AN EFFECTIVE METHOD FOR IDENTIFYING CLUSTERS OF ROBOT STRENGTHS

BY JEN-CHIEH TENG¹,a, CHIN-TSANG CHIANG²b and ALVIN LIM³4,c

¹Data Science Degree Program, National Taiwan University, aD09948011@ntu.edu.tw
²Institute of Applied Mathematical Sciences, National Taiwan University, bchiangct@ntu.edu.tw
³Data Science Department, Measured, Inc.
⁴Goizueta Business School, Emory University, calvin.lim@emory.edu

In the analysis of qualification stage data from FIRST Robotics Competition (FRC) championships, the ratio (1.67 – 1.68) of the number of observations (110 – 114 matches) to the number of parameters (66 – 68 robots) in each division has been found to be quite small for the most commonly used winning margin power rating (WMPR) model. This usually leads to imprecise estimates and inaccurate predictions in such three-on-three matches that FRC tournaments are composed of. With the recognition of a clustering feature in estimated robot strengths, a more flexible model with latent clusters of robots was proposed to alleviate overparameterization of the WMPR model. Since its structure can be regarded as a dimension reduction of the parameter space in the WMPR model, the identification of clusters of robot strengths is naturally transformed into a model selection problem. Instead of comparing a huge number of competing models (7.76 × 1067 to 3.66 × 1070), we develop an effective method to estimate the number of clusters, clusters of robots and robot strengths in the format of qualification stage data from the FRC championships. The new method consists of two parts: (i) a combination of hierarchical and non-hierarchical classifications to determine candidate models; and (ii) variant goodness-of-fit criteria to select optimal models. In contrast to existing hierarchical classification, each step of our proposed non-hierarchical classification is based on estimated robot strengths from a candidate model in the preceding non-hierarchical classification step. A great advantage of the proposed methodology is its ability to consider the possibility of reassigning robots to other clusters. To reduce overestimation of the number of clusters by the mean squared prediction error criteria, corresponding Bayesian information criteria are further established as alternatives for model selection. With a coherent assembly of these essential elements, a systematic procedure is presented to perform the estimation of parameters. In addition, we propose two indices to measure the nested relation between clusters from any two models and monotonic association between robot strengths from any two models. Data from the 2018 and 2019 FRC championships and a simulation study are also used to illustrate the applicability and superiority of our proposed methodology.

1. Introduction. The FIRST Robotics Competition (FRC) is a sports-like robotic competition sponsored by the international youth organization For Inspiration and Recognition of Science and Technology (FIRST). According to the 2022 FRC Game Manual (FIRST, 2022), “FRC combines the excitement of sport with the rigors of science and technology. Teams of students are challenged to design, build and program industrial-size robots and compete for awards, while they also create a team identity, raise funds, hone teamwork skills and advance
respect and appreciation for STEM within the local community.” Wikipedia (2018) provides a more detailed description of the FIRST organization and its history.

Each year’s FRC season begins with a kickoff event in January when a new and challenging game is introduced with specific rules and mechanics. Teams of high school students around the world then design and build a robot to participate in local, regional and international tournaments. While a new game is introduced each year, the format of a match remains constant over the years. A match involves an alliance of three robots from three teams going against another alliance of three other robots from three other teams in a game defined by the rules and mechanics of the year. Winning a match requires an alliance to score more points over the opposing alliance according to game rules and mechanics. The opposing alliances in a match are randomly designated as a “blue alliance” and a “red alliance” and robots representing each alliance will have bumpers of their alliance’s color (see Figure 1 for visual representation of the game field).

![The 2019 FRC game field. Source: FIRST (2019).](image)

All local, regional and world championship tournaments are divided into a qualification stage and a play-off stage. In the qualification stage, the participating teams’ robots are assigned by an algorithm into a predetermined number of matches and in each match, a robot is randomly assigned to one of two opposing alliances. After each match, the three robots in an alliance are assigned the same number of ranking points (different from scores earned in the match) according to game rules and the outcome of the match. At the end of the qualification stage, a predetermined number of robots with the most ranking points then qualify for the playoff stage. The robots’ average scores across qualification matches are used as an auxiliary measure to break ties when necessary.

In contrast to the qualification stage, where robots are randomly assigned alliances for each match, a predetermined number of top teams whose robots qualified for the playoff stage are asked to form their own alliances of four robots each. These alliances remain throughout matches in the playoff stage, where alliances attempt to eliminate each other across matches. As in the qualification stage, each match in the playoff stage will involve three robots each from opposing alliances with one robot each in reserve. Since each robot is identified with the team that built and owns the robot, we will be using the terms robot and team (or robotics team) interchangeably henceforth. More detailed descriptions of FRC games, matches and tournaments can be found in Lim, Chiang and Teng (2021).

On the basis of very limited official statistics (regular scores, bonuses and penalties), one of the most challenging tasks for the teams when forming alliances for the playoff stage
is to accurately assess the true “strengths” of robots based on their contributions in winning or losing matches. There is still no reliable robot rating/ranking system, albeit the FRC and robotics teams might have their own measures or models to quantify the strengths of competing robots. For example, as mentioned earlier, official rankings of robotics teams are determined by a system of ranking points with the average scores as an auxiliary measure to break ties when necessary. However, a strong (weak) robot may be underrated (overrated) by both the measures since the same number of ranking points and the same score is assigned to all robots in an alliance after a qualification match regardless of each individual robot’s actual contribution to the performance of the alliance.

Due to the importance of objectively assessing a robot’s potential contribution to an alliance in the playoff stage of a tournament and the absence of such accurate and reliable assessment measures beyond those proposed in Lim, Chiang and Teng (2021), we devote this study to building a more accurate model of robot strengths and providing a more effective estimation of model parameters. For the remainder of this section, we review existing rating models that are currently in use by robotics teams and traditional sports teams.

In robotics forums, the calculated contribution to winning margin (CCWM), offensive power rating (OPR) and winning margin power rating (WMPR) models have been popularly used to assess robot strengths. As shown by Law (2008), the CCWM model is a special case of the WMPR model and is meaningless in application. Based on the calculated contribution measure of Karthik Kanagasabapthy in 2004, Fang (2017) modeled the average alliance score as the sum of robot strengths and proposed a least squares estimation procedure. Before this work, Weingart (2006) termed such a measure as the OPR and explained its detailed computation. Due to ignoring actions of robots (as quantified by final match scores) in the opposing alliance, the OPR model tends to have biased estimates. By considering an interaction between alliances and latent factors in each match, Gardner (2015) proposed the WMPR model to characterize the effects of robot strengths on the difference in final match scores. Similar to this model, the adjusted plus/minus rating model and its generalizations have been found to be useful for paired comparison data in traditional sports and esports (cf. Stefani, 1977, 1980; Clarke and Norman, 1995; Rosenbaum, 2004; Macdonald, 2011; Schuckers et al., 2011; Sæbø and Hvattum, 2015; Hass and Craig, 2018; Hvattum, 2019; Clark, Macdonald and Kloo, 2020, among others). In practice, there is no need to introduce a home-court advantage and clutch time/garbage time play in the WMPR model for the FRC.

In an application to qualification stage data from the 2018 FRC Houston and Detroit championships, Lim, Chiang and Teng (2021) demonstrated the poor performance of the WMPR model in estimating robot strengths and predicting match scores (or outcomes). This can be explained by a relatively large number of parameters (number of robots) as compared to the number of observations (number of matches) in each division. With the recognition of a clustering feature in estimated robot strengths, the WMPRC model, which is an extension of the WMPR model with the consideration of latent clusters of robots, was proposed by Lim, Chiang and Teng (2021) to alleviate the overparameterization problem. The number of candidate cluster regression functions is huge for each given number of clusters, albeit the WMPRC model has the form of clusterwise linear regression of Späth (1979). Therefore, the high computational complexity of the conditional mixture method of DeSarbo and Cron (1988) makes the estimation of robot strengths challenging. In the format of qualification stage data, there is also no advantage in introducing a pairwise fusion penalty method in the spirit of Bondell and Reich (2008). By applying the central linkage clustering method of Sokal (1958) on estimated robot strengths from the WMPR model, Lim, Chiang and Teng (2021) developed a simple method to determine WMPRC candidate models for different numbers of clusters. However, the performance of these competing models is highly related to imprecise initial estimates. To reduce this influence, the authors improved their first method by sequentially
merging two clusters based on estimated robot strengths from a WMPRC candidate model determined in the preceding hierarchical classification step. Two types of mean squared prediction error (MSPE) criteria were further established to select optimal models from these WMPRC candidate models. As shown in the simulation study, the MSPE measures are underestimated by their cross-validation estimates. Even though estimated MSPE curves exhibit the U-shaped patterns, the number of clusters is generally overestimated by their minimizers.

For the hierarchical classification in the second method of Lim, Chiang and Teng (2021), robots may be classified into incorrect clusters. To tackle this problem, a non-hierarchical classification is proposed to examine the possibility of reassigning robots to other clusters. With the determined clusters of robots in each hierarchical classification step, we fit all possible WMPRC models with a robot in a new cluster and the rest robots in the original clusters. The clusters of robots are further updated by applying the central linkage clustering method on estimated robot strengths corresponding to new clusters in these fitting WMPRC models. This process is repeated until updated clusters of robots are the same as those in the preceding iterations. Different from existing hierarchical classifications, some steps of our hierarchical classification are based on estimated robot strengths from WMPRC candidate models in their preceding non-hierarchical classification steps. To reduce overestimation of the number of clusters by the MSPE criteria, the corresponding MSPE-based Bayesian information (MSPEB) criteria are established as alternatives to select optimal models from WMPRC candidate models. The conducted simulation shows that the performance of the MSPE criteria can be enhanced by their MSPEB criteria when the proportion of sufficiently separated robot strengths, which are assessed by the ratios of pairwise differences between different robot strengths to the standard deviation of the error, is higher or lower than some threshold value. Briefly speaking, the developed method is formed by two parts: one is a combination of hierarchical and non-hierarchical classifications and the other is any one of the MSPE and MSPEB criteria. This new method generally outperforms its competitors in the format of qualification stage data from the FRC championships. In this study, we also propose an index to measure the nested relation between clusters from any two models and modify the rank correlation of Han (1987) to measure the monotonic association between robot strengths from any two models.

The rest of this article is organized as follows. Section 2 presents the WMPRC model and reviews existing estimation methods. An effective method is developed in Section 3 to estimate the model parameters. In Section 4, the methodology is further applied to qualification stage and playoff stage data from the 2018 and 2019 FRC Houston and Detroit championships. Moreover, a simulation is conducted to assess and compare the performance of our method and its competitors. Section 5 summarizes the achievements and findings and discusses future research directions.

2. WMPRC Model and Existing Estimation Methods. We define the following notations. Let $Y_s$ denote the difference in scores between the red and blue alliances in match $s$; $D_s$ a binary outcome ($1=$win, $0=$loss) in match $s$; $O_i$ robot $i$; and $x_{si}$ the designated covariate of $O_i$ in match $s$ taking a value of $1$ if $O_i$ is assigned to the red alliance, $-1$ if $O_i$ is assigned to the blue alliance and $0$ otherwise, $s = 1, \ldots, M, i = 1, \ldots, K$.

