Predicting object-oriented software maintainability using multivariate adaptive regression splines

Yuming Zhou, Hareton Leung

The Journal of Systems and Software, 2006
Contents

- Introduction
- Background
- Related work
- Multivariate Adaptive Regression Splines (MARS) model
- Model evaluation
- Case study
- Conclusion
- Discussion
Introduction

Motivation

- Prediction to software maintainability is very important
  - Various aid to both developers and managers
- Lack of a technique with high prediction accuracy
  - Power of MARS for building prediction models
    - Adapt the unknown functional form
    - Demonstrated in many applications
  - Little works with MARS in software engineering

Research goal

- Investigation into the applicability of MARS
  - To object-oriented software maintainability prediction
  - With object-oriented software metrics
## Related work

- **About issues for the prediction with OO metrics**
  - OO metrics as predictors of maintenance effort
    - Uses 10 metrics (= above 9 + SIZE1)
    - Chidamber and Kemerer, [TSE’94] \(\text{WMC, DIT, RFC, NOC, and LCOM}\)
    - Li and Henry, [JSS’93] \(\text{MPC, DAC, NOM, SIZE2}\)
    - Uses 10 metrics (= above 9 + SIZE1)

- **Way to measure the maintainability of a S/W system**
  - Jeffery et al., [IST’00] Time required to make changes
  - Emam et al., [JSS’01] Time to understand, develop, and implement modifications
  - Number of changes to lines in the code during a maintenance period
  - Represented with a metric “CHANGE”
MARS model (1/4)

- **Representation**
  - Weighted combination of basis functions
    \[
    \hat{y} = c_0 + \sum_{m=1}^{M} c_m \prod_{k=1}^{K_m} b_{km}(x_{v(k,m)})
    \]
  - **Basis function**
    \[
    b_{km+1}(x) = (t_{km} - x)^+ q = \begin{cases} 
    (t_{km} - x)^q & \text{if } t_{km} > x \\
    0 & \text{otherwise}
    \end{cases}
    \]
    \[
    b_{km}(x) = (x - t_{km})^+ = \begin{cases} 
    (x - t_{km})^q & \text{if } x > t_{km} \\
    0 & \text{otherwise}
    \end{cases}
    \]

- **Variables**:
  - \(x_{v(k,m)}\): Value of metric
  - \(M\): Number of non constant terms
  - \(K_m\): order of interactions, determined by user as 1 or 2

- **Notes**:
  - \(t_{km}\): knot of the splines, one of \(x_{v(k,m)}\)
  - \(q\): manipulate the smoothness of splines
  - \(+\): if the calculated value of this function is negative, it becomes the zero
MARS model (2/4)

- **Construction**
  - ✓ Forward stepwise selection process
    - Starts with the constant basis function
    - Adds the basis function which has a good GCV value
  - ✓ Backward pruning process
    - Eliminates basis functions from the model which has the least GCV

- Generalized Cross-validation Criterion (GCV)

\[
\hat{y} = c_0 + \sum_{m=1}^{M} c_m \prod_{k=1}^{K_m} b_{km}(x_{v(k,m)})
\]

\[
GCV(M) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y})^2 \left/ \left[ 1 - \frac{C(M)}{n} \right]^2 \right.
\]

\[
C(M) = M + cd
\]

- **Legend**
  - \( n \): Number of observations in the data set
  - \( y \): Actual value of CHANGE
  - \( \hat{y} \): Predicted value of CHANGE
  - \( C(M) \): Complexity penalty function
  - \( M \): Number of non-constant terms in the model
  - \( c \): user-defined cost penalty factor
  - \( d \): number of basis functions in the model
**MARS model (3/4)**

