 0.04,  6.00)
sODPR:  26.0:  |  85.4:  56.0:   4.8   0.0 ( 0.00,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  44.1:  |  79.8
OPR  :  33.4:  |  58.0:  55.6:   4.2       ( 0.00,  2.00)
sCPR :  24.4:  |  80.8:  55.6:   4.2   0.0 ( 0.00,  2.00)
CPR  :  24.4:  |  80.8:  65.6: -13.1 -18.0 ( 0.00,  2.00)
CCWM :  37.1:  |  87.0:  70.9: -22.3 -27.6 ( 0.06,  4.00)
EPR  :  27.0:  |  61.5:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  14.8:  |  17.3:  16.0
sCPR :  12.3:  |  21.0:  16.0
CPR  :  12.3:  |  21.0:  23.5
CCWM :  14.8:  |  29.6:  32.1
EPR  :  14.8:  |  18.5:
```

```
2013: mnmi2
Teams = 60, Matches = 80, Matches Per Team = 1.333

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  34.2:  |  56.6
Oa+Da:  54.4:  |  87.6
OPR  :  21.4:  |  54.7:  48.3:  11.6        ( 0.00,  2.00)
O+DPR:  33.9:  | 111.9:  65.4: -19.6        ( 0.00,  4.00)
sODPR:   9.7:  | 152.4:  48.3:  11.6   0.0  ( 0.00,  2.00)

Mw: Percent of variance of winning margins left
Wa   :  72.1:  |  78.8
OPR  :  17.0:  |  41.5:  38.3:   7.7        ( 0.00,  1.00)
sCPR :   6.5:  | 107.1:  38.3:   7.7   0.0  ( 0.00,  1.00)
CPR  :   6.5:  | 107.1:  51.5: -24.2 -34.6  ( 0.00,  1.00)
CCWM :  33.8:  |  81.3:  54.5: -31.3 -42.3  ( 0.00,  2.00)
EPR  :  11.5:  |  45.3:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  11.5:  |  17.9:  17.9
sCPR :   5.1:  |  25.6:  17.9
CPR  :   5.1:  |  25.6:  21.8
CCWM :  11.5:  |  29.5:  24.4
EPR  :   7.7:  |  19.2:
```

Wow, look at the overfitting of LS sODPR, sCPR, and CPR!

2013 FRC Championship Divisions

```
2013: archi
Teams = 100, Matches = 134, Matches Per Team = 1.340

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa    :  41.5:  |  69.7
Oa+Da:  67.0:  | 105.7
OPR  :  23.0:  |  58.7:  52.8:  10.1        ( 0.00,  2.00)
O+DPR:  42.8:  | 122.9:  68.5: -16.7        ( 0.00,  5.00)
sODPR:  12.1:  | 186.9:  52.8:  10.1   0.0  ( 0.00,  2.00)

Mw: Percent of variance of winning margins left
Wa   :  88.2:  | 114.9
OPR  :  21.6:  |  52.5:  48.2:   8.1        ( 0.00,  1.00)
sCPR :  10.9:  | 168.0:  48.2:   8.1   0.0  ( 0.00,  1.00)
CPR  :  10.9:  | 168.0:  69.2: -31.9 -43.5  ( 0.10,  3.00)
CCWM :  40.6:  | 111.1:  69.6: -32.7 -44.4  ( 0.10,  4.00)
EPR  :  15.8:  |  61.0:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  13.6:  |  24.2:  22.0
sCPR :   9.1:  |  35.6:  22.0
CPR  :   9.1:  |  35.6:  34.8
CCWM :  18.9:  |  33.3:  28.8
EPR  :  11.4:  |  25.8:
```

Wow, look at the overfitting of LS sODPR, sCPR, and CPR!
Only having 1.34 matches per team really creates overfitting in 2013!

```
2013: curie
Teams = 100, Matches = 134, Matches Per Team = 1.340

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING    |  TESTING SET
        LS      |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  39.9:  |  66.4
Oa+Da:  56.9:  |  91.5
OPR  :  22.0:  |  55.9:  50.8:   9.1        ( 0.00,  1.00)
O+DPR:  32.7:  | 100.4:  63.5: -13.6        ( 0.00,  4.00)
sODPR:   8.1:  | 117.9:  50.8:   9.2   0.0  ( 0.02,  1.00)

Mw: Percent of variance of winning margins left
Wa   :  75.2:  |  97.3
OPR  :  21.8:  |  52.7:  47.7:   9.5        ( 0.00,  1.00)
sCPR :   6.9:  |  99.5:  46.9:  11.0   1.7  ( 0.06,  1.00)
CPR  :   6.9:  |  99.5:  55.1:  -4.6 -15.6  ( 0.04,  1.00)
CCWM :  29.8:  |  88.7:  64.4: -22.2 -35.0  ( 0.00,  2.00)
EPR  :  14.1:  |  53.7:

Match Winner Data
    TRAINING    |  TESTING SET
        LS      |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  11.2:  |  21.6:  23.9
sCPR :   7.5:  |  22.4:  21.6
CPR  :   7.5:  |  22.4:  25.4
CCWM :  16.4:  |  32.1:  33.6
EPR  :  10.4:  |  17.9:
```

Wow, look at the overfitting of LS sODPR, sCPR, and CPR!
Only having 1.34 matches per team really creates overfitting in 2013!

```
2013: galileo
Teams = 100, Matches = 134, Matches Per Team = 1.340

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  43.6:  |  78.2
Oa+Da:  61.4:  | 121.6
OPR  :  29.5:  |  75.4:  64.7:  14.2        ( 0.00,  3.00)
O+DPR:  43.0:  | 149.1:  81.3:  -7.7        ( 0.00,  6.00)
sODPR:  14.5:  | 216.0:  64.7:  14.2  -0.0  ( 0.00,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  72.0:  | 127.7
OPR  :  27.5:  |  66.8:  60.4:   9.7        ( 0.00,  2.00)
sCPR :  11.0:  | 160.2:  60.4:   9.7   0.0  ( 0.00,  2.00)
CPR  :  11.0:  | 160.2:  82.1: -22.7 -35.9  ( 0.10,  4.00)
CCWM :  40.4:  | 130.3:  83.3: -24.7 -38.0  ( 0.10,  5.00)
EPR  :  19.4:  |  74.4:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  16.7:  |  29.5:  27.3
sCPR :  10.6:  |  37.1:  27.3
CPR  :  10.6:  |  37.1:  34.1
CCWM :  15.9:  |  31.8:  33.3
EPR  :  13.6:  |  30.3:
```

Wow, look at the overfitting of LS sODPR, sCPR, and CPR!
Only having 1.34 matches per team really creates overfitting in 2013!

```
2013: newton
Teams = 100, Matches = 134, Matches Per Team = 1.340

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  41.7:  |  72.4
Oa+Da:  63.5:  | 107.7
OPR  :  27.8:  |  70.9:  60.7:  14.3        ( 0.00,  3.00)
O+DPR:  43.7:  | 132.8:  73.2:  -3.2        ( 0.00,  6.00)
sODPR:  13.8:  | 216.9:  60.7:  14.3   0.0  ( 0.00,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  73.4:  | 113.1
OPR  :  24.9:  |  61.0:  55.6:   8.9        ( 0.00,  2.00)
sCPR :  10.1:  | 154.9:  55.5:   9.1   0.2  ( 0.02,  2.00)
CPR  :  10.1:  | 154.9:  69.0: -13.1 -24.1  ( 0.10,  2.00)
CCWM :  38.5:  | 105.7:  72.0: -18.0 -29.5  ( 0.08,  3.00)
EPR  :  16.4:  |  64.0:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  20.9:  |  29.9:  28.4
sCPR :  12.7:  |  37.3:  29.9
CPR  :  12.7:  |  37.3:  34.3
CCWM :  16.4:  |  32.8:  32.1
EPR  :  16.4:  |  29.9:
```

Wow, look at the overfitting of LS sODPR, sCPR, and CPR!
Only having 1.34 matches per team really creates overfitting in 2013!

2014 FRC Regional Tournaments

```
2014: casa
Teams = 54, Matches = 108, Matches Per Team = 2.000

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  44.4:  |  70.5
Oa+Da:  40.6:  |  86.0
OPR  :  41.0:  |  72.9:  67.2:   7.8       ( 0.00,  4.00)
O+DPR:  37.0:  |  95.4:  75.2:  -3.3       ( 0.00,  6.00)
sODPR:  27.0:  | 109.0:  67.2:   7.8   0.0 ( 0.00,  4.00)

Mw: Percent of variance of winning margins left
Wa   :  42.2:  |  82.4
OPR  :  37.0:  |  66.4:  62.9:   5.4       ( 0.00,  3.00)
sCPR :  24.0:  |  95.8:  62.8:   5.5   0.2 ( 0.02,  3.00)
CPR  :  24.0:  |  95.8:  70.4:  -6.0 -12.0 ( 0.08,  3.00)
CCWM :  35.8:  |  87.9:  71.7:  -8.0 -14.1 ( 0.10,  4.00)
EPR  :  28.1:  |  69.7:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  16.8:  |  27.1:  26.2
sCPR :   8.4:  |  22.4:  25.2
CPR  :   8.4:  |  22.4:  19.6
CCWM :  10.3:  |  23.4:  24.3
EPR  :  10.3:  |  21.5:
```