2.1. Model Structure. Let $c_o$ and $g^o = (g^o_1, \ldots, g^o_K)$, respectively, represent the corresponding number of clusters and clusters of robots in the WMPRC model with $g^o_1, \ldots, g^o_K \in \{1, \ldots, c_o\}$. The relation between the difference in scores and robot strengths in this model is specified as follows:

$$Y = X\beta^o + \varepsilon,$$

(1)
where \( Y = (Y_1, \ldots, Y_M)^T \) is an \( M \times 1 \) response vector, \( X = (x_{si}) \) is an \( M \times K \) covariate matrix, \( \beta^o = (\beta^o_{g_1}, \ldots, \beta^o_{g_K})^T \) is a \( K \times 1 \) parameter vector of robot strengths and \( \varepsilon = (\varepsilon_1, \ldots, \varepsilon_M)^T \) is an \( M \times 1 \) vector of independent and identically distributed random errors from an unknown distribution function \( F^o(\cdot) \) with mean zero and variance \( \sigma^2 \). It is noted that the WMPR model is in the form of model (1) with \( c_o = K \). Since \( X \) is not a full-rank matrix, the constraint \( \sum_{i=1}^K \beta^o_{g_i} = 0 \) is imposed as most existing paired comparison models (see Bradley and Terry, 1952; Bradley, 1953; Cattelan, Varin and Firth, 2013). The regression coefficients \( \beta^o_{g_1}, \ldots, \beta^o_{g_K} \) are, thus, interpreted as the relative robot strengths. Just like in application to traditional sports, the primary research interest in the rating of robots in the FRC lies in predicting a match outcome. It follows from model (1) that the probability of \( \{D_s = 1\} \), given \( x_s = (x_{s1}, \ldots, x_{sK})^T \), can be expressed as

\[
P(D_s = 1|x_s) = 1 - F^o(-x_s^T \beta^o), \quad s = 1, \ldots, M.
\]

Since no assumption is made on the functional form of \( F^o \), the Bradley-Terry model (Bradley and Terry, 1952; Bradley, 1953) and the Thurstone-Mosteller model (Thurstone, 1927; Mosteller, 1951) are special cases of the above binary response model.

In fact, the perspective of the WMPRC model can be adopted to classify players into same skill levels in the United States Chess Federation (USCF) and the National Wheelchair Basketball Association (NWBA). In contrast with the rating systems in these applications, the number of clusters and clusters of robots in the WMPRC model are unknown in advance of data analysis. Furthermore, it may be inappropriate to determine the clusters of robots with some degrees of subjectivity. As shown by Lim, Chiang and Teng (2021), the estimation of \( \beta^o \) can be transformed into model selection from WMPRC candidate models of the form:

\[
Y = X \beta^c + \varepsilon \quad \text{with} \quad g^c,
\]

where \( \beta^c = (\beta^c_{g_1}, \ldots, \beta^c_{g_K})^T \) is a \( K \times 1 \) parameter vector of robot strengths, \( g^c = (g^c_1, \ldots, g^c_K) \) is a \( K \times 1 \) vector of clusters of robots with \( g^c_1, \ldots, g^c_K \in \{1, \ldots, c\} \) and \( \varepsilon_1, \ldots, \varepsilon_M \) in \( \varepsilon \) are from an unknown distribution \( F^c(\cdot) \), \( c = 2, \ldots, K \). The probability of \( \{D_s = 1\} \), given \( x_s \), is directly derived as

\[
P(D_s = 1|x_s) = 1 - F^c(-x_s^T \beta^c), \quad s = 1, \ldots, M.
\]

### 2.2. Classification and Model Selection

Although the estimation of parameters in (1) and (2) can be transformed into a model selection problem, it is impractical to fit a huge number of models \( (\sum_{c=1}^K \sum_{j=1}^{c-1} (-1)^j \binom{c}{j} (c-j)^K / c!) \) in the forms of (3) and (4). Especially, this requires much higher computational cost and longer computational time. To overcome such a limitation, Lim, Chiang and Teng (2021) proposed two simple methods to estimate the parameters \( c_o, g^o \) and \( \beta^o \). For a WMPRC candidate model \( Y = X \beta^c + \varepsilon \) with \( g^c \), the authors estimated \( \beta^c \) with the LSE \( \hat{\beta}^c \) and \( F^c(\cdot) \) with the empirical distribution function \( \hat{F}^c(\cdot) \) of the residuals \( e^c_s = Y_s - x_s^T \beta^c \), \( s = 1, \ldots, M \).

Let \( Y_0 \) be a future observation of the difference in scores on a given covariate vector \( x_0 \), which is defined as \( x_{1}, \ldots, x_M \), with the corresponding \( D_0 \) for a match outcome. For the predictors \( \hat{Y}_0 = x_0^T \hat{\beta}^c, \hat{p}_0 = 1 - \hat{F}^c(\hat{Y}_0) \) and \( D_0 = I(\hat{p}_0 > 0.5) + 0.5I(\hat{p}_0 = 0.5) \), their predictive capacities can be assessed by

\[
\begin{align*}
\text{MSPE}_Y(c, g^c) &= E\left[ (Y_0 - \hat{Y}_0)^2 \right], \\
\text{MSPE}_D(c, g^c) &= E\left[ (D_0 - \hat{p}_0)^2 \right], \text{ and} \\
\text{MSPE}_E(c, g^c) &= E\left[ (D_0 - D_0)^2 \right],
\end{align*}
\]

respectively.
It is noted that the measure \((1 - \text{MSPE}_D(c, g^c))\) is an alternative expression for the probability of correct prediction (PCP), i.e.
\[
\text{PCP}(c, g^c) = P\left( (\hat{p}_0 - 0.5)Y_0 > 0 \right) + 0.5P\left( (\hat{p}_0 - 0.5)Y_0 = 0 \right).
\]

Based on qualification stage data \(\{(Y_s, x_s)\}_{s=1}^M\), the following cross-validation estimators are naturally proposed for the corresponding MSPE measures in (5):
\[
\begin{align*}
\text{MSPE}_Y(c, g^c) &= \frac{1}{M} \sum_{s=1}^M \left( Y_s - \hat{Y}_s \right)^2, \\
\text{MSPE}_D(c, g^c) &= \frac{1}{M} \sum_{s=1}^M \left( D_s - \hat{D}_s \right)^2,
\end{align*}
\]

where \(\hat{Y}_s = x_s^T \hat{\beta}_s^c, \hat{p}_s = 1 - \hat{F}_s^c(-\hat{Y}_s), \hat{D}_s = I(\hat{p}_s - 0.5 > 0) + 0.5I(\hat{p}_s - 0.5 = 0)\) and \((\hat{\beta}_s^c, \hat{F}_s^c)\) is computed as \((\hat{\beta}^c, \hat{F}^c)\) with \((Y_s, x_s)\) excluded, \(s = 1, \ldots, M\). From the context of regression analysis, these counterparts can be directly obtained by using the algebraic relation between \(\hat{\beta}_s^c\) and \((\hat{\beta}^c, X, e^c)\) for each \(s\).

For the estimation of the WMPRC model, Lim, Chiang and Teng (2021) applied the centroid linkage clustering method of Sokal (1958) on the LSE \(\hat{\beta}^K\) of \(\beta^K\) in the WMPR model \(Y = X\beta^K + \varepsilon\) to determine the clusters of robots \(g^c\) in \(Y = X\beta^c + \varepsilon\) for each \(c\). An optimal model is further selected from these WMPRC candidate models with \(\max_c \text{PCP}(c, g^c) (= 1 - \min_c \text{MSPE}_D(c, g^c))\) or \(\min_c \text{MSPE}_Y(c, g^c)\). In practice, the estimated function \(\text{MSPE}_Y(c, g^c)\) can be used as another criterion for model selection. To avoid over-relying on an imprecise estimator \(\hat{\beta}^K\), the first method was improved by sequentially determining the clusters of robots \(g^c\) in \(Y = X\beta^c + \varepsilon\) based on the LSE \(\hat{\beta}^{c+1}\) of \(\beta^{c+1}\) in \(Y = X\beta^{c+1} + \varepsilon\) with \(g^{c+1}\) for \(c\) from \(K - 1\) to 2. In an application to playoff data from the 2018 FRC Houston and Detroit championships, the second method was found to have better performance in the prediction of match scores and outcomes. The data and R codes for both the methods are also available at http://www.runmycode.org/companion/view/4012.

**Remark 2.1.** Given the number of clusters \(c_0\), the WMPRC model has the form of clusterwise linear regression of Späth (1979). With a specified probability density function \(f^o(\cdot)\) of the error \(\varepsilon\), the unconditional mixture likelihood function of qualification stage data in each division of the FRC championships can be derived as follows:
\[
L(\beta^o_1, \ldots, \beta^o_{c_0}, f^o, p_1, \ldots, p_{c_0})
= \sum_{g_1=1}^{c_0} \cdots \sum_{g_{K}=1}^{c_0} \left( \prod_{s=1}^{M} f^o(Y_s - \beta^o_{g_1, s} \cdot x_{s1} - \cdots - \beta^o_{g_K, s} \cdot x_{sK}) \right) p_k^c \sum_{k=1}^K I(g_i = k),
\]
where \(p_k \in (0, 1)\) is the probability of \(\{g^o_i = k\}\) for each \(i = 1, \ldots, M, k = 1, \ldots, c_0\), with the constraint \(\sum_{k=1}^{K} p_k = 1\). However, the maximum likelihood method is unsuitable for estimating robot strengths due to the unknown \(c_0\) and \(f^o\) in the proposed model and the computational complexity in the maximization of (8).

**3. New Estimation Method.** We first introduce the essential elements of the developed method for identifying clusters of robot strengths. A systematic estimation procedure and indices of nested relation and monotonic association are further proposed in the succeeding subsections. Given clusters of robots \(g^c\), the \(k\)th cluster \(\{O_i : g^c_i = k, i = 1, \ldots, K\}\) is hereinafter denoted by \(G^c_k, k = 1, \ldots, c\).
3.1. Background. In the second method of Lim, Chiang and Teng (2021), which is referred to as the LCT method, the clusters of robots are sequentially updated based on estimated robot strengths from a WMPRC candidate model determined in the preceding hierarchical classification step. For each WMPRC model determined by our hierarchical classification, which is different from that in the LCT method, a non-hierarchical classification is further proposed to examine the possibility of reassigning robots to other clusters. As alternatives to the MSPE criteria, the MSPEB criteria are also established to select optimal models from WMPRC candidate models determined by the proposed classification system, a combination of hierarchical and non-hierarchical classifications.

In contrast with the hierarchical classification in the LCT method, our hierarchical classification is performed on the WMPR model $Y = X\beta^c + \varepsilon$ and WMPRC models $Y = X\beta^c + \varepsilon$ with $g^c$ determined in the preceding hierarchical classification step for $c = K - 2$ and $K - 3$ and the preceding non-hierarchical classification step for $c$ from $K - 4$ to 2. Based on the distances between $\hat{\beta}_i^{c+1}$ and $\hat{\beta}_j^{c+1}$, $1 \leq i < j \leq c + 1$, in the LSE $\hat{\beta}^{c+1}$ of $\beta^{c+1}$, a pair of clusters, say $(G^c_{\ell}, G^c_m)$, with the minimal distance is merged into one cluster. The clusters of robots $g^c$ are, thus, determined by the clusters $\{G^c_1, \ldots, G^c_c\}$ with $G^c_{\ell} = G^c_{\ell} + G^c_m$ and $G^c_k$, $k = 2, \ldots, c$, in one-to-one correspondence with the rest $G^c_{k+1}$s.

For a WMPRC model $Y = X\beta^c + \varepsilon$ with $g^c$, $\beta^c$ is naturally estimated by the LSE $\hat{\beta}^c$. A full description of the proposed hierarchical classification is given in Algorithm 1.