- **Construction (cont’d)**

  \[
  y = c_0 + c_1'(a-x_1) + c_2'(r-x_2) + c_3'(s-x_3)
  \]

  \[
  y = c_0 + c_1''(Metric_1-x_1) + c_2''(Metric_2-x_1)
  \]

  \[
  y = c_0 + c_1'''(Metric_1-x_1) + c_2'''(Metric_2-x_1)
  \]

  \[
  \hat{y} = c_0 + \sum_{m=1}^{M} \prod_{k=1}^{K_m} b_{km}(x_{v(k,m)})
  \]

  \[
  GCV(M) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y})^2 \left/ \left[ 1 - \frac{C(M)}{n} \right]^2 \right.
  \]

  - **CHANGE (=y)**
    - Data1 Data2 Data3 Data n-1
    - Metric1 a b r s ...
    - Metric2 ...
    - Metric10 ...
    - CHANGE (=y) ...

  - \[ \square \]: Value of a metric

  - n: Total number of classes in a dataset

  - GCV (M): Generalized Cross Validation

  - n: Number of observations in the data set

  - \( \hat{y} \): Actual value of CHANGE

  - \( y_M \): Predicted value of CHANGE

  - C(M): Complexity penalty function

  - M: Number of non-constant terms in the model

  - c: User-defined cost penalty factor

  - d: Number of basis functions in the model
Example of construction

[ Distribution of data ]

x-axis : Independent variable
y-axis : dependent variable

MARS model (4/4)
Model evaluation (1/2)

- Prediction accuracy
  - Absolute Residual Error (ARE)
    - **Resi = yi − ŷi**
    - **AREi = |Resi|**
  - Magnitude of Relative Error (MRE)
    - **MREi = |yi − ŷi| / yi**
  - Pred(0.25) and Pred(0.30)
    - Identifies if a model is acceptable

<table>
<thead>
<tr>
<th>Type</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>SumARE</td>
<td>Measure total residuals over the data</td>
</tr>
<tr>
<td>MedARE</td>
<td>Measure the central tendency of the ARE distribution</td>
</tr>
<tr>
<td>SDARE</td>
<td>Measure the dispersion of the ARE distribution</td>
</tr>
<tr>
<td>MaxMRE</td>
<td>Measure the maximum relative discrepancy</td>
</tr>
<tr>
<td>MMRE</td>
<td>De facto standard as an accuracy measure</td>
</tr>
</tbody>
</table>

*Note: k : number of observations whose MRE is less than l
n : Total number of observations*
Model evaluation (2/2)

- Cross validation
  - Enables to get the predictive power of a model
  - Uses Leave-one-out (LOO) in this paper
    - Good enough to get an estimate as accurate as possible

- Wilcoxon signed-rank test as a significance test
  - Free from the distribution of data

\[ H_0: \text{ARE}_{MARS} = \text{ARE}_X \quad H_0: \text{MRE}_{MARS} = \text{MRE}_X \]
\[ H_1: \text{ARE}_{MARS} \neq \text{ARE}_X \quad H_1: \text{MRE}_{MARS} \neq \text{MRE}_X \]

\[ Z^+ = 1 + 2.5 + 2.5 = 6, \quad Z^- = 4, \quad Z = \min(Z^+, Z^-) = 4 \]
## Object-Oriented metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHANGE</td>
<td>Number of lines changed in the class</td>
</tr>
<tr>
<td>DIT</td>
<td>Depth of Inheritance Tree</td>
</tr>
<tr>
<td>DAC</td>
<td>Data Abstraction Coupling</td>
</tr>
<tr>
<td>LCOM</td>
<td>Lack of COhesion in Methods</td>
</tr>
<tr>
<td>MPC</td>
<td>Message-Passing Coupling</td>
</tr>
<tr>
<td>NOC</td>
<td>Number Of Children</td>
</tr>
<tr>
<td>NOM</td>
<td>Number Of Methods</td>
</tr>
<tr>
<td>RFC</td>
<td>Response For a Class</td>
</tr>
<tr>
<td>WMC</td>
<td>Weighted Methods / Class</td>
</tr>
<tr>
<td>SIZE1</td>
<td>Lines of code</td>
</tr>
<tr>
<td>SIZE2</td>
<td>Number of properties</td>
</tr>
</tbody>
</table>