```
2014: casb
Teams = 40, Matches = 87, Matches Per Team = 2.175

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  48.3:  |  75.2
Oa+Da:  38.5:  |  86.0
OPR  :  44.9:  |  75.7:  71.2:   5.9       ( 0.00,  3.00)
O+DPR:  37.5:  |  93.6:  80.7:  -6.7       ( 0.10,  5.00)
sODPR:  28.0:  |  94.9:  69.3:   8.4   2.7 ( 0.10,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  28.4:  |  63.0
OPR  :  30.1:  |  52.3:  52.3:   0.0       ( 0.00,  0.00)
sCPR :  17.1:  |  56.8:  46.6:  11.0  11.0 ( 0.10,  1.00)
CPR  :  17.1:  |  56.8:  50.1:   4.4   4.4 ( 0.00,  1.00)
CCWM :  27.2:  |  68.5:  62.2: -18.8 -18.8 ( 0.10,  2.00)
EPR  :  20.9:  |  48.0:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  16.1:  |  26.4:  26.4
sCPR :  12.6:  |  25.3:  23.0
CPR  :  12.6:  |  25.3:  24.1
CCWM :  16.1:  |  32.2:  32.2
EPR  :  14.9:  |  24.1:
```

(11% gain in winning margin prediction for MMSE sCPR vs. MMSE OPR)

```
2014: gadu
Teams = 64, Matches = 96, Matches Per Team = 1.500

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  45.9:  |  79.7
Oa+Da:  49.5:  | 104.9
OPR  :  34.8:  |  78.6:  68.4:  12.9       ( 0.00,  3.00)
O+DPR:  31.0:  | 108.9:  78.1:   0.6       ( 0.00,  5.00)
sODPR:  14.4:  | 125.2:  68.0:  13.4   0.6 ( 0.06,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  63.0:  | 108.6
OPR  :  32.4:  |  71.0:  64.4:   9.4       ( 0.00,  3.00)
sCPR :  14.4:  | 118.6:  63.5:  10.5   1.3 ( 0.06,  2.00)
CPR  :  14.4:  | 118.6:  71.7:  -1.0 -11.4 ( 0.10,  2.00)
CCWM :  29.3:  |  96.1:  76.6:  -7.9 -19.0 ( 0.00,  3.00)
EPR  :  21.0:  |  70.5:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :   9.5:  |  16.8:  20.0
sCPR :  14.7:  |  32.6:  20.0
CPR  :  14.7:  |  32.6:  27.4
CCWM :  15.8:  |  28.4:  26.3
EPR  :   8.4:  |  18.9:
```

2014 FRC Championship Divisions

```
2014: archi
Teams = 100, Matches = 167, Matches Per Team = 1.670

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  50.8:  |  83.5
Oa+Da:  61.8:  | 104.2
OPR  :  41.2:  |  84.2:  73.3:  13.0        ( 0.00,  4.00)
O+DPR:  42.6:  | 116.3:  77.7:   7.8        ( 0.10,  6.00)
sODPR:  23.0:  | 141.3:  72.6:  13.8   1.0  ( 0.10,  4.00)

Mw: Percent of variance of winning margins left
Wa   :  70.9:  |  94.6
OPR  :  33.1:  |  65.2:  61.1:   6.4        ( 0.00,  2.00)
sCPR :  16.4:  | 100.1:  58.9:   9.6   3.5  ( 0.10,  2.00)
CPR  :  16.4:  | 100.1:  62.0:   4.9  -1.5  ( 0.00,  2.00)
CCWM :  40.8:  |  91.2:  62.9:   3.5  -3.0  ( 0.00,  3.00)
EPR  :  22.9:  |  65.0:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  14.5:  |  23.5:  24.1
sCPR :   7.2:  |  25.9:  23.5
CPR  :   7.2:  |  25.9:  25.9
CCWM :  12.0:  |  25.3:  26.5
EPR  :  10.8:  |  25.9:
```

```
2014: curie
Teams = 100, Matches = 167, Matches Per Team = 1.670

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  53.4:  |  87.3
Oa+Da:  79.1:  | 129.1
OPR  :  42.2:  |  85.9:  74.8:  12.8        ( 0.00,  5.00)
O+DPR:  46.4:  | 128.5:  82.8:   3.5        ( 0.00,  6.00)
sODPR:  28.0:  | 172.7:  74.8:  12.8   0.0  ( 0.00,  5.00)