**Algorithm 1:** The hierarchical classification of robots

```plaintext
if $c = K$ then
    $g^c \leftarrow (1, \ldots, c)^\top$;
    Fit $Y = X\hat{\beta}^c + \varepsilon$ and compute $\hat{\beta}^c$, $\hat{\text{MSPE}}(c, g^c)$ and $\hat{\text{MSPEB}}(c, g^c)$;
else
    $(\ell, m) \leftarrow \arg \min_{i < j} |\hat{\beta}_i^{c+1} - \hat{\beta}_j^{c+1}|$;
    Determine $g^c$ by $\{G^c_1, \ldots, G^c_c\}$ with $G^c_{\ell} = G^c_{\ell} + G^c_m$ and $G^c_k$, $k = 2, \ldots, c$, in one-to-one correspondence with the rest $G^c_{k+1}$s;
    Fit $Y = X\hat{\beta}^c + \varepsilon$ with $g^c$ and compute $\hat{\beta}^c$, $\hat{\text{MSPE}}(c, g^c)$ and $\hat{\text{MSPEB}}(c, g^c)$;
end
```

To reduce the misclassification of robots in our hierarchical classification, a non-hierarchical classification is performed on WMPRC models $Y = X\beta^c + \varepsilon$ with $g^c$ determined in the hierarchical classification step for $c$ from $K - 2$ to 2. In this part, estimated robot strengths are retained the same for robots in $\{G^c_k : |G^c_k| = 1, k = 1, \ldots, c\}$. The strength of robot $i$ in $\{G^c_k : |G^c_k| \geq 2, k = 1, \ldots, c\}$ is reestimated by the $i$th element of the LSE of $\beta^{c+1}$ in a WMPRC model $Y = X\beta^{c+1} + \varepsilon$ with $g^{c+1}$ set to $g^c$ and the $i$th element of $g^{c+1}$ replaced by $c + 1$. The original clusters of robots $g^c$ are further updated by performing the centroid linkage clustering method on these estimated robot strengths. Thus, new estimated robot strengths can be directly obtained from the LSE of the regression coefficients in the resulting WMPRC model. This process is repeated until updated clusters of robots are the same as those in the preceding iterations. In practice, such a non-hierarchical classification greatly reduces the misclassification of robots when the number of clusters $c$ $(\geq c_o)$ is relatively small compared to the number of matches $M$. The details of our non-hierarchical classification are described in Algorithm 2.
WMPRC candidate models. The basic idea behind such criteria is to introduce penalties for overparameterized models. Thus, we can estimate the parameters $c_0$, $g^o$ and $\beta^o$ with $\tilde{c} = \text{arg min}_c \overline{\text{MSPE}}(c, g^c)$, $g^c$ and the LSE $\tilde{\beta}^c$, respectively, where $\overline{\text{MSPE}}(c, g^c)$ is one of the estimated functions in (9). In our data analysis and simulation study, all estimated MSPE and MSPEB functions exhibit U-shaped patterns over $c$. The established criteria are also found to be useful for selecting optimal models from WMPRC candidate models.

**Algorithm 2:** The non-hierarchical classification of robots

```plaintext
r \leftarrow 0, \tilde{\beta}^c_{(0)} \leftarrow \tilde{\beta}^c, \epsilon \leftarrow \text{tolerance value};
```

repeat

```
r \leftarrow r + 1;
\tilde{\beta}^{(r)} \leftarrow \tilde{\beta}^{(r-1)};
```

```for i \in \{ C^c_k : |C^c_k| \geq 2, k = 1, \ldots, c \} do
```

```
g^{c+1} = \text{arg min}_g \text{MSPEB}(\hat{\beta}(r), g);
g^c_i \leftarrow c + 1;
```

```Replace the ith element of $\tilde{\beta}(r)$ by the ith element of the LSE of $\beta^{c+1}$ in $Y = X\beta^{c+1} + \epsilon$
```

```with $g^{c+1}$;
```

```end
```

Update $g^c$ by performing the centroid linkage clustering method on $\tilde{\beta}(r)$ and compute the LSE $\tilde{\beta}^c(r)$ of $\beta^c$ in $Y = X\beta^c + \epsilon$ with $g^c$;

```
until \min_{0 < r_o < r} \left\| \tilde{\beta}^c(r) - \tilde{\beta}^c(r_o) \right\| < \epsilon;
```

```
Fit Y = X\beta^c + \epsilon with g^c determined by $\tilde{\beta}^c(r)$ (the LSE $\tilde{\beta}^c$ of $\beta^c$);
```

**Remark 3.1.** Similar to the mean-regression-based method of Lin and Ng (2012), any one of the MSPE estimates in (7) can be used as an assessment for merging two clusters and reassigning robots to other clusters. However, there may be more than one WMPRC model with the same minimal MSPE estimate for some numbers of clusters. Furthermore, more computational time is needed to obtain WMPRC candidate models. As shown in the simulation study, all the MSPE measures are also underestimated by their cross-validation estimates in the format of qualification stage data from the 2018 and 2019 FRC Houston and Detroit championships. Therefore, the MSPE estimates are inappropriate to serve as assessments for the classification of robots.

By performing the proposed classification system, we can determine WMPRC candidate models $Y = X\beta^c + \epsilon$ with $g^c$, $c = 2, \ldots, K$. Same as the model selection criteria in the LCT method, the parameters $c_0$, $g^c$ and $\beta^o$ in model (1) are further estimated by $\tilde{c} = \text{arg min}_c \overline{\text{MSPE}}(c, g^c)$, $g^c$ and the LSE $\tilde{\beta}^c$, respectively, where $\overline{\text{MSPE}}(c, g^c)$ is one of the estimated functions in (7). Since these MSPE criteria tend to overestimate $c_0$, their alternatives are established based on the following estimated MSPEB functions:

$$\overline{\text{MSPEB}}_{\gamma}(c, g^c) = \ln \left( \overline{\text{MSPE}}_{\gamma}(c, g^c) \right) + \frac{c \ln(M)}{M},$$

$$\overline{\text{MSPEB}}_{\nu}(c, g^c) = \ln \left( \overline{\text{MSPE}}_{\nu}(c, g^c) \right) + \frac{c \ln(M)}{M},$$

and

$$\overline{\text{MSPEB}}_{\varnothing}(c, g^c) = \ln \left( \overline{\text{MSPE}}_{\varnothing}(c, g^c) \right) + \frac{c \ln(M)}{M}, c = 2, \ldots, K.$$ 

As in the Schwarz criterion (Schwarz, 1978) for linear models with normal errors, the basic idea behind such criteria is to introduce penalties for overparameterized models. Thus, we can estimate the parameters $c_0$, $g^o$ and $\beta^o$ with $\tilde{c} = \text{arg min}_c \overline{\text{MSPE}}(c, g^c)$, $g^c$ and the LSE $\tilde{\beta}^c$, respectively, where $\overline{\text{MSPE}}(c, g^c)$ is one of the estimated functions in (9). In our data analysis and simulation study, all estimated MSPE and MSPEB functions exhibit U-shaped patterns over $c$. The established criteria are also found to be useful for selecting optimal models from WMPRC candidate models.
3.2. Estimation Procedure. To enhance performance of the LCT method, we develop a more effective method. The first part of this method is to perform a classification system for the determination of candidate models. Another part is to establish the MSPE and MSPEB criteria for the selection of optimal models from these candidate models.

To avoid verbosity, the subscripts Y, P and D in the MSPE and MSPEB functions (including their estimates) are dropped in the following algorithm of the proposed estimation procedure:

**Algorithm 3:** Algorithm of the proposed estimation procedure

```
begin
for c ← K to 2 do
    if c = K then
        \( g^c \leftarrow (1, \ldots, c)^\top \);
        Fit \( Y = X\hat{\beta}^c + \varepsilon \) and compute \( \hat{\beta}^c, \text{MSPE}(c, g^c) \) and \( \text{MSPEB}(c, g^c) \);
    else
        \( (\ell, m) \leftarrow \arg\min_{i<j} |\hat{\beta}_{i}^{c+1} - \hat{\beta}_{j}^{c+1}| ; \)
        Determine \( g^c \) by \( \{G^c_1, \ldots, G^c_c\} \) with \( G^c_1 = G^c_c \cup G^c_m \) and \( G^c_k, k = 2, \ldots, c \), in one-to-one correspondence with the rest \( G^c_k \)’s;
        Fit \( Y = X\hat{\beta}^c + \varepsilon \) with \( g^c \) and compute \( \hat{\beta}^c, \text{MSPE}(c, g^c) \) and \( \text{MSPEB}(c, g^c) \);
    end
    if \( c \leq K - 2 \) then
        \( r \leftarrow 0, \tilde{\beta}^{(r)}_0 \leftarrow \tilde{\beta}^c, \epsilon \leftarrow \text{tolerance value}; \)
        repeat
            \( r \leftarrow r + 1; \)
            \( \tilde{\beta}^{(r)} \leftarrow \tilde{\beta}^{(r-1)}; \)
            for \( i \in \{G^c_k : |G^c_k| \geq 2, k = 1, \ldots, c\} \) do
                \( g^{c+1} \leftarrow g^c; \)
                \( g^{c+1}_{i} \leftarrow \epsilon + 1; \)
                Replace the \( i \)th element of \( \tilde{\beta}^{(r)} \) by the \( i \)th element of the LSE of \( \beta^{c+1} \) in \( Y = X\beta^{c+1} + \varepsilon \) with \( g^{c+1} \);
            end
        until \( \min_{0 \leq r < \infty} \|\tilde{\beta}^{(r)} - \tilde{\beta}^{(r)}_{(0)}\| < \epsilon; \)
    Fit \( Y = X\beta^c + \varepsilon \) with \( g^c \) determined by \( \tilde{\beta}^{(r)} \) (the LSE \( \hat{\beta}^c \) of \( \beta^c \)) and compute \( \text{MSPE}(c, g^c) \) and \( \text{MSPEB}(c, g^c) \);
end
\hat{c} \leftarrow \arg\min_{c} \text{MSPE}(c, g^c), \tilde{c} \leftarrow \arg\min_{c} \text{MSPEB}(c, g^c);
Estimate \((c_o, g^{o}, \beta^{o})\) by \( (\hat{c}, \tilde{\beta}^c, \tilde{\beta}^c) \) or \( (\hat{c}, \hat{\beta}^c, \hat{\beta}^c) \);
end
```

The classification system in Algorithm 3 has been described in Algorithms 1 and 2.

Different from the proposed hierarchical classification, the clusters of robots \( g^c \) in the LCT method are determined by estimated robot strengths in the preceding hierarchical classification step for \( c \) from \( K - 1 \) to 2. By incorporating our non-hierarchical classification into the
LCT method, an alternative method can also be used to estimate the parameters \( c_0, g^0 \) and \( \beta^0 \). This estimation procedure is displayed in Algorithm 4 for clarity. The results are not presented because its performance is generally inferior to that of the developed method in the simulation study.

**Algorithm 4:** Algorithm of an alternative estimation procedure

begin
for \( c \leftarrow K \) to 2 do
  if \( c = K \) then
    \( g^c \leftarrow (1, \ldots, c)^\top; \)
    Fit \( Y = X \beta^c + \epsilon \) and compute \( \hat{\beta}^c \), \( \text{MSPE}(c, g^c) \) and \( \text{MSPEB}(c, g^c) \);
  else
    \((\ell, m) \leftarrow \arg \min_{i < j} |\hat{\beta}^c_i - \hat{\beta}^c_j|; \)
    Determine \( g^\ell \) by \( \{G^\ell_1, \ldots, G^\ell_c\} \) with \( G^\ell_1 = G^{c+1}_\ell \cup G^{c+1}_m \) and \( G^\ell_k, k = 2, \ldots, c \), in one-to-one correspondence with the rest \( G^{c+1}_k \)’s;
    Fit \( Y = X \beta^c + \epsilon \) with \( g^c \) and compute \( \hat{\beta}^c \), \( \text{MSPE}(c, g^c) \) and \( \text{MSPEB}(c, g^c) \);
  end
end
for \( c \leftarrow K-2 \) to 2 do
  \( r \leftarrow 0, \tilde{c}(0) \leftarrow \hat{\beta}^c, \epsilon \leftarrow \) tolerance value;
  repeat
    \( \hat{\beta}(r) \leftarrow \hat{\beta}(r-1) \);
    for \( i \in \{G^c_k : |G^c_k| \geq 2, k = 1, \ldots, c\} \) do
      \( g^{c+1} \leftarrow g^c \);
      \( g^{c+1}_i \leftarrow c + 1; \)
      Replace the \( i \)th element of \( \hat{\beta}(r) \) by the \( i \)th element of the LSE of \( \beta^{c+1} \) in 
      \( Y = X \beta^{c+1} + \epsilon \) with \( g^{c+1} \);
    end
    Update \( g^c \) by performing the centroid linkage clustering method on \( \hat{\beta}(r) \) and compute the
    LSE \( \hat{\beta}^c(r) \) of \( \beta^c \) in \( Y = X \beta^c + \epsilon \) with \( g^c \);
  until \( \min_{0 \leq r_0 < r} \| \hat{\beta}^c(r) - \hat{\beta}^c(r_0) \| < \epsilon; \)
  Fit \( Y = X \beta^c + \epsilon \) with \( g^c \) determined by \( \hat{\beta}^c(r) \) (the LSE \( \hat{\beta}^c \) of \( \beta^c \)) and compute \( \text{MSPE}(c, g^c) \)
  and \( \text{MSPEB}(c, g^c) \);
end
\( \hat{c} \leftarrow \arg \min_c \text{MSPE}(c, g^c), \tilde{c} \leftarrow \arg \min_c \text{MSPEB}(c, g^c); \)
Estimate \((c_0, g^0, \beta^0)\) by \( (\hat{c}, g^\hat{c}, \tilde{\beta}^c) \) or \( (\tilde{c}, g^\tilde{c}, \tilde{\beta}^c) \);
end

**Remark 3.2.** In the spirit of a pairwise fusion penalized least squares method of Bondell and Reich (2008), the robot strengths \( \beta^0 \) in model (1) can also be estimated by a minimizer of the following objective function with the constraint \( \sum_{i=1}^K \beta_i = 0; \)

\[
\frac{1}{2} (Y - X \beta)^\top (Y - X \beta) + \sum_{1 \leq i < j \leq K} p(|\beta_i - \beta_j|, \lambda),
\]

(10)
where $\lambda$ is a non-negative tuning parameter and $p(\cdot, \lambda)$ is a penalty function such as the smoothly clipped absolute deviation of Fan and Li (2001) and the minimax concave penalty of Zhang (2010). In implementation, this minimization is achieved by the alternating direction method of multipliers of Boyd et al. (2011) and the penalty parameter is chosen by the traditional BIC of Schwarz (1978). A model selected by the best criterion in the developed method outperforms its estimated model in the data analysis (see Table S1), albeit an estimator obtained from the criterion in (10) enjoys the oracle properties under some regularity conditions. The conducted simulation also shows that our method provides better estimation of parameters and more accurate prediction of match scores (or outcomes) in the format of qualification stage data (see Table S2). The data analysis and simulation results of this pairwise fusion method are displayed in the supplementary material.