- **CHANGE**: Insertion and deletion = 1, change of contents = 2
- **DIT**: Length of the longest path from the root to a class
- **DAC**: Number of abstract data types defined in a class
- **LCOM**: Number of pairs of methods using no attribute in common
- **MPC**: Number of send statements defined in a class
- **NOC**: Number of classes that directly inherit from a class
- **NOM**: Number of methods implemented within a class
- **RFC**: Number of methods executed by a received message
- **WMC**: McCabe’s cyclometric complexity of a class
- **SIZE1**: Number of semicolons in a class
- **SIZE2**: Number of attributes and local methods in a class
Case study (2/7)

- Data set description

  ✓ User Interface Management System (UIMS) data set

<table>
<thead>
<tr>
<th>Metric</th>
<th>Maximum</th>
<th>75%</th>
<th>Median</th>
<th>25%</th>
<th>Minimum</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMC</td>
<td>69</td>
<td>12</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>11.38</td>
<td>15.90</td>
</tr>
<tr>
<td>DIT</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2.15</td>
<td>0.90</td>
</tr>
<tr>
<td>RFC</td>
<td>101</td>
<td>20</td>
<td>17</td>
<td>11</td>
<td>2</td>
<td>23.21</td>
<td>20.19</td>
</tr>
<tr>
<td>NOC</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.95</td>
<td>2.01</td>
</tr>
<tr>
<td>LCOM</td>
<td>31</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>7.49</td>
<td>6.11</td>
</tr>
<tr>
<td>MPC</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4.33</td>
<td>3.41</td>
</tr>
<tr>
<td>DAC</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2.41</td>
<td>4.00</td>
</tr>
<tr>
<td>NOM</td>
<td>40</td>
<td>13</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>11.58</td>
<td>10.21</td>
</tr>
<tr>
<td>SIZE1</td>
<td>439</td>
<td>131</td>
<td>74</td>
<td>27</td>
<td>4</td>
<td>106.44</td>
<td>114.65</td>
</tr>
<tr>
<td>SIZE2</td>
<td>61</td>
<td>16</td>
<td>9</td>
<td>6</td>
<td>1</td>
<td>13.97</td>
<td>13.47</td>
</tr>
<tr>
<td>CHANGE</td>
<td>289</td>
<td>39</td>
<td>18</td>
<td>10</td>
<td>2</td>
<td>46.82</td>
<td>71.89</td>
</tr>
</tbody>
</table>

  ✓ QUality Evaluation System (QUES) data set

<table>
<thead>
<tr>
<th>Metric</th>
<th>Maximum</th>
<th>75%</th>
<th>Median</th>
<th>25%</th>
<th>Minimum</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMC</td>
<td>83</td>
<td>22</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>14.96</td>
<td>17.06</td>
</tr>
<tr>
<td>DIT</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1.92</td>
<td>0.53</td>
</tr>
<tr>
<td>RFC</td>
<td>156</td>
<td>62</td>
<td>40</td>
<td>34</td>
<td>17</td>
<td>54.44</td>
<td>32.62</td>
</tr>
<tr>
<td>NOC</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>LCOM</td>
<td>33</td>
<td>14</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>9.18</td>
<td>7.34</td>
</tr>
<tr>
<td>MPC</td>
<td>42</td>
<td>21</td>
<td>17</td>
<td>12</td>
<td>2</td>
<td>17.75</td>
<td>8.33</td>
</tr>
<tr>
<td>DAC</td>
<td>25</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3.44</td>
<td>3.91</td>
</tr>
<tr>
<td>NOM</td>
<td>57</td>
<td>21</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>13.41</td>
<td>12.00</td>
</tr>
<tr>
<td>SIZE1</td>
<td>1009</td>
<td>333</td>
<td>211</td>
<td>172</td>
<td>115</td>
<td>275.58</td>
<td>171.60</td>
</tr>
<tr>
<td>SIZE2</td>
<td>82</td>
<td>25</td>
<td>10</td>
<td>7</td>
<td>4</td>
<td>18.03</td>
<td>15.21</td>
</tr>
<tr>
<td>CHANGE</td>
<td>217</td>
<td>85</td>
<td>52</td>
<td>35</td>
<td>6</td>
<td>64.23</td>
<td>43.13</td>
</tr>
</tbody>
</table>