Mw: Percent of variance of winning margins left
Wa   :  96.4:  | 125.4
OPR  :  34.8:  |  68.7:  64.6:   6.1        ( 0.00,  2.00)
sCPR :  22.5:  | 137.1:  64.6:   6.1   0.0  ( 0.00,  2.00)
CPR  :  22.5:  | 137.1:  72.7:  -5.8 -12.6  ( 0.00,  4.00)
CCWM :  43.2:  | 105.6:  71.5:  -4.0 -10.8  ( 0.04,  5.00)
EPR  :  27.9:  |  79.4:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  20.5:  |  30.1:  31.3
sCPR :  18.7:  |  39.8:  31.3
CPR  :  18.7:  |  39.8:  34.3
CCWM :  21.7:  |  36.1:  34.3
EPR  :  19.9:  |  34.3:
```

```
2014: galileo
Teams = 100, Matches = 167, Matches Per Team = 1.670

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  53.9:  |  88.6
Oa+Da:  79.3:  | 127.1
OPR  :  44.7:  |  91.1:  77.6:  14.8        ( 0.00,  6.00)
O+DPR:  49.8:  | 134.7:  83.7:   8.1        ( 0.04,  6.00)
sODPR:  27.4:  | 168.1:  77.6:  14.8   0.0 ( 0.00,  6.00)

Mw: Percent of variance of winning margins left
Wa   :  94.0:  | 124.7
OPR  :  36.9:  |  73.0:  68.6:   6.1        ( 0.00,  3.00)
sCPR :  20.8:  | 125.4:  68.6:   6.1   0.0 ( 0.00,  3.00)
CPR  :  20.8:  | 125.4:  74.0:  -1.3  -7.9 ( 0.00,  4.00)
CCWM :  45.9:  | 112.5:  73.1:  -0.1  -6.7 ( 0.10,  6.00)
EPR  :  28.4:  |  81.0:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  21.6:  |  31.1:  29.3
sCPR :  18.6:  |  37.1:  29.3
CPR  :  18.6:  |  37.1:  30.5
CCWM :  21.0:  |  31.1:  29.3
EPR  :  19.2:  |  34.1:
```

```
2014: newton
Teams = 100, Matches = 167, Matches Per Team = 1.670

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  55.7:  |  91.1
Oa+Da:  74.4:  | 119.7
OPR  :  46.7:  |  95.2:  79.5:  16.5        ( 0.00,  6.00)
O+DPR:  43.5:  | 119.3:  82.0:  13.9        ( 0.10,  6.00)
sODPR:  23.2:  | 139.5:  78.5:  17.5   1.1  ( 0.10,  6.00)

Mw: Percent of variance of winning margins left
Wa   :  83.8:  | 112.6
OPR  :  35.6:  |  70.7:  68.1:   3.7        ( 0.00,  2.00)
sCPR :  18.6:  | 108.9:  67.6:   4.5   0.8  ( 0.10,  3.00)
CPR  :  18.6:  | 108.9:  69.5:   1.8  -2.0  ( 0.02,  3.00)
CCWM :  40.6:  |  96.4:  68.7:   2.9  -0.8  ( 0.04,  4.00)
EPR  :  25.3:  |  72.6:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  21.6:  |  31.7:  32.3
sCPR :  14.4:  |  33.5:  31.7
CPR  :  14.4:  |  33.5:  30.5
CCWM :  20.4:  |  30.5:  31.1
EPR  :  18.6:  |  28.7:
```

2015 FRC Regional Tournaments

```
2015: gadu
Teams = 66, Matches = 88, Matches Per Team = 1.333

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  48.5:  |  91.9
Oa+Da:  50.0:  | 118.6
OPR  :  35.6:  |  93.8:  77.8:  17.0        ( 0.00,  4.00)
O+DPR:  39.7:  | 149.8:  87.5:   6.7        ( 0.00,  6.00)
sODPR:  16.5:  | 263.6:  77.8:  17.0   0.0  ( 0.00,  4.00)

Mw: Percent of variance of winning margins left
Wa   :  62.9:  | 116.4
OPR  :  36.4:  |  88.9:  69.9:  21.4        ( 0.00,  4.00)
sCPR :   8.3:  | 117.7:  68.0:  23.5   2.7  ( 0.10,  3.00)
CPR  :   8.3:  | 117.7:  70.0:  21.3  -0.2  ( 0.00,  1.00)
CCWM :  30.9:  |  93.7:  74.9:  15.7  -7.2  ( 0.10,  2.00)
EPR  :  20.3:  |  80.6:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  24.4:  |  30.2:  32.6
sCPR :   8.1:  |  22.1:  29.1
CPR  :   8.1:  |  22.1:  25.6
CCWM :  16.3:  |  33.7:  33.7
EPR  :  15.1:  |  29.1:
```