3.3. Matching Index of the Nested Relation and Rank Correlation. To measure the nested relation between clusters of the WMPRC models $Y = X\beta^c + \varepsilon$ with $g^c$ and $Y = X\beta^d + \varepsilon$ with $g^d$, a matching index of the nested relation (MINR) is proposed as follows:

\begin{equation}
\text{MINR}(g^c, g^d) = \frac{1}{K} \sum_{i=1}^{K} \left( I(\hat{g}^c_i = g^d_i) I(c > d) + I(\hat{g}^c_i = g^d_i) I(c < d) + \frac{I(\hat{g}^c_i = g^d_i) + I(\hat{g}^c_i = g^d_i)}{2} I(c = d) \right),
\end{equation}

where $\hat{g}^c_i = \ell$ if $|\hat{\beta}^c_{g^c_i} - \hat{\beta}^d_\ell| = \min_{1 \leq j \leq d} |\hat{\beta}^c_{g^c_i} - \hat{\beta}^d_j|$ and $\hat{g}^d_i = m$ if $|\hat{\beta}^d_{g^d_i} - \hat{\beta}^c_\ell| = \min_{1 \leq j \leq c} |\hat{\beta}^d_{g^d_i} - \hat{\beta}^c_j|$, $i = 1, \ldots, K$, $\ell = 1, \ldots, d$, $m = 1, \ldots, c$. Given the information of $c$ and $d$, one of the models with a smaller number of clusters is treated as a reference model. The MINR investigates whether another model is in a more general form. When both models have the same number of clusters, they serve as a reference model of each other. In our simulation study, the true WMPRC model is treated as a reference model. The main purpose is to introduce penalties for underparameterized models. With this consideration, the MINR is modified as

\begin{equation}
\text{MINR}(g^c, g^d) = \frac{1}{K} \sum_{i=1}^{K} I(\hat{g}^c_i = g^d_i),
\end{equation}

where $\hat{g}^c_i = \ell$ if $|\hat{\beta}^c_{g^c_i} - \hat{\beta}^d_\ell| = \min_{1 \leq j \leq c} |\hat{\beta}^c_{g^c_i} - \hat{\beta}^d_j|$, $i = 1, \ldots, K$, $\ell = 1, \ldots, c_0$.

In application, the rank correlation (RC) of Han (1987) has been used by Lim, Chiang and Teng (2021) to measure the monotonic association between robot strengths from any two models. Since robot strengths are related to their intrinsic ranks, it is inappropriate to take a weight of 0.5 for those robots assigned to the same clusters. For the considered ordinal scale measurements, we modify their rank correlation estimate of $\beta^c$ and $\beta^d$ as follows:

\begin{equation}
\hat{RC}(\beta^c, \beta^d) = \frac{1}{K(K-1)} \sum_{i \neq j} \left( I\left( \left( \hat{\beta}^c_{g^c_i} - \hat{\beta}^c_{g^c_j} \right) \left( \hat{\beta}^d_{g^d_i} - \hat{\beta}^d_{g^d_j} \right) > 0 \right) + I(\hat{\beta}^c_{g^c_i} = \hat{\beta}^c_{g^c_j} \text{ and } \hat{\beta}^d_{g^d_i} = \hat{\beta}^d_{g^d_j}) \right).
\end{equation}

Naturally, a rank correlation estimate of $\beta^c$ and $\beta^o$ can be computed as in (13) with $\hat{\beta}^d_{g^d_i}$'s replaced by $\hat{\beta}^o_{g^c_i}$'s.

**Remark 3.3.** In principle, the rank-based index is not designed to measure the nested relation between clusters from any two models. As shown in Figures 6 and 7, the strength of monotonic association is not necessarily implied by the strength of nested relation and vice versa.
4. Data Analysis and Simulation Study. We applied the developed method (hereinafter referred to as the TCL method) and the LCT method of Lim, Chiang and Teng (2021) to qualification stage and playoff stage data from the 2018 and 2019 FRC Houston and Detroit championships. A Monte Carlo simulation was further conducted to investigate the performance of both the methods in identifying clusters of robot strengths and predicting an outcome, outcome probabilities and a difference in scores.

4.1. Application. In the divisions of the 2018 and 2019 FRC Houston and Detroit championships, there were 110 to 114 scheduled matches involving 66 to 68 robots in the qualification stage. Under the playoff system of the FRC, 14 to 18 matches were played by top eight robotics teams and their self-selected alliance partners. Details of the data description and processing can be found in Lim, Chiang and Teng (2021). In the data analysis, one match with extremely high scores in the Turing division of the 2018 FRC Houston championship was removed to ensure better quality of data. We further computed the MSPE and MSPEB estimates of each WMPRC candidate model as in (7) and (9) and the MINR and RC of two selected optimal models as in (11) and (13).

In Figures 2 – 5, it can be seen that the estimated MSPE and MSPEB curves are slightly serrated convex functions of the number of clusters \( c \). Except for the MSPE\( D \) criterion, the rest criteria have unique minimizers. For those WMPRC candidate models with the minimal MSPE\( D \) estimate, a model with the smallest number of clusters is selected in implementation. The parameters of selected models may be very different, albeit any one of estimated MSPE curves in the LCT method is close to the corresponding one in the TCL method. With the same MSPE criterion for model selection, Tables 1 – 6 show that the averages of estimated numbers of clusters, which were computed over the divisions in the same tournament, are comparable in both the LCT and TCL methods. In the TCL method, the corresponding MSPEB criteria of the MSPE criteria are further found to have noticeably smaller averages of estimated numbers of clusters. The MSPE estimates of a model selected by any one of the MSPE criteria in the LCT method are comparable to the corresponding estimates of a model selected by the same criterion in the TCL method. Moreover, models selected by the MSPE\( D \) (or PCP), MSPE\( P \) and MSPE\( Y \) criteria have the minimal MSPE\( D \) (or maximal PCP), MSPE\( P \) and MSPE\( Y \) estimates, respectively. It is noted that the MSPE estimates of the WMPR model (see Table 7) are much larger than the corresponding estimates of selected optimal models in both the LCT and TCL methods.

As shown in Tables 3, 6 and 7, models selected by the MSPE\( Y \) criterion in both the LCT and TCL methods generally have better predictions of semifinalists, finalists and champions in the playoffs than the WMPR model. Compared to models selected by any one of the MSPE criteria in the LCT method, models selected by the same criterion in the TCL method have comparable or even better predictive performance for the 2018 and 2019 FRC championships. For models selected by one of the MSPE\( D \) and MSPE\( Y \) criteria in the TCL method, those selected by the corresponding MSPEB criteria have better predictive capacities for the 2018 FRC championships but relatively poor predictive capacities for the 2019 FRC championships. Since models selected by the MSPEB\( P \) criterion in the TCL method are accompanied by extremely large MSPE\( Y \) estimates for the Galileo and Darwin divisions of the 2018 FRC championships, the MSPE\( P \) criterion was used to select optimal models in these two divisions. For the 2019 FRC championships, models selected by the MSPEB\( P \) criterion also outperform those selected by the MSPEB\( P \) criterion. Overall, the findings are consistent with the simulation results in the next subsection.

By borrowing a rule of Hosmer, Lemeshow and Sturdivant (2013), clusters of two selected optimal models are regarded to have an outstanding nested relation for \( \text{MINR} \geq 0.9 \); an excellent nested relation for \( 0.8 \leq \text{MINR} < 0.9 \); an acceptable nested relation for \( 0.7 \leq \text{MINR} < 0.8 \); and a poor nested relation for \( 0.6 \leq \text{MINR} < 0.7 \).
FIG 2. The estimated MSPE curves (black line) and MSPEB curves (red line), accompanied with their minima (black and red dots), in the TCL method for the divisions of the 2018 FRC Houston championship.

FIG 3. The estimated MSPE curves (black line) and MSPEB curves (red line), accompanied with their minima (black and red dots), in the TCL method for the divisions of the 2018 FRC Detroit championship.
Fig 4. The estimated MSPE curves (black line) and MSPEB curves (red line), accompanied with their minima (black and red dots), in the TCL method for the divisions of the 2019 FRC Houston championship.

Fig 5. The estimated MSPE curves (black line) and MSPEB curves (red line), accompanied with their minima (black and red dots), in the TCL method for the divisions of the 2019 FRC Detroit championship.
MINR < 0.8; and a poor nested relation for MINR < 0.7. It is further used to classify the strength of the RC for the monotonic association between robot strengths of two selected optimal models. Figures 6 and 7 display, respectively, the averages of MINR’s and RC’s over the divisions in Houston and Detroit. There exist outstanding MINR’s of models selected by the MSPE criteria in the LCT method for the 2018 and 2019 FRC championships. These values are further found to be higher than those of models, which have outstanding or excellent MINR’s, selected by the criteria in the TCL method. Moreover, the MINR of models selected by any two of the MSPE criteria in the LCT method for the 2018 FRC champi-
onships is generally higher than that of models selected by the same criteria for the 2019 FRC championships. This conclusion can also be made for the MINR of models selected by the same MSPE criterion in both the LCT and TCL methods for the 2018 and 2019 FRC championships. For the 2019 FRC championships, clusters of some selected optimal models have only acceptable nested relations. Figures 6 and 7 show that robot strengths from any two models with an outstanding or excellent MINR have outstanding or excellent monotonic associations. Any two selected optimal models with an acceptable MINR are further found to have an excellent RC. As for models selected by any two criteria in either the LCT method

---

### Table 3

| Season   | Tournament | Division | \( \hat{c} \) | PCP | MSPEP | MSPEY | Semifinalists | Finalists | Champions |
|----------|------------|----------|--------------|-----|-------|-------|---------------|-----------|-----------|
| 2018     | Houston    | Carver   | 11           | 85.1% | 0.091 | 5641.4 | 3              | 0         | 1         |
|          |            | Galileo  | 8            | 87.7% | 0.101 | 7088.9 | 2              | 1         | 1         |
|          |            | Hopper   | 10           | 86.8% | 0.089 | 6634.4 | 3              | 2         | 0         |
|          |            | Newton   | 10           | 87.9% | 0.078 | 5700.7 | 3              | 2         | 1         |
|          |            | Roebling | 11           | 90.6% | 0.090 | 6105.8 | 2              | 1         | 0         |
|          |            | Turing   | 12           | 92.8% | 0.061 | 6285.1 | 2              | 1         | 1         |
| Detroit  | Archimedes| 9         | 86.8%        | 0.100 | 5716.8 |       | 4              | 0         | 1         |
|          |            | Carson   | 10           | 85.5% | 0.101 | 5984.5 | 4              | 0         | 1         |
|          |            | Cuine    | 10           | 84.8% | 0.092 | 4993.6 | 4              | 1         | 1         |
|          |            | Daly     | 11           | 88.2% | 0.072 | 6479.8 | 2              | 2         | 0         |
|          |            | Darwin   | 8            | 85.3% | 0.118 | 7499.7 | 0              | 1         | 0         |
|          |            | Tesla    | 12           | 91.1% | 0.071 | 5117.4 | 2              | 1         | 0         |
| Average  |            |           | 10.2         | 87.7% | 0.089 | 6109.0 | 2.6            | 1.1       | 0.5       |