[39 classes from the UIMS]

[71 classes from the QUES]
Case study (3/7)

- Data set description (cont’d)
  - Correlation between the metrics of two data sets
    - Pearson’s correlation coefficient
      - Indicate the strength of a relationship between two variables
    - Results of analysis

<table>
<thead>
<tr>
<th></th>
<th>WMC</th>
<th>DIT</th>
<th>RFC</th>
<th>NOC</th>
<th>LCOM</th>
<th>MPC</th>
<th>DAC</th>
<th>NOM</th>
<th>SIZE1</th>
<th>SIZE2</th>
<th>CHANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMC</td>
<td>1</td>
<td>-0.22</td>
<td>0.91*</td>
<td>0.23</td>
<td>0.80*</td>
<td>0.63*</td>
<td>0.44*</td>
<td>0.84*</td>
<td>0.97*</td>
<td>0.77*</td>
<td>0.65*</td>
</tr>
<tr>
<td>DIT</td>
<td>-0.13</td>
<td>1</td>
<td>-0.23</td>
<td>-0.47</td>
<td>-0.19</td>
<td>0.06</td>
<td>-0.43*</td>
<td>-0.36*</td>
<td>-0.19</td>
<td>-0.41*</td>
<td>-0.43*</td>
</tr>
<tr>
<td>RFC</td>
<td>0.74*</td>
<td>0.11</td>
<td>1</td>
<td>0.21</td>
<td>0.79*</td>
<td>0.74*</td>
<td>0.62*</td>
<td>0.93*</td>
<td>0.91*</td>
<td>0.89*</td>
<td>0.64*</td>
</tr>
<tr>
<td>NOC</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>0.13</td>
<td>0.03</td>
<td>0.32*</td>
<td>0.23</td>
<td>0.17</td>
<td>0.27</td>
<td>0.56*</td>
</tr>
<tr>
<td>LCOM</td>
<td>0.57*</td>
<td>0.12</td>
<td>0.82*</td>
<td>NA</td>
<td>1</td>
<td>0.50*</td>
<td>0.36*</td>
<td>0.75*</td>
<td>0.82*</td>
<td>0.68*</td>
<td>0.57*</td>
</tr>
<tr>
<td>MPC</td>
<td>0.14</td>
<td>0.02</td>
<td>0.33*</td>
<td>NA</td>
<td>-0.10</td>
<td>1</td>
<td>0.44*</td>
<td>0.55*</td>
<td>0.67*</td>
<td>0.55*</td>
<td>0.45*</td>
</tr>
<tr>
<td>DAC</td>
<td>0.57*</td>
<td>0.39*</td>
<td>0.64*</td>
<td>NA</td>
<td>0.56*</td>
<td>0.02</td>
<td>1</td>
<td>0.75*</td>
<td>0.52*</td>
<td>0.87*</td>
<td>0.63*</td>
</tr>
<tr>
<td>NOM</td>
<td>0.70*</td>
<td>0.13</td>
<td>0.81*</td>
<td>NA</td>
<td>0.88*</td>
<td>-0.11</td>
<td>0.81*</td>
<td>1</td>
<td>0.87*</td>
<td>0.98*</td>
<td>0.64*</td>
</tr>
<tr>
<td>SIZE1</td>
<td>0.89*</td>
<td>0.01</td>
<td>0.80*</td>
<td>NA</td>
<td>0.54*</td>
<td>0.37*</td>
<td>0.64*</td>
<td>0.69*</td>
<td>1</td>
<td>0.82*</td>
<td>0.63*</td>
</tr>
<tr>
<td>SIZE2</td>
<td>0.69*</td>
<td>0.20</td>
<td>0.81*</td>
<td>NA</td>
<td>0.84*</td>
<td>-0.08</td>
<td>0.89*</td>
<td>0.99*</td>
<td>0.71*</td>
<td>1</td>
<td>0.67*</td>
</tr>
<tr>
<td>CHANGE</td>
<td>0.43*</td>
<td>-0.09</td>
<td>0.39*</td>
<td>NA</td>
<td>0.05</td>
<td>0.46*</td>
<td>0.08</td>
<td>0.14</td>
<td>0.64*</td>
<td>0.15</td>
<td>1</td>
</tr>
</tbody>
</table>