```
2015: incmp
Teams = 31, Matches = 68, Matches Per Team = 2.194

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  50.6:  |  77.0
Oa+Da:  43.7:  |  89.8
OPR  :  48.5:  |  80.2:  74.7:   6.8        ( 0.00,  4.00)
O+DPR:  41.3:  |  97.9:  83.6:  -4.3        ( 0.10,  6.00)
sODPR:  34.5:  | 114.2:  74.5:   7.1   0.3  ( 0.04,  4.00)

Mw: Percent of variance of winning margins left
Wa   :  32.1:  |  66.7
OPR  :  32.9:  |  57.2:  56.1:   1.9        ( 0.00,  2.00)
sCPR :  19.7:  |  64.8:  51.4:  10.2   8.4  ( 0.08,  1.00)
CPR  :  19.7:  |  64.8:  54.8:   4.2   2.4  ( 0.10,  2.00)
CCWM :  27.5:  |  69.6:  63.7: -11.3 -13.5  ( 0.02,  2.00)
EPR  :  23.3:  |  53.0:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  17.9:  |  26.9:  26.9
sCPR :  13.4:  |  29.9:  26.9
CPR  :  13.4:  |  29.9:  29.9
CCWM :  19.4:  |  31.3:  29.9
EPR  :  13.4:  |  31.3:
```

```
2015: micmp
Teams = 102, Matches = 204, Matches Per Team = 2.000

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  57.2:  |  89.6
Oa+Da:  52.2:  | 101.4
OPR  :  51.3:  |  91.7:  81.4:  11.2       ( 0.00,  6.00)
O+DPR:  44.9:  | 112.5:  84.3:   8.1       ( 0.10,  6.00)
sODPR:  32.9:  | 130.8:  80.8:  11.9   0.8 ( 0.10,  6.00)

Mw: Percent of variance of winning margins left
Wa   :  56.6:  |  86.5
OPR  :  45.4:  |  79.0:  70.4:  10.9       ( 0.00,  5.00)
sCPR :  21.5:  |  83.9:  64.7:  18.2   8.1 ( 0.10,  1.00)
CPR  :  21.5:  |  83.9:  63.0:  20.2  10.4 ( 0.00,  2.00)
CCWM :  34.5:  |  79.8:  66.2:  16.2   5.9 ( 0.00,  3.00)
EPR  :  29.7:  |  69.7:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  25.2:  |  35.1:  33.7
sCPR :  17.3:  |  33.2:  31.7
CPR  :  17.3:  |  33.2:  32.7
CCWM :  19.3:  |  32.2:  32.2
EPR  :  22.3:  |  34.2:
```

2015 FRC Championship Divisions

```
2015: archimedes
Teams = 76, Matches = 127, Matches Per Team = 1.671

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  57.5:  |  95.5
Oa+Da:  61.8:  | 130.2
OPR  :  48.0:  |  98.7:  83.3:  15.6       ( 0.00,  6.00)
O+DPR:  49.9:  | 153.2:  98.3:   0.4       ( 0.00,  6.00)
sODPR:  33.7:  | 207.9:  83.3:  15.6   0.0 ( 0.00,  6.00)

Mw: Percent of variance of winning margins left
Wa   :  73.2:  | 134.2
OPR  :  43.6:  |  86.4:  74.0:  14.4       ( 0.00,  5.00)
sCPR :  24.7:  | 144.8:  74.0:  14.4   0.0 ( 0.00,  5.00)
CPR  :  24.7:  | 144.8:  88.3:  -2.2 -19.4 ( 0.10,  6.00)
CCWM :  45.1:  | 127.3:  89.2:  -3.2 -20.5 ( 0.00,  6.00)
EPR  :  31.2:  |  90.3:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  22.4:  |  33.6:  34.4
sCPR :  17.6:  |  38.4:  34.4
CPR  :  17.6:  |  38.4:  34.4
CCWM :  20.8:  |  40.0:  36.0
EPR  :  19.2:  |  28.8:
```