### Table 4

| Season   | Tournament | Division | \( \hat{c} \) | PCP | MSPEP | MSPEY | Semifinalists | Finalists | Champions |
|----------|------------|----------|--------------|-----|-------|-------|---------------|-----------|-----------|
| 2018     | Houston    | Carver   | 16           | 88.6% | 0.084 | 6087.9 | 6              | 86.8%     | 0.102     |
|          |            | Galileo  | 9            | 88.6% | 0.098 | 7541.4 | 9              | 88.6%     | 0.098     | 7541.4    | 2.2       | 2         | 2.1       |
|          |            | Hopper   | 10           | 86.8% | 0.100 | 6447.7 | 6              | 87.7%     | 0.102     | 7933.3    | 3.3       | 2         | 2.0       |
|          |            | Newton   | 19           | 92.4% | 0.069 | 6426.7 | 10             | 90.6%     | 0.067     | 5803.8    | 3.3       | 2         | 2.2       |
|          |            | Roebing  | 9            | 78.1% | 0.140 | 1159.2 | 9              | 87.3%     | 0.109     | 137.1     | 1         | 2         | 1.3       |
|          |            | Turing   | 9            | 87.3% | 0.091 | 137.1  | 9              | 87.3%     | 0.109     | 137.1     | 1         | 2         | 1.2       |
| Detroit  | Archimedes| 10        | 82.5%        | 0.111 | 160.1  |       | 4              | 86.8%     | 0.103     | 88.8      |
|          |            | Carson   | 9            | 86.0% | 0.103 | 6027.1 | 6              | 87.9%     | 0.100     | 7393.4    | 2.2       | 2         | 1.0       |
|          |            | Cuine    | 12           | 84.2% | 0.096 | 144.2  | 12             | 85.1%     | 0.097     | 114.4     | 0         | 1         | 0         |
|          |            | Daly     | 8            | 85.1% | 0.097 | 144.2  | 12             | 84.0%     | 0.105     | 89.1      | 2.3       | 1         | 0.0       |
|          |            | Darwin   | 10           | 85.1% | 0.104 | 128.4  | 10             | 85.1%     | 0.104     | 128.4     | 2         | 1         | 1.1       |
| Average  |            |           | 9.7          | 84.8% | 0.108 | 126.3  | 2.3            | 1.3       | 0.7       |

The numbers of clusters in models selected by the MSPE criterion in the LCT method, the PCP, MSPEP and MSPEY estimates of these selected optimal models and the numbers of correct predictions of semifinalists, finalists and champions in the playoffs for the 2018 and 2019 FRC championships.
or the TCL method, the RC for the 2018 FRC championships is generally higher than that for the 2019 FRC championships.

4.2. Monte Carlo Simulation. In this simulation study, data were similar in nature to qualification stage data for the divisions of the 2019 Houston and Detroit championships. We designed $c_\beta$, $\beta^2$, $\sigma^2$ and $X$ in the proposed WMPRC model based on $\hat{c}$, $\hat{\beta}$, $\hat{\sigma}^2 = \text{MSPE}_Y(\hat{c}, \hat{\beta}, \hat{\sigma}^2)$ and $X$ from the following selected optimal models:

### Table 5

| Season | Tournament | Division | Semifinalists | Finalists | Champions |
|--------|------------|----------|---------------|-----------|-----------|
| 2018   | Houston    | Carver   | 16            | 88.6%     | 0.084     |
|        |            | Galileo  | 9             | 88.6%     | 0.098     |
|        |            | Hopper   | 9             | 88.6%     | 0.083     |
|        |            | Newton   | 14            | 91.5%     | 0.067     |
|        |            | Rockefeller | 3          | 88.8%     | 0.094     |
|        |            | Turing   | 9             | 92.8%     | 0.059     |
|        | Detroit    | Archimedes | 9           | 87.7%     | 0.090     |
|        |            | Carson   | 13            | 87.3%     | 0.098     |
|        |            | Cane     | 9             | 85.7%     | 0.096     |
|        |            | Daly     | 8             | 89.9%     | 0.066     |
|        |            | Darwin   | 10            | 85.5%     | 0.112     |
|        |            | Tesla    | 7             | 87.8%     | 0.068     |
|        | Detroit    | Archimedes | 9           | 87.7%     | 0.100     |
|        |            | Carson   | 11            | 80.7%     | 0.133     |
|        |            | Galileo  | 9             | 90.6%     | 0.079     |
|        |            | Hopper   | 15            | 90.4%     | 0.093     |
|        |            | Newton   | 12            | 83.9%     | 0.107     |
|        |            | Rockefeller | 5          | 84.4%     | 0.124     |
|        |            | Turing   | 7             | 88.2%     | 0.090     |
|        | Detroit    | Archimedes | 11          | 87.7%     | 0.099     |
|        |            | Carson   | 10            | 83.3%     | 0.108     |
|        |            | Cane     | 10            | 86.6%     | 0.087     |
|        |            | Daly     | 8             | 89.5%     | 0.093     |
|        |            | Darwin   | 11            | 84.6%     | 0.107     |
|        |            | Tesla    | 9             | 87.7%     | 0.099     |

### Table 6

| Season | Tournament | Division | Semifinalists | Finalists | Champions |
|--------|------------|----------|---------------|-----------|-----------|
| 2019   | Houston    | Carver   | 11            | 80.7%     | 0.133     |
|        |            | Galileo  | 9             | 90.6%     | 0.107     |
|        |            | Hopper   | 15            | 90.4%     | 0.093     |
|        |            | Newton   | 12            | 83.9%     | 0.107     |
|        |            | Rockefeller | 5          | 84.4%     | 0.124     |
|        |            | Turing   | 7             | 88.2%     | 0.090     |
|        | Detroit    | Archimedes | 11          | 87.7%     | 0.100     |
|        |            | Carson   | 10            | 83.3%     | 0.108     |
|        |            | Cane     | 10            | 86.6%     | 0.087     |
|        |            | Daly     | 8             | 89.5%     | 0.093     |
|        |            | Darwin   | 11            | 84.6%     | 0.107     |
|        |            | Tesla    | 9             | 87.7%     | 0.099     |

Average | 10.1 | 84.9% | 0.106 | 127.3 | 7.8 | 84.7% | 0.110 | 133.6 | 2.3 | 2.5 | 1.2 | 1.1 | 0.7 | 0.6 |
FIG 6. The averages of the MINR’s of two selected optimal models over the divisions in the Houston and Detroit tournaments of the 2018 (upper triangular part) and 2019 (lower triangular part) FRC championships.

FIG 7. The averages of the RC’s of two selected optimal models over the divisions in the Houston and Detroit tournaments of the 2018 (upper triangular part) and 2019 (lower triangular part) FRC championships.
The number of robots, the PCP, MSPE and MSPEY estimates of the WMPP model and the numbers of correct predictions of semifinalists, finalists and champions in the playoffs for the 2018 and 2019 FRC championships.

| Season | Tournament | Division | \( K \) | \( \beta \) | MSPE | MSPEY | Semifinalists | Finalists | Champions |
|---------|------------|----------|--------|--------|-------|-------|--------------|-----------|-----------|
| 2018    | Houston    | Carver   | 68     | 0.229  | 26782.8 | 26782.8 | 3            | 0         | 1         |
|         |            | Galileo  | 68     | 0.241  | 33306.8 | 33306.8 | 2            | 2         | 1         |
|         |            | Hopper   | 68     | 0.233  | 30922.8 | 30922.8 | 3            | 2         | 0         |
|         |            | Newton   | 67     | 0.178  | 25571.7 | 25571.7 | 3            | 2         | 0         |
|         |            | Roebling | 67     | 0.251  | 28662.4 | 28662.4 | 2            | 1         | 0         |
|         |            | Turing   | 67     | 0.199  | 30199.0 | 30199.0 | 2            | 1         | 0         |
| Detroit | Archimedes | 68     | 0.248  | 26676.4 | 26676.4 | 4            | 0         | 1         |
|         | Carson     | 68     | 0.275  | 28789.7 | 28789.7 | 4            | 0         | 1         |
|         | Cuse       | 68     | 0.277  | 23039.6 | 23039.6 | 3            | 5         | 2         |
|         | Daly       | 68     | 0.216  | 29722.2 | 29722.2 | 2            | 1         | 0         |
|         | Darwin     | 68     | 0.295  | 36873.5 | 36873.5 | 0            | 1         | 0         |
|         | Tesla      | 68     | 0.189  | 25379.2 | 25379.2 | 3            | 1         | 0         |
| Average |            |         | 67.5   | 0.236  | 28827.2 | 28827.2 | 2.6          | 1.0       | 0.4       |
| 2019    | Houston    | Carver   | 68     | 0.312  | 912.2   | 912.2   | 4            | 1         | 1         |
|         |            | Galileo  | 67     | 0.234  | 498.2   | 498.2   | 2            | 2         | 0         |
|         |            | Hopper   | 68     | 0.257  | 469.6   | 469.6   | 4            | 1         | 1         |
|         |            | Newton   | 67     | 0.257  | 735.3   | 735.3   | 3            | 1         | 1         |
|         |            | Roebling | 67     | 0.283  | 559.0   | 559.0   | 3            | 2         | 0         |
|         |            | Turing   | 66     | 0.248  | 679.3   | 679.3   | 1            | 2         | 1         |
| Detroit | Archimedes | 68     | 0.292  | 718.3   | 718.3   | 2            | 1         | 1         |
|         | Carson     | 68     | 0.233  | 425.9   | 425.9   | 2            | 1         | 0         |
|         | Cuse       | 68     | 0.228  | 622.9   | 622.9   | 4            | 0         | 1         |
|         | Daly       | 68     | 0.266  | 556.4   | 556.4   | 0            | 1         | 0         |
|         | Darwin     | 68     | 0.222  | 388.2   | 388.2   | 1            | 1         | 1         |
|         | Tesla      | 68     | 0.233  | 609.8   | 609.8   | 2            | 1         | 1         |
| Average |            |         | 67.6   | 0.256  | 599.4   | 599.4   | 2.3          | 1.2       | 0.7       |

M1. The model, which is the same as those selected by the MSPEB and MSPEBY criteria, is selected by the MSPEB method in the TCL method for the Roebling division with \( \hat{\sigma} = 4, (\hat{\beta}^1, \ldots, \hat{\beta}^4) = (-15.07, -4.75, 4.76, 14.52), (|\hat{G}^1_1|, \ldots, |\hat{G}^4_1|) = (9, 25, 24, 9) \) and \( \hat{\sigma} = 11.056. \)

M2. The model, which is the same as that selected by the MSPEB criterion, is selected by the MSPEB method in the TCL method for the Daly division with \( \hat{\sigma} = 8, (\hat{\beta}^1, \ldots, \hat{\beta}^8) = (-18.16, -13.57, -9.81, -0.58, 4.88, 7.90, 11.16, 18.20), (|\hat{G}^8_1|, \ldots, |\hat{G}^8_8|) = (5, 3, 8, 22, 17, 6, 5, 2) \) and \( \hat{\sigma} = 10.275. \)

For models M1 and M2, the corresponding ratios of the differences between adjacent estimated robot strengths to \( \hat{\sigma} \) are \((0.93, 0.86, 0.88)\) and \((0.45, 0.37, 0.9, 0.53, 0.29, 0.31, 0.68)\). The error was further specified to follow a normal distribution with mean zero and variance \( \sigma^2 \). With this setup, the difference in scores \( Y \) was generated from model (1) with the designed model M1 or M2 for \( X \beta^o \) and one of the values in \( \{0.25\hat{\sigma}, 0.5\hat{\sigma}, \hat{\sigma}, 2\hat{\sigma}, 4\hat{\sigma}\} \) for \( \sigma \). To ensure the numerical stability, the results were based on 5000 replications.