- Meaning of values
  - < 0.1 : trivial
  - 0.1 ~ 0.3 : minor
  - 0.3 ~ 0.5 : moderate
  - 0.5 ~ 0.7 : large
  - 0.7 ~ 0.9 : very large
  - 0.9 ~ 1.0 : almost perfect

\[ r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}} \]


2007-10-10 Software Engineering Lab, KAIST
Case study (4/7)

- Results from UIMS data set

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Importance</th>
<th>–gcv</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMC</td>
<td>100.000</td>
<td>3277.827</td>
</tr>
<tr>
<td>NOC</td>
<td>85.673</td>
<td>3123.369</td>
</tr>
<tr>
<td>DAC</td>
<td>59.394</td>
<td>2902.030</td>
</tr>
</tbody>
</table>

- [ Variable importance ]

- [ Residual boxplots ]

MLR : Multivariate Linear Regression
SVR : Support Vector Regression
RT : Regression Trees
ANN : Artificial Neural Networks
Case study (5/7)

- Results from UIMS data set (cont’d)
  - Results about the prediction accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>MaxMRE</th>
<th>MMRE</th>
<th>Pred(0.25)</th>
<th>Pred(0.30)</th>
<th>SumARE</th>
<th>MedARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARS</td>
<td>14.06</td>
<td>1.86</td>
<td>0.28</td>
<td>0.28</td>
<td>1532.78</td>
<td>9.26</td>
</tr>
<tr>
<td>MLR</td>
<td>18.88</td>
<td>2.70</td>
<td>0.15</td>
<td>0.21</td>
<td>1457.26</td>
<td>13.64</td>
</tr>
<tr>
<td>SVR</td>
<td>9.13</td>
<td>1.68</td>
<td>0.31</td>
<td>0.36</td>
<td>1242.59</td>
<td>11.03</td>
</tr>
<tr>
<td>ANN</td>
<td>19.63</td>
<td>1.95</td>
<td>0.15</td>
<td>0.15</td>
<td>1473.14</td>
<td>10.84</td>
</tr>
<tr>
<td>RT</td>
<td>24.57</td>
<td>4.95</td>
<td>0.10</td>
<td>0.10</td>
<td>1988.96</td>
<td>29.54</td>
</tr>
</tbody>
</table>

- Results of Wilcoxon signed-rank test

<table>
<thead>
<tr>
<th>Modeling method</th>
<th>MLR ARE</th>
<th>MLR MRE</th>
<th>SVR ARE</th>
<th>SVR MRE</th>
<th>ANN ARE</th>
<th>ANN MRE</th>
<th>RT ARE</th>
<th>RT MRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARS</td>
<td>0.907a</td>
<td>2.610a</td>
<td>0.809b</td>
<td>0.777b</td>
<td>0.112b</td>
<td>0.334b</td>
<td>2.344a</td>
<td>2.686a</td>
</tr>
<tr>
<td></td>
<td>(0.364)</td>
<td>(0.009)</td>
<td>(0.418)</td>
<td>(0.437)</td>
<td>(0.911)</td>
<td>(0.739)</td>
<td>(0.019)</td>
<td>(0.007)</td>
</tr>
</tbody>
</table>

\[a \quad T_+ < T_-\]
\[b \quad T_+ > T_-\]