```
2015: curie
Teams = 76, Matches = 127, Matches Per Team = 1.671

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  46.0:  |  72.8
Oa+Da:  55.4:  |  92.7
OPR  :  33.5:  |  68.5:  62.2:   9.1        ( 0.00,  3.00)
O+DPR:  36.5:  | 101.2:  69.5:  -1.5        ( 0.00,  5.00)
sODPR:  17.6:  | 110.3:  61.8:   9.7   0.6  ( 0.06,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  75.3:  |  97.7
OPR  :  32.4:  |  64.3:  56.7:  11.9        ( 0.00,  3.00)
sCPR :  15.3:  |  92.6:  54.4:  15.4   4.0  ( 0.10,  2.00)
CPR  :  15.3:  |  92.6:  59.8:   7.1  -5.4  ( 0.10,  2.00)
CCWM :  34.9:  |  81.7:  62.3:   3.2  -9.8  ( 0.10,  3.00)
EPR  :  21.8:  |  62.9:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  19.2:  |  24.0:  24.8
sCPR :  15.2:  |  32.8:  23.2
CPR  :  15.2:  |  32.8:  25.6
CCWM :  12.0:  |  24.8:  24.8
EPR  :  12.8:  |  25.6:
```

```
2015: galileo
Teams = 76, Matches = 127, Matches Per Team = 1.671

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  43.3:  |  69.9
Oa+Da:  39.8:  |  87.2
OPR  :  31.2:  |  63.4:  58.8:   7.2        ( 0.00,  2.00)
O+DPR:  31.8:  |  96.2:  70.5: -11.2        ( 0.00,  5.00)
sODPR:  15.8:  |  97.4:  58.1:   8.3   1.2  ( 0.06,  2.00)

Mw: Percent of variance of winning margins left
Wa   :  55.7:  | 110.7
OPR  :  37.1:  |  72.8:  60.7:  16.7        ( 0.00,  4.00)
sCPR :  14.3:  |  86.7:  60.4:  17.1   0.5  ( 0.10,  2.00)
CPR  :  14.3:  |  86.7:  70.6:   3.1 -16.3  ( 0.10,  1.00)
CCWM :  29.5:  |  91.6:  81.7: -12.2 -34.7  ( 0.02,  3.00)
EPR  :  22.9:  |  65.3:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  19.7:  |  26.0:  28.3
sCPR :  11.8:  |  26.8:  26.8
CPR  :  11.8:  |  26.8:  29.1
CCWM :  20.5:  |  35.4:  37.0
EPR  :  15.0:  |  26.8:
```

```
2015: newton
Teams = 76, Matches = 127, Matches Per Team = 1.671

MMSE search parameters
VarD/VarO =  0.00 to  0.10, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 1.000

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  43.7:  |  69.8
Oa+Da:  49.7:  |  94.6
OPR  :  33.5:  |  69.1:  62.0:  10.3        ( 0.00,  3.00)
O+DPR:  36.2:  | 108.1:  75.3:  -9.0        ( 0.00,  6.00)
sODPR:  21.9:  | 135.7:  62.0:  10.3   0.0  ( 0.00,  3.00)

Mw: Percent of variance of winning margins left
Wa   :  59.4:  |  87.0
OPR  :  27.1:  |  53.7:  49.7:   7.4        ( 0.00,  2.00)
sCPR :  13.4:  |  81.6:  49.0:   8.8   1.5  ( 0.06,  2.00)
CPR  :  13.4:  |  81.6:  58.4:  -8.8 -17.5  ( 0.10,  2.00)
CCWM :  27.6:  |  75.6:  64.0: -19.1 -28.7  ( 0.00,  2.00)
EPR  :  18.7:  |  53.8:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  18.1:  |  28.3:  29.9
sCPR :  11.0:  |  31.5:  31.5
CPR  :  11.0:  |  31.5:  30.7
CCWM :  19.7:  |  30.7:  32.3
EPR  :  16.5:  |  29.1:
```

2015 FTC Tournaments

```
Ftc: /2015EastTesla
Teams = 36, Matches = 81, Matches Per Team = 2.250

MMSE search parameters
VarD/VarO =  0.00 to  0.30, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 0.500

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa    :  54.7:  |  84.3
Oa+Da:  43.0:  |  91.8
OPR  :  52.3:  |  86.3:  79.2:   8.2        ( 0.00,  4.00)
O+DPR:  41.3:  |  98.0:  82.0:   5.0        ( 0.26,  5.50)
sODPR:  36.1:  | 113.3:  77.9:   9.7   1.6 ( 0.14,  3.50)

Mw: Percent of variance of winning margins left
Wa   :  42.0:  |  90.0
OPR  :  49.4:  |  80.7:  75.9:   6.0        ( 0.00,  3.00)
sCPR :  32.9:  | 102.2:  73.3:   9.2   3.4 ( 0.20,  3.00)
CPR  :  32.9:  | 102.2:  76.4:   5.4  -0.7 ( 0.22,  3.00)
CCWM :  37.0:  |  89.7:  79.8:   1.2  -5.1 ( 0.00,  2.50)
EPR  :  38.0:  |  80.6:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  19.8:  |  24.7:  27.2
sCPR :  13.6:  |  30.9:  25.9
CPR  :  13.6:  |  30.9:  32.1
CCWM :  18.5:  |  35.8:  32.1
EPR  :  16.0:  |  28.4:
```