In both the LCT and TCL methods, the MSPE measures of model M1 are almost the same as the minima of the corresponding MSPE curves for \( \sigma \leq 5.528 \) (Figure 8), whereas the MSPE measures of model M2 are close to the minima of the corresponding MSPE curves only for \( \sigma = 2.569 \) (Figure 10). This is mainly because the proportion of sufficiently separated robot strengths in model M1 is higher than that in model M2. As expected, the deviations between the MSPE measures of models M1 and M2 and the minima of the corresponding MSPE curves become larger as \( \sigma \) increases. Except for fairly large numbers of clusters, it can be seen in Figures 8 – 11 that the MSPE and MSPEB curves are higher than their estimated curves under models M1 and M2. With any one of the MSPE criteria in the TCL method, the MSPE measures of selected optimal models approach the minima of the corresponding MSPE curves. The minimizers of the MSPEB curves are also smaller than those of their estimated curves and the corresponding MSPE curves. The parameters in their selected optimal models may be very different, albeit each estimated MSPE curve in the LCT method is close to that in the TCL method. Especially, the MSPE measures and estimates of models M1 and M2 are, respectively, smaller and larger than the corresponding MSPE measures and
estimates of selected optimal models. Therefore, these estimates are inappropriate to assess
the goodness-of-fit for selected optimal models.

With the same MSPE criterion in both the LCT and TCL methods, their estimated numbers
of clusters in Table 8 are quite close to each other. The number of clusters in a model selected
by the MSPE$_D$ criterion is further found to be smaller than those in models selected by the
MSPE$_P$ and MSPE$_Y$ criteria, which have similar performance in estimating the number of
clusters. In the TCL method, the numbers of clusters in models selected by the MSPEB criteria
are shown to be noticeably smaller than those in models selected by the corresponding
MSPE criteria. Moreover, the MSPE$_P$ criterion produces the smallest number of clusters
and both the MSPE$_B$ and MSPE$_Y$ criteria have comparable numbers of clusters. As we
can see, the number of clusters in model M1 is overestimated by all the model selection crite-
rria but that in model M2 is only overestimated by the MSPE criteria. Due to a relatively low
proportion of sufficiently separated robot strengths in model M2, the number of clusters is
underestimated by the MSPE$_D$ and MSPE$_Y$ criteria for $\sigma \geq 20.549$ and by the MSPE$_P$
criterion for $\sigma \geq 10.275$. To achieve the minimal MSPE$_D$, MSPE$_P$, or MSPE$_Y$, this underes-
timation is attributed to the increase of biases and decrease of variances of estimated robot
strengths.

Tables 9 – 14 summarize the performance of the WMPR model in the estimation of robot
strengths and prediction of match scores (or outcomes). In Table 9, the mean squared error
(MSE) of estimated robot strengths from the WMPR model is significantly improved by the
TCL method with the MSPE$_P$ criterion for $\sigma \leq 5.528$ or $\sigma \geq 22.112$ under model M1 and
for $\sigma = 2.569$ or $\sigma \geq 20.549$ under model M2. The MSE of estimated robot strengths from
the WMPR model is, however, slightly smaller than that of estimated robot strengths from a

![Diagram](image-url)
IDENTIFYING CLUSTERS OF ROBOT STRENGTHS

**Fig 9.** The MSPEB curves (red line) and the averages of their 5000 estimated curves (black line) in the TCL method under model M1.

**Fig 10.** The MSPE curves and the averages of their 5000 estimated curves in both the LCT and TCL methods (LCT: red and black dashed lines; TCL: red and black solid lines) under model M2. The red and black dots, respectively, denote the MSPE measures of model M2 and the averages of their 5000 estimates.
model selected by the MSPE\(_Y\) criterion in the TCL method for \(\sigma = 11.056\) under model M1 and for \(5.137 \leq \sigma \leq 10.275\) under model M2. In an attempt to alleviate overparameterization of the WMPR model, the gain in identifying clusters of robot strengths still exceeds the loss in introducing additional biases and variances of estimated number of clusters and clusters of robots. The same conclusion can be further made for the MINR’s in Table 10. Compared to the RC of the true model and a model selected by the best criterion in the TCL method, the RC of the true model and the WMPR model (see Table 11) is notably lower for \(\sigma \leq 11.056\)
under model M1 and for $\sigma \leq 5.137$ under model M2; slightly lower for $\sigma = 22.112$ under model M1 and for $\sigma = 10.275$ under model M2; slightly higher for $\sigma = 44.223$ under model M1 and for $\sigma \geq 20.549$ under model M2. To reduce the MSE of estimated robot strengths from the WMPR model for the case of extremely large values of $\sigma$, the established criteria in the TCL method tend to increase the bias and decrease the variance of estimated robot strengths from a selected optimal model. As a result, it will lower the RC of the true model and a selected optimal model. In Tables 12 – 14, the MSPE measures of the WMPR model are found to be overestimated for all values of $\sigma$ under models M1 and M2. The conclusion for the MSE of estimated robot strengths can also be made for the predictive performance of the WMPR model and our selected optimal models in terms of the MSPE measures.

With the same MSPE criterion for model selection, Table 9 shows that estimated robot strengths in the LCT method can be improved by the proposed classification system in terms of the MSE. In the TCL method, the MSE of estimated robot strengths from a model selected by any one of the MSPEB criteria is smaller than that of estimated robot strengths from a model selected by the corresponding MSPE criterion for $\sigma \leq 5.528$ or $\sigma \geq 22.112$ under model M1 and for $\sigma = 2.569$ or $\sigma \geq 20.549$ under model M2. On the basis of this assessment, the MSPEB$P$ criterion outperforms the MSPEB$D$ and MSPEB$Y$ criteria for these values of $\sigma$ under models M1 and M2 and the MSPE$Y$ criterion outperforms the MSPE$D$ and MSPE$P$ criteria for the other values of $\sigma$. The MINR and RC of the true model and a model selected by any one of the MSPE criteria in the TCL method are generally comparable to or even higher than those of the true model and a model selected by the same criterion in the LCT method. Together with the information of estimated numbers of clusters in Table 8, the MSPEB criteria outperform the corresponding MSPE criteria in terms of the MINR in the TCL method for $\sigma \leq 5.528$ or $\sigma \geq 22.112$ under model M1 and for $\sigma \geq 20.549$ under model M2. Especially, the MSPEB$P$ criterion outperforms the MSPEB$D$ and MSPEB$Y$ criteria for these values of $\sigma$ under models M1 and M2 and the MSPE$Y$ criterion outperforms the MSPE$D$ and MSPE$P$ criteria for the other values of $\sigma$. Compared to any one of the MSPE criteria in the TCL method, the corresponding MSPE criterion has a higher RC for $\sigma \leq 11.056$ under model M1 and for $\sigma \leq 5.137$ under model M2. Except for $\sigma = 11.056$ under model M1, the MSPEB$P$ criterion outperforms the MSPEB$D$ and MSPEB$Y$ criteria in terms of the RC. Moreover, the MSPE$Y$ criterion produces the highest RC for the other values of $\sigma$.

| Model | $\sigma$ | WMPR | LCT method | TCL method |
|-------|---------|------|------------|------------|
|       |         | MSPE$P$ | MSPE$D$ | MSPE$Y$ | MSPE$P$ | MSPE$D$ | MSPE$Y$ | MSPE$P$ | MSPE$D$ | MSPE$Y$ |
| WMPR  |         | 2.764 | 65.529 | 72.692 | 82.870 | 60.595 | 47.971 | 68.194 | 39.004 | 79.212 | 52.500 |
|       |         | 5.528 | 344.336 | 353.795 | 368.416 | 303.266 | 257.720 | 312.996 | 229.984 | 354.855 | 291.340 |
|       |         | 11.056 | 1688.947 | 1599.061 | 1577.577 | 1588.265 | 1647.903 | 1576.448 | 1657.808 | 1573.168 | 1599.839 |
|       |         | 22.112 | 5983.565 | 6037.026 | 6072.589 | 5677.451 | 5195.265 | 5177.731 | 5145.647 | 5564.496 | 5562.063 |
|       |         | 44.223 | 22723.638 | 23151.075 | 23501.749 | 21053.862 | 17324.722 | 21898.330 | 16600.093 | 22884.581 | 19397.953 |

| M2    |         | 2.569 | 53.063 | 72.497 | 73.750 | 79.880 | 63.613 | 61.119 | 67.955 | 58.758 | 76.556 | 59.334 |
|       |         | 5.137 | 340.250 | 368.003 | 358.792 | 352.319 | 350.230 | 339.390 | 338.392 | 396.763 | 351.848 | 364.888 |
|       |         | 10.275 | 1360.710 | 1408.038 | 1397.063 | 1387.490 | 1391.760 | 1446.551 | 1383.382 | 1459.470 | 1379.467 | 1400.618 |
|       |         | 20.549 | 5442.841 | 5275.136 | 5319.028 | 5352.931 | 5022.987 | 4673.683 | 5137.169 | 4619.494 | 5246.074 | 4926.585 |
|       |         | 41.098 | 21776.030 | 20025.404 | 20417.138 | 20728.255 | 18601.295 | 15644.076 | 19293.720 | 14995.979 | 20118.878 | 17202.609 |

In Tables 12 – 14, the MSPE measures of selected optimal models are underestimated in both the LCT and TCL methods. This underestimation becomes severe as $\sigma$ increases. For selected optimal models with the minimal MSPE measures, the closest estimates are further found in the TCL method. Moreover, the MSPE$D$ of a model selected by any one of the MSPEB criteria is smaller than that selected by the corresponding MSPE criterion.
TABLE 10
The averages of 5000 MINR’s of the true model and the WMPR model and those of the true model and selected optimal models in both the LCT and TCL methods under models M1 and M2.

| Model | σ    | WMPR LCT method | | | | | TCL method | | | | |
|-------|------|-----------------|---|---|---|---|---|---|---|---|---|---|
|       |      | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE |
| M1    | 2.764 | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% |
|       | 5.528 | 96.4%  | 96.1%  | 96.2%  | 96.3%  | 96.9%  | 97.0%  | 96.7%  | 96.8%  | 96.5%  | 96.9%  | 96.8%  |
|       | 11.056 | 73.5%  | 72.2%  | 72.5%  | 72.9%  | 72.5%  | 72.6%  | 72.5%  | 72.5%  | 72.2%  | 72.7%  | 72.1%  |
|       | 22.112 | 47.3%  | 47.1%  | 47.2%  | 47.2%  | 47.7%  | 47.8%  | 47.4%  | 47.4%  | 47.2%  | 47.4%  | 47.7%  |
|       | 44.223 | 30.9%  | 31.1%  | 31.0%  | 31.0%  | 31.8%  | 32.6%  | 31.5%  | 32.7%  | 31.2%  | 32.0%  | 31.8%  |
| M2    | 2.569 | 94.4%  | 93.5%  | 94.0%  | 94.2%  | 94.4%  | 93.8%  | 94.6%  | 93.6%  | 94.5%  | 94.9%  | 94.9%  |
|       | 5.137 | 75.1%  | 73.0%  | 73.7%  | 74.1%  | 73.0%  | 71.2%  | 73.5%  | 71.0%  | 74.0%  | 72.8%  | 72.8%  |
|       | 10.275 | 48.5%  | 47.3%  | 47.6%  | 47.9%  | 47.4%  | 46.1%  | 47.7%  | 45.9%  | 47.9%  | 47.2%  | 47.2%  |
|       | 20.549 | 28.3%  | 28.4%  | 28.4%  | 28.4%  | 28.8%  | 29.0%  | 28.8%  | 29.0%  | 28.6%  | 28.8%  | 28.8%  |
|       | 41.098 | 16.9%  | 17.2%  | 17.1%  | 17.0%  | 17.6%  | 18.3%  | 17.5%  | 18.4%  | 17.3%  | 17.8%  | 17.8%  |