\textsuperscript{x}.xxx \textsuperscript{a} or \textsuperscript{b} : z-value, means null hypothesis will be wrong as this value increase
\(\textsuperscript{x}.xxx\) : p-value, indicates the probability that the meaning of z-value is extremely wrong
Case study (6/7)

- Results from QUES data set

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Importance</th>
<th>gev</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIZE1</td>
<td>100.000</td>
<td>1471.475</td>
</tr>
<tr>
<td>WMC</td>
<td>82.008</td>
<td>1171.803</td>
</tr>
<tr>
<td>DAC</td>
<td>36.684</td>
<td>679.524</td>
</tr>
<tr>
<td>MPC</td>
<td>33.426</td>
<td>658.626</td>
</tr>
</tbody>
</table>

[ Variable importance ]

MLR: Multivariate Linear Regression
SVR: Support Vector Regression
RT: Regression Trees
ANN: Artificial Neural Networks

[ Residual boxplots ]
Case study (7/7)

- Results from QUES data set (cont’d)
  - Results about the prediction accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>MaxMRE</th>
<th>MMRE</th>
<th>Pred(0.25)</th>
<th>Pred(0.30)</th>
<th>SumARE</th>
<th>MedARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARS</td>
<td>1.91</td>
<td>0.32</td>
<td>0.48</td>
<td>0.59</td>
<td>1153.10</td>
<td>13.45</td>
</tr>
<tr>
<td>MLR</td>
<td>2.03</td>
<td>0.42</td>
<td>0.37</td>
<td>0.41</td>
<td>1566.08</td>
<td>16.71</td>
</tr>
<tr>
<td>SVR</td>
<td>2.07</td>
<td>0.43</td>
<td>0.34</td>
<td>0.46</td>
<td>1563.67</td>
<td>15.57</td>
</tr>
<tr>
<td>ANN</td>
<td>3.07</td>
<td>0.59</td>
<td>0.37</td>
<td>0.45</td>
<td>2396.85</td>
<td>18.48</td>
</tr>
<tr>
<td>RT</td>
<td>4.82</td>
<td>0.58</td>
<td>0.41</td>
<td>0.45</td>
<td>1983.88</td>
<td>16.67</td>
</tr>
</tbody>
</table>

- Results of Wilcoxon signed-rank test

<table>
<thead>
<tr>
<th>Modeling method</th>
<th>MLR ARE</th>
<th>MLR MRE</th>
<th>SVR ARE</th>
<th>SVR MRE</th>
<th>ANN ARE</th>
<th>ANN MRE</th>
<th>RT ARE</th>
<th>RT MRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARS</td>
<td>3.217^a</td>
<td>3.134^a</td>
<td>3.226^a</td>
<td>3.155^a</td>
<td>3.037^a</td>
<td>3.158^a</td>
<td>3.134^a</td>
<td>2.854^a</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

^a \( T_+ < T_- \)

x.xxx^a or ^b: z-value, means null hypothesis will be wrong as this value increase
(x.xxx): p-value, indicates the probability that the meaning of z-value is extremely wrong
Conclusion

- Contribution
  - Proposes another prediction model
    - MARS is very suitable for modeling complex relationships
  - Showed that MARS can be useful for maintainability prediction
    - Various empirical results were presented

- Future work
  - Replicates this study across programming languages
    - It allows them to further investigate the capability of MARS
  - Develop more accurate prediction models
    - They will combine MARS with other prediction techniques
Discussion

- **Characteristic**
  - Supports the persuasive power of this approach well
    - By referencing many papers
      - Defines the coverage of this paper obviously
    - By comparing the results with other methods

- **Limitation**
  - Collected data only from systems implemented with Ada
  - Result with UIMS which don’t so much outperformed