```
Ftc: /2015OregonTech
Teams = 23, Matches = 29, Matches Per Team = 1.261

MMSE search parameters
VarD/VarO =  0.00 to  0.30, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 0.500

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  48.0:  | 105.5
Oa+Da:  30.7:  |  77.9
OPR  :  39.7:  | 107.9:  89.2:  17.4        ( 0.00,  2.50)
O+DPR:  30.9:  |  87.0:  60.9:  43.6        ( 0.30,  1.50)
sODPR:   6.7:  | 137.5:  62.8:  41.8  29.6  ( 0.30,  0.50)

Mw: Percent of variance of winning margins left
Wa   :  57.0:  | 108.0
OPR  :  65.8:  | 167.8:  95.3:  43.2        ( 0.00,  6.00)
sCPR :  12.0:  | 208.0:  76.1:  54.7  20.2  ( 0.30,  1.50)
CPR  :  12.0:  | 208.0:  68.3:  59.3  28.3  ( 0.22,  1.50)
CCWM :  29.6:  | 103.3:  74.8:  55.4  21.5  ( 0.16,  2.00)
EPR  :  30.5:  | 135.9:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  17.2:  |  37.9:  31.0
sCPR :   3.4:  |  31.0:  27.6
CPR  :   3.4:  |  31.0:  20.7
CCWM :   6.9:  |  20.7:  24.1
EPR  :  17.2:  |  31.0:
```

```
Ftc: /2015OregonTimber
Teams = 23, Matches = 29, Matches Per Team = 1.261

MMSE search parameters
VarD/VarO =  0.00 to  0.30, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 0.500

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  35.4:  |  80.3
Oa+Da:  25.5:  | 102.1
OPR  :  24.4:  |  68.3:  62.8:   8.0       ( 0.00,  1.00)
O+DPR:  25.8:  | 109.8:  76.7: -12.4       ( 0.00,  2.50)
sODPR:   8.7:  | 167.5:  57.2:  16.3   9.0 ( 0.06,  0.50)

Mw: Percent of variance of winning margins left
Wa   :  31.2:  | 157.0
OPR  :  36.5:  |  96.6:  73.6:  23.8       ( 0.00,  2.00)
sCPR :   8.5:  | 147.7:  52.4:  45.8  28.8 ( 0.18,  0.50)
CPR  :   8.5:  | 147.7:  72.5:  25.0   1.5 ( 0.30,  0.50)
CCWM :  18.7:  |  94.5:  93.9:   2.8 -27.5 ( 0.30,  0.50)
EPR  :  15.6:  |  69.5:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  27.6:  |  34.5:  31.0
sCPR :  20.7:  |  41.4:  31.0
CPR  :  20.7:  |  41.4:  44.8
CCWM :  20.7:  |  51.7:  58.6
EPR  :  24.1:  |  34.5:
```

2015 FTC Championship Divisions

```
Ftc: /2015WorldsEdison
Teams = 64, Matches = 144, Matches Per Team = 2.250

MMSE search parameters
VarD/VarO =  0.00 to  0.30, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 0.500

Match Prediction Data
    TRAINING    |  TESTING SET
        LS      |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  63.6:  |  97.5
Oa+Da:  58.4:  | 109.2
OPR  :  58.5:  |  96.6:  86.9:  10.1        ( 0.00,  6.00)
O+DPR:  47.6:  | 107.6:  84.4:  12.6        ( 0.30,  6.00)
sODPR:  37.3:  | 119.6:  83.9:  13.2   3.4  ( 0.30,  5.00)

Mw: Percent of variance of winning margins left
Wa   :  64.6:  | 113.0
OPR  :  54.0:  |  88.7:  83.3:   6.1        ( 0.00,  4.00)
sCPR :  37.2:  | 118.0:  81.6:   8.0   2.0  ( 0.26,  5.00)
CPR  :  37.2:  | 118.0:  82.9:   6.6   0.5  ( 0.20,  5.00)
CCWM :  47.6:  | 106.3:  82.7:   6.8   0.7  ( 0.22,  5.50)
EPR  :  42.5:  |  90.2:

Match Winner Data
    TRAINING    |  TESTING SET
        LS      |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  25.0:  |  36.8:  34.7
sCPR :  23.6:  |  38.2:  36.8
CPR  :  23.6:  |  38.2:  33.3
CCWM :  20.8:  |  36.8:  31.3
EPR  :  22.2:  |  36.1:
```

```
Ftc: /2015WorldsFranklin
Teams = 64, Matches = 144, Matches Per Team = 2.250

MMSE search parameters
VarD/VarO =  0.00 to  0.30, in steps of 0.020
VanN/VarO =  0.00 to  6.00, in steps of 0.500

Match Prediction Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE        %gn1  %gn2  ( VarD,  VarN)
Mo: Percent of variance of match scores left
Oa   :  50.8:  |  77.0
Oa+Da:  54.4:  |  99.9
OPR  :  44.6:  |  73.8:  69.1:   6.3        ( 0.00,  2.50)
O+DPR:  44.5:  |  99.9:  77.8:  -5.4        ( 0.00,  5.00)
sODPR:  31.6:  | 101.1:  69.1:   6.3  -0.0  ( 0.00,  2.50)

Mw: Percent of variance of winning margins left
Wa   :  58.8:  |  99.6
OPR  :  41.5:  |  68.1:  64.7:   5.0        ( 0.00,  2.00)
sCPR :  30.4:  |  97.3:  64.7:   5.0   0.0  ( 0.00,  2.00)
CPR  :  30.4:  |  97.3:  74.5:  -9.4 -15.2  ( 0.18,  3.00)
CCWM :  44.2:  |  95.7:  75.7: -11.2 -17.0  ( 0.06,  3.50)
EPR  :  34.1:  |  72.4:

Match Winner Data
    TRAINING   |  TESTING SET
        LS     |  LS     MMSE
Probability of incorrectly predicting match winner
OPR  :  20.8:  |  29.9:  27.8
sCPR :  18.8:  |  37.5:  27.8
CPR  :  18.8:  |  37.5:  34.0
CCWM :  22.2:  |  33.3:  31.3
EPR  :  17.4:  |  30.6:
```

Appendix D - MMSE Equations for OPR with A Priori estimates of OPR Values

Instead of assuming a priori that all of the OPR values are identically distributed with mean *Oave* and with variance σ_o^2, MMSE techniques can also be used to estimate the OPR values given *different* a priori estimates for each team's OPR value.

If we are given an a priori estimate of the mean OPR values as the vector $\boldsymbol{O_e}$ and estimates of the variances of the *i*th different team's OPR as σ_{oi}^2, then the MMSE estimate for the OPRs becomes

$$\underline{\hat{\mathbf{O}}}_{opr} = \left(\mathbf{A}_o'\mathbf{A}_o + diag\left(\frac{\sigma_n^2}{\sigma_{o1}^2}, \frac{\sigma_n^2}{\sigma_{o2}^2}, \ldots\right)\right)^{-1} \left(\mathbf{A}_o'\mathbf{M}_o + diag\left(\frac{\sigma_n^2}{\sigma_{o1}^2}, \frac{\sigma_n^2}{\sigma_{o2}^2}, \ldots\right)\mathbf{O}_e\right)$$

and if $\sigma_{oi}^2 = \sigma_o^2$ for all *i* (i.e., the predicted variance for all OPRs is estimated to be the same), then the estimate becomes

$$\underline{\hat{\mathbf{O}}}_{opr} = \left(\mathbf{A}_o'\mathbf{A}_o + \frac{\sigma_n^2}{\sigma_o^2}\mathbf{I}\right)^{-1} \left(\mathbf{A}_o'\mathbf{M}_o + \frac{\sigma_n^2}{\sigma_o^2}\mathbf{O}_e\right).$$

Before a tournament starts, $\boldsymbol{A}o$ = 0 and the estimate is simply the a priori OPR estimate,

$$\underline{\hat{\mathbf{O}}}_{opr} = \mathbf{O}_e$$

As the number of matches in a tournament grows large (which causes the first term in each set of parentheses to dominate) or if a priori knowledge of **O** is not available (i.e., if σ_o^2 is infinite, which makes the second term in each set of parentheses zero), then the equation converges to the standard OPR equation,

$$\underline{\hat{\mathbf{O}}}_{opr} = (\mathbf{A}_o'\mathbf{A}_o)^{-1}\,\mathbf{A}_o'\mathbf{M}_o.$$

Based on the experimental results shown in this paper, a reasonable choice for σ_n^2/σ_o^2 is 3 for FRC tournaments and 2 for FTC tournaments. These numbers are equivalent to assuming that roughly ½ of the variance in each match is due to differences in team strength (OPRs) and ½ is due to non-predictable randomness.