TABLE 11
The averages of 5000 RC’s of the true model and the WMPR model and those of the true model and selected optimal models in both the LCT and TCL methods under models M1 and M2.

| Model | σ    | WMPR LCT method | | | | | TCL method | | | | |
|-------|------|-----------------|---|---|---|---|---|---|---|---|---|---|
|       |      | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE | MSPE |
| M1    | 2.764 | 70.7%  | 83.9%  | 81.9%  | 79.6%  | 84.6%  | 88.1%  | 82.6%  | 90.5%  | 80.2%  | 87.0%  | 80.2%  |
|       | 5.528 | 70.6%  | 82.2%  | 80.6%  | 78.8%  | 83.4%  | 87.3%  | 81.7%  | 89.5%  | 79.4%  | 85.2%  | 79.4%  |
|       | 11.056 | 67.0%  | 71.7%  | 71.5%  | 71.3%  | 71.5%  | 71.8%  | 71.4%  | 72.0%  | 71.3%  | 72.3%  | 71.3%  |
|       | 22.112 | 58.0%  | 58.6%  | 58.7%  | 58.9%  | 58.6%  | 57.8%  | 58.8%  | 57.6%  | 58.8%  | 58.6%  | 58.6%  |
|       | 44.223 | 48.5%  | 47.5%  | 47.7%  | 47.9%  | 47.6%  | 46.5%  | 47.8%  | 46.1%  | 48.0%  | 47.1%  | 47.1%  |
| M2    | 2.569 | 80.8%  | 90.7%  | 88.9%  | 87.1%  | 89.6%  | 93.3%  | 89.9%  | 94.1%  | 87.6%  | 92.4%  | 87.6%  |
|       | 5.137 | 79.0%  | 83.4%  | 83.0%  | 82.7%  | 83.4%  | 83.7%  | 83.2%  | 84.0%  | 82.9%  | 84.1%  | 82.9%  |
|       | 10.275 | 73.2%  | 73.2%  | 73.4%  | 73.7%  | 73.0%  | 71.6%  | 73.4%  | 71.4%  | 73.6%  | 73.0%  | 73.0%  |
|       | 20.549 | 63.9%  | 60.8%  | 61.3%  | 61.8%  | 60.9%  | 58.4%  | 61.3%  | 57.6%  | 61.8%  | 59.9%  | 59.9%  |
|       | 41.098 | 54.3%  | 49.9%  | 50.4%  | 50.8%  | 49.9%  | 47.0%  | 50.3%  | 46.0%  | 50.8%  | 48.2%  | 48.2%  |

for $\sigma \leq 5.528$ under model M1 and for $\sigma = 2.569$ under model M2. Compared to a model selected by any one of the MSPE criteria, a model selected by the corresponding MSPEB criteria has smaller MSPE$^P$ and MSPE$^Y$ for $\sigma \leq 5.528$ or $\sigma \geq 22.112$ under model M1 and for $\sigma = 2.569$ or $\sigma \geq 20.549$ under model M2. Especially, a model selected by the MSPE$^B_P$ criterion has the minimal MSPE$^Y$ and the closest MSPE$^Y$ estimate. A model selected by the MSPE$^Y$ criterion also has this advantage for the other values of $\sigma$.

5. Concluding Remarks and Future Challenges. Compared to existing methods, a more effective method is developed to identify clusters of robot strengths in an application to qualification stage data from the 2018 and 2019 FRC Houston and Detroit championships. In our method, a classification system is first proposed to determine WMPRC candidate models. The designed non-hierarchical classification is found to be particularly useful for examining possible misclassification of robots in each hierarchical classification step. The MSPE and MSPEB criteria are further established to select optimal models from these candidate models. By coherently assembling the two parts, a systematic procedure is, thus, presented to estimate the number of clusters, clusters of robots and robot strengths. In addition to these achievements, the MINR and modified RC are provided, respectively, to measure the nested relation between clusters from any two models and monotonic association between robot strengths from any two models.

With the same MSPE criterion for model selection, the conducted simulation shows that our classification system can reduce the MSE of estimated robot strengths in the LCT method. The MINR and RC in the TCL method are further found to be comparable to or even higher than those in the LCT method. For selected optimal models with the minimal MSPE measures, which are underestimated in both the methods, the TCL method tends to produce the
Identifying Clusters of Robot Strengths

Table 12
The MSPE_D’s and the averages of their 5000 estimates (within parenthesis) of the WMPR model and selected optimal models in both the LCT and TCL methods under models M1 and M2.

| Model | σ  | WMPR | LCT method | TCL method |
|-------|----|------|-------------|------------|
|       |    |      | MSPE_D | MSPE_P | MSPE_Y | MSPE_D | MSPE_P | MSPE_Y | MSPE_D | MSPE_P | MSPE_Y |
| M1    | 2.764 | 0.101 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 |
|       |      | (0.102) | (0.041) | (0.048) | (0.055) | (0.043) | (0.046) | (0.056) | (0.055) | (0.061) | (0.056) |
| 5.528 | 0.121 | 0.121 | 0.120 | 0.120 | 0.120 | 0.121 | 0.120 | 0.120 | 0.120 | 0.120 |
|       | (0.142) | (0.055) | (0.063) | (0.070) | (0.054) | (0.058) | (0.061) | (0.072) | (0.070) | (0.075) |
| 11.056 | 0.213 | 0.217 | 0.216 | 0.216 | 0.216 | 0.215 | 0.216 | 0.215 | 0.215 | 0.215 |
|       | (0.252) | (0.104) | (0.114) | (0.120) | (0.103) | (0.111) | (0.113) | (0.124) | (0.121) | (0.126) |
| 22.112 | 0.340 | 0.343 | 0.342 | 0.342 | 0.342 | 0.340 | 0.341 | 0.341 | 0.341 | 0.341 |
|       | (0.376) | (0.162) | (0.175) | (0.181) | (0.161) | (0.172) | (0.174) | (0.191) | (0.182) | (0.190) |
| 44.223 | 0.439 | 0.440 | 0.440 | 0.440 | 0.440 | 0.439 | 0.440 | 0.440 | 0.440 | 0.440 |
|       | (0.459) | (0.203) | (0.217) | (0.223) | (0.202) | (0.215) | (0.216) | (0.238) | (0.224) | (0.235) |

Table 13
The MSPE_P’s and the averages of their 5000 estimates (within parenthesis) of the WMPR model and selected optimal models in both the LCT and TCL methods under models M1 and M2.

| Model | σ  | WMPR | LCT method | TCL method |
|-------|----|------|-------------|------------|
|       |    |      | MSPE_D | MSPE_P | MSPE_Y | MSPE_D | MSPE_P | MSPE_Y | MSPE_D | MSPE_P | MSPE_Y |
| M1    | 2.764 | 0.071 | 0.066 | 0.068 | 0.069 | 0.066 | 0.064 | 0.068 | 0.062 | 0.069 | 0.065 |
|       |      | (0.081) | (0.036) | (0.033) | (0.037) | (0.035) | (0.036) | (0.032) | (0.036) | (0.037) | (0.039) |
| 5.528 | 0.114 | 0.116 | 0.115 | 0.114 | 0.116 | 0.116 | 0.119 | 0.115 | 0.120 | 0.114 | 0.116 |
|       | (0.136) | (0.055) | (0.063) | (0.069) | (0.054) | (0.059) | (0.061) | (0.069) | (0.069) | (0.073) | (0.073) |
| 10.275 | 0.214 | 0.217 | 0.216 | 0.215 | 0.217 | 0.217 | 0.222 | 0.216 | 0.224 | 0.215 | 0.218 |
|       | (0.250) | (0.105) | (0.115) | (0.121) | (0.103) | (0.111) | (0.113) | (0.126) | (0.122) | (0.128) | (0.128) |
| 20.549 | 0.338 | 0.340 | 0.339 | 0.339 | 0.340 | 0.345 | 0.340 | 0.346 | 0.338 | 0.342 |
|       | (0.373) | (0.163) | (0.176) | (0.181) | (0.161) | (0.171) | (0.174) | (0.190) | (0.182) | (0.189) | (0.189) |
| 41.098 | 0.436 | 0.437 | 0.437 | 0.436 | 0.437 | 0.439 | 0.436 | 0.440 | 0.436 | 0.438 |
|       | (0.455) | (0.202) | (0.217) | (0.222) | (0.201) | (0.213) | (0.215) | (0.236) | (0.224) | (0.234) | (0.234) |

closest MSPE estimates. Especially, models selected by the MSPEB criteria are shown to have better performance than those selected by the corresponding MSPE criteria when the proportion of sufficiently separated robot strengths is higher or lower than some threshold value. Moreover, the MSPE measures and their estimates of a selected optimal model are larger and smaller than the corresponding MSPE measures and estimates of the true WMPRC model. This underestimation becomes more severe as the standard deviation of the error increases. The established MSPE or MSPEB criteria are still useful for selecting optimal models, albeit MSPE estimates are inappropriate to serve as assessments in the proposed classification system. It is notable that a pairwise fusion method, which is introduced in Remark 3.2, is superior to the developed method for strongly separated clusters of robot strengths. However, our method has been demonstrated to perform better on data similar to those of the qualification stage of the FRC championships and is recommended for adoption by FRC teams to evaluate potential alliance partners in the playoff stage.
Table 14

The MSPE'y's and the averages of their 5000 estimates (within parenthesis) of the WMPR model and selected optimal models in both the LCT and TCL methods under models M1 and M2.

| Model | \( \sigma \) | WMPR | LCT method | TCL method |
|-------|-------------|-------|-------------|-------------|
|       |             |       | MSPE \_d   | MSPE \_p   | MSPE \_Y   | MSPE \_d | MSPE \_p | MSPE \_Y |
| M1    | 2.764       | 12.152| 11.410      | 11.782      | 11.112      | 10.606   | 11.394  | 10.925  |
|       | (18.897)    | (5.261)| (5.069)     | (4.644)     | (5.257)     | (5.084)  | (5.888) | (4.645) |
|       | 5.528       | 48.608| 48.138      | 48.432      | 46.769      | 45.346   | 47.294  | 48.151  |
|       | (20.667)    | (18.666)| (18.166)   | (18.166)   | (20.626)   | (22.484) | (18.161) | (19.871) |
|       | 11.056      | 194.433| 198.969     | 197.359     | 200.611     | 197.107  | 197.880 |
|       | (302.355)   | (72.959)| (69.716)   | (66.651)   | (73.846)   | (83.688) | (70.063) | (70.688) |
|       | 22.112      | 777.312| 779.768     | 779.469     | 777.174     | 777.966  | 777.880 |
|       | (1209.420)  | (266.464)| (259.957) | (251.761) | (274.191)  | (264.440) | (254.748) | (268.698) |
|       | 44.223      | 3100.927| 3063.919    | 3075.108    | 3051.233    | 3066.674 | 3083.676 | 3018.019 |
|       | (4873.680)  | (1027.002)| (1004.256)| (961.964) | (1056.786) | (1150.639)| (995.138) | (1025.745) |
| M2    | 2.569       | 10.506| 10.412      | 10.411      | 10.176      | 10.333   | 10.192  | 10.307  |
|       | (16.182)    | (5.092)| (5.114)     | (4.443)     | (5.563)     | (6.295)  | (5.155) | (6.535) |
|       | 5.137       | 42.025| 43.461      | 42.740      | 44.243      | 46.575   | 43.539  | 46.768  |
|       | (64.728)    | (18.436)| (17.434)   | (16.131)   | (18.765)    | (21.493) | (17.847) | (21.712) |
|       | 10.275      | 168.102| 171.344     | 170.052     | 172.653     | 178.761  | 171.302 | 179.692 |
|       | (258.911)   | (60.034)| (60.892)   | (58.129)   | (64.607)    | (72.424) | (62.056) | (73.029) |
|       | 20.549      | 672.407| 672.921     | 671.913     | 670.860     | 670.683  | 670.228 | 670.457 |
|       | (1055.643)  | (231.244)| (224.506) | (218.217) | (236.383)  | (257.357) | (228.701) | (220.202) |
|       | 41.088      | 2689.630| 2651.076    | 2645.535    | 2630.466    | 2659.382 | 2643.275 | 2655.320 |
|       | (4142.571)  | (888.842)| (868.555) | (848.412) | (908.573)  | (969.931) | (894.930) | (857.026) |

Regarding the model structure, the WMPRC model describes the linear dependence of the difference in scores on robot strengths. Compared to most of the pairwise differences between different robot strengths, the standard deviation of the error is considerably large in qualification stage data from the FRC championships. Much of the variation may be attributed to the current scoring system in the FRC and/or the degree of model misspecification. A possible way to reduce the variation of the scoring system is to increase the number of matches in the qualification stage for model training. Although robot strengths in the WMPRC model can be generalized to random robot strengths, robot-specific-effects are negligible on the variation of the error in our data analysis. The results are not shown due to space limitations. For the model misspecification problem, another avenue is to borrow the neural network or deep learning framework (cf. Rosenblatt, 1962; LeCun, Bengio and Hinton, 2015) for the relation between the difference in scores and robot strengths with a monotonicity constraint on the link function. Our method should be applicable for parameter estimation in this model approximation. We note that the variation in scores may also be caused by some factors such as robot characteristics, penalty scores and environmental factors. A more thorough study warrants future research.

SUPPLEMENTARY MATERIAL

Supplementary data and codes. The data and R codes used in this article are available at https://github.com/d09948011/Identifying-clusters.

Data analysis and simulation results. The supplementary tables referenced in Remark 3.2.

REFERENCES

Bonelli, H. D. and Reich, B. J. (2008). Simultaneous regression shrinkage, variable selection, and supervised clustering of predictors with OSCAR. Biometrics 64 115–123.

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. 3 1–122.
The numbers of clusters in models estimated by the criterion in (10) as well as BIC, the PCP, MSPE_p and \(\text{MSPE}_Y\) estimates of these estimated models and the numbers of correct predictions of semifinalists, finalists and champions in the playoffs for the 2018 and 2019 FRC championships.

| Season | Tournament | Division | \(\hat{c}\) | PCP | \(\text{MSPE}_p\) | \(\text{MSPE}_Y\) | Semifinalists | Finalists | Champions |
|--------|------------|----------|-------------|------|----------------|----------------|---------------|-----------|-----------|
| 2018   | Houston    | Carver   | 9           | 82.5%| 0.126         | 8896.8         | 3             | 0          | 1         |
|        |            | Galileo  | 10          | 86.0%| 0.115         | 9180.4         | 3             | 2          | 0         |
|        |            | Hopper   | 8           | 84.2%| 0.114         | 9326.7         | 3             | 2          | 0         |
|        |            | Newton   | 7           | 89.7%| 0.083         | 7348.6         | 3             | 2          | 0         |
|        |            | Roehling | 16          | 86.2%| 0.118         | 7901.9         | 2             | 1          | 0         |
|        |            | Turing   | 8           | 89.2%| 0.083         | 9046.2         | 3             | 1          | 0         |
| Houston | Archimedes | 7         | 85.8%       | 0.119| 82.8%         |                | 3             | 0          | 1         |
|        | Carson     | 10       | 76.8%       | 0.139| 9831.4        |                | 4             | 1           | 1         |
|        | Curie      | 6        | 89.3%       | 0.094| 6853.5        |                | 2             | 1           | 1         |
|        | Daly       | 10       | 89.0%       | 0.082| 8195.1        |                | 3             | 2           | 0         |
|        | Darwin     | 8        | 83.5%       | 0.128| 9815.1        |                | 1             | 2           | 0         |
|        | Tesla      | 11       | 86.6%       | 0.116| 8537.8        |                | 2             | 1           | 1         |
| Average |            |          | 10          |      | 85.1%         | 0.111          | 8610.7        | 2.7        | 1.3       | 0.4       |
| 2019   | Houston    | Carver   | 9           | 81.6%| 0.128         | 186.4          | 4             | 1           | 1         |
|        |            | Galileo  | 10          | 89.7%| 0.095         | 124.5          | 4             | 1           | 1         |
|        | Hopper     | 11       | 83.3%       | 0.101| 112.2         |                | 3             | 1           | 1         |
|        | Newton     | 12       | 82.1%       | 0.121| 173.7         |                | 3             | 1           | 1         |
|        | Roehling   | 9        | 77.2%       | 0.145| 164.4         |                | 2             | 1           | 0         |
|        | Turing     | 7        | 82.7%       | 0.115| 153.3         |                | 1             | 1           | 1         |
| Houston | Archimedes | 7         | 78.9%       | 0.130| 227.3         |                | 2             | 1           | 1         |
|        | Carson     | 13       | 80.7%       | 0.122| 111.2         |                | 1             | 1           | 0         |
|        | Curie      | 7        | 82.5%       | 0.118| 175.3         |                | 4             | 1           | 1         |
|        | Daly       | 5        | 74.6%       | 0.152| 189.9         |                | 2             | 1           | 0         |
|        | Darwin     | 5        | 75.0%       | 0.160| 148.1         |                | 3             | 1           | 0         |
|        | Tesla      | 7        | 81.6%       | 0.135| 178.4         |                | 2             | 1           | 1         |
| Average |            |          | 8.5         |      | 80.8%         | 0.128          | 162.1         | 2.4        | 1.1       | 0.6       |

The means of 5000 estimated numbers of clusters, mean squared errors of estimated robot strengths, averages of 5000 MINR’s between the true and estimated clusters, averages of 5000 RC’s between the true and estimated robot strengths and \(\text{MSPE}_D\), \(\text{MSPE}_P\) and \(\text{MSPE}_Y\) of models estimated by the criterion in (10).

| Model | \(\sigma\) | \(\hat{c}\) | MSE | MINR | RC | \(\text{MSPE}_P\) | \(\text{MSPE}_D\) | \(\text{MSPE}_Y\) |
|-------|-------------|-------------|-----|------|----|----------------|----------------|----------------|
| M1    | 2.764       | 4.3         | 3.195| 100.0% | 99.9%| 0.093          | 0.051          | 7.899         |
|       | 5.528       | 9.0         | 120.898 | 98.3% | 96.1%| 0.119          | 0.073          | 38.235        |
|       | 11.056      | 8.3         | 2205.229 | 61.9% | 66.9%| 0.255          | 0.173          | 242.470       |
|       | 22.112      | 7.1         | 5439.416 | 46.0% | 53.8%| 0.363          | 0.245          | 788.623       |
|       | 44.223      | 6.1         | 15676.612 | 34.1% | 42.5%| 0.452          | 0.292          | 2855.011      |
| M2    | 2.569       | 18.0        | 47.067| 94.1% | 95.8%| 0.062          | 0.042          | 9.632         |
|       | 5.137       | 17.0        | 450.986 | 98.3% | 96.1%| 0.119          | 0.073          | 49.023        |
|       | 10.225      | 8.3         | 1789.001 | 61.9% | 66.9%| 0.255          | 0.173          | 242.470       |
|       | 20.549      | 7.1         | 4747.166 | 46.0% | 53.8%| 0.363          | 0.245          | 788.623       |
|       | 41.098      | 6.1         | 13005.827 | 34.1% | 42.5%| 0.452          | 0.292          | 2855.011      |

Bradley, R. A. (1953). Some statistical methods in taste testing and quality evaluation. Biometrics 9 22–38.
Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. the method of paired comparisons. Biometrika 39 324–345.
Cattelan, M., Varin, C. and Firth, D. (2013). Dynamic Bradley-Terry modelling of sports tournaments. J. R. Stat. Soc. C 62 135–150.
Clark, N., Macdonald, B. and Kloo, I. (2020). A Bayesian adjusted plus-minus analysis for the esport Dota 2. J. Quant. Anal. Sports 16 325–341.
Clarke, S. R. and Norman, J. M. (1995). Home ground advantage of individual clubs in English soccer. Statistica 44 509–521.
DeSarbo, W. S. and Cron, W. L. (1988). A maximum likelihood methodology for clusterwise linear regression. J. Classif. 5 249–282.
Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. J. Am. Stat. Assoc. 96 1348–1360.
Fang, E. (2017). The math behind OPR – an introduction. The Blue Alliance Blog. https://blog.thebluealliance.com/2017/10/05/the-math-behind-opr-an-introduction/.
FIRST (2019). 2019 FRC game season manual. https://firstfrc.blob.core.windows.net/frc2019/Manual/2019FRCGameSeasonManual.pdf.
FIRST (2022). 2022 FIRST robotics competition game manual. https://firstfrc.blob.core.windows.net/frc2022/Manual/2022FRCGameManual.pdf.
GARDNER, W. (2015). An overview and analysis of statistics used to rate FIRST robotics teams. https://www.chiefdelphi.com/uploads/default/original/3X/e/b/ebd6f208f128bcc2d44267c7b6d25a4ae53aa8a4.pdf.

HAN, A. K. (1987). Non-parametric analysis of a generalized regression model: the maximum rank correlation estimator. *J. Econom.* **35** 303–316.

HASS, Z. and CRAIG, B. (2018). Exploring the potential of the plus/minus in NCAA women’s volleyball via the recovery of court presence information. *J. Syst. Architet.* **4** 285–295.

HOSMER, D. W. JR, LEMESHOW, S. and STURDIVANT, R. X. (2013). *Applied Logistic Regression*. Wiley, New York.

HVATUM, L. M. (2019). A comprehensive review of plus-minus ratings for evaluating individual players in team sports. *Int. J. Comput. Sci. Sport* **18** 1–23.

LAW, E. (2008). New scouting database from team 2834. https://www.chiefdelphi.com/forums/showpost.php?p=835222&postcount=48.

LICUN, Y., BENGIO, Y. and HINTON, G. (2015). Deep learning. *Nature* **521** 436–444.

LIM, A., CHIANG, C.-T. and TENG, J.-C. (2021). Estimating robot strengths with application to selection of alliance members in FIRST robotics competitions. *Comput. Stat. Data Anal.* **158** 107181.

LIN, C.-C. and NG, S. (2012). Estimation of panel data models with parameter heterogeneity when group membership is unknown. *J. Econ. Methods* **1** 42–55.

MACDONALD, B. (2011). A regression-based adjusted plus-minus statistic for NHL players. *J. Quant. Anal. Sports* **7** 4.

MOSTELLER, F. (1951). Remarks on the method of paired comparisons: I. the least squares solution assuming equal standard deviations and equal correlations. *Psychometrika* **16** 3–9.

ROSENBAUM, D. T. (2004). Measuring how NBA players help their teams win. https://www.82games.com/comm30.htm.

ROSENBLATT, F. (1962). *Principles of Neurodynamics*. Spartan, New York.

SÆBØ, O. D. and HVATUM, L. M. (2015). Evaluating the efficiency of the association football transfer market using regression based player ratings. *NIK-2015 Conference* **4** 285–295.

SCHUCKERS, M. E., LOCK, D. F., WELLS, C., KNICKERBOCKER, C. and LOCK, R. H. (2011). National Hockey League skater ratings based upon all on-ice events: An adjusted minus/plus probability (AMPP) approach. Unpublished Manuscript. http://myslu.stlawu.edu/~msch/sports/lockschuckersprobplussminus113010.pdf.

SCHWARZ, G. (1978). Estimating the dimension of a model. *Ann. Stat.* **6** 461–464.

SOKAL, R. R. (1958). A statistical method for evaluating systematic relationships. *Univ. Kansas, Sci. Bull.* **38** 1409–1438.

SPATH, H. (1979). Algorithm 39: Clusterwise linear regression. *Computing* **22** 367–373.

STEFANI, R. T. (1977). Football and basketball predictions using least squares. *IEEE Trans. Syst. Man Cyber.* **7** 117–21.

STEFANI, R. T. (1980). Improved least squares football, basketball, and soccer predictions. *IEEE Trans. Syst. Man Cyber.* **10** 116–123.

THURSTONE, L. L. (1927). A law of comparative judgment. *Psychol. Rev.* **34** 273–286.

WEINGART, S. (2006). Offense/defense rankings for 1043 teams. https://www.chiefdelphi.com/forums/showpost.php?p=484220&postcount=19.

WIKIPEDIA (2018). FIRST robotics competition. Retrieved August 30, 2018 from https://en.wikipedia.org/wiki/FIRST_Robotics_Competition.

ZHANG, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Stat.* **38** 894–942.