COMPUTATIONAL BASED AUTOMATED PIPELINE CORROSION DATA ASSESSMENT

MAZURA MAT DIN

A thesis submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

DECEMBER 2015
To arwah Ayah, Abah, Mak\(^2\), Sanorazman, Adam, Aman.
ACKNOWLEDGMENTS

In the name of Allah, Most Gracious, Most Merciful, I thank Allah s.w.t for granting me perseverance and strength I needed to complete this thesis. In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main supervisor, Associate Professor Dr. Norafida Ithnin, for encouragement, guidance and critics. I am also very thankful to my co-supervisor Associate Professor Dr. Azlan Mohd Zain and for his guidance, advice and motivation. I must also acknowledge Associate Professor Dr. Norhazilan from Faculty of Civil Engineering and Professor Kee Eung Kim at Computer Science Department, Korea Advance Institute of Science and Technology, Daejeon, Korea, for their assistance given during my internship. His comments and suggestions have helped a lot. I am also indebted to Ministry of Higher Education and Universiti Teknologi Malaysia for study leave and funding my study. My sincere appreciation also extends to all my fellow postgraduate and my colleagues for their support and assistance at various occasions. I am grateful to all my family members, especially my parents, the late Mat Din Ahmad, Mohd. Isa Hassan, Robiah Mohd Noor, Nursiah Mohd Zain for their prayers and moral support. To Sanorazman Mohd Isa, Adam Rizqan and Muhammad Razman, thank you for the inspiration to complete my journey.
Corrosion is a complex process influenced by the surrounding environment and operational systems which cannot be interpreted by deterministic approach as in the industry codes and standards. The advancement of structural inspection technologies and tools has produced a huge amount of corrosion data. Unfortunately, available corrosion data are still under-utilized. Complicated assessment code, and manual analysis which is tedious and error prone has overburdened pipeline operators. Moreover, the current practices produce a negative corrosion growth data defying the nature of corrosion progress, and consuming a lot of computational time during the reliability assessment. Therefore, this research proposes a computational based automated pipeline corrosion data assessment that provides complete assessment in terms of statistical and computational. The purpose is to improve the quality of corrosion data as well as performance of reliability simulation. To accomplish this, .Net framework and Hypertext Preprocessor (PHP) language is used for an automated matching procedure. The alleviation of deterministic value in corrosion data is gained by using statistical analysis. The corrosion growth rate prediction and comparison is utilized using an Artificial Neural Network (ANN) and Support Vector Machine (SVM) model. Artificial Chemical Reaction Optimization Algorithm (ACROA), Particle Swarm Optimization (PSO), and Differential Evolution (DE) model is used to improve the reliability simulation based on the matched and predicted corrosion data. A computational based automated pipeline corrosion data assessment is successfully experimented using multiple In-Line Inspection (ILI) data from the same pipeline structure. The corrosion data sampling produced by the automated matching is consistent compared to manual sampling with the advantage of timeliness and elimination of tedious process. The computational corrosion growth prediction manages to reduce uncertainties and negative rate in corrosion data with SVM prediction is superior compared to ANN. The performance value of reliability simulation by ACROA outperformed the PSO and DE models which show an applicability of computational optimization models in pipeline reliability assessment. Contributions from this research are a step forward in the realization of computational structural reliability assessment.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
<td>xvii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td></td>
<td>xxi</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td></td>
<td>xxiii</td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Overview</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1.2 Research Motivation</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>1.3 Problem Background</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>1.4 Problem Statement</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>1.5 Research Objectives</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>1.6 Research Scopes</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>
1.7 Research Significance

1.8 Summary

2 LITERATURE REVIEW

2.1 Introduction

2.2 Reliability-based Corrosion Management Systems
   2.2.1 Periodic Inline Inspection (ILIs)
   2.2.2 Corrosion Defect Assessment (CDA)
   2.2.3 Mitigation of Defects (MoD)

2.3 Challenges and Problems in CDA
   2.3.1 Interpretation of ILI Data
      2.3.1.1 Matching and aligning multiple ILI data
      2.3.1.2 Interpretation of ILI Data
   2.3.2 Modelling of Corrosion Growth
   2.3.3 Modelling of Reliability

2.4 Existing Reliability Based Corrosion Defect Assessment
   2.4.1 ILI Data sampling and analysis
      2.4.1.1 Matching Multiple ILI data
      2.4.1.2 Existing ILI Data analysis
      2.4.1.3 Existing Corrosion Growth Model
   2.4.2 Modelling of Reliability Assessment
      2.4.2.1 Deterministic Model
      2.4.2.2 Statistical Model
      2.4.2.3 Computational Model

2.5 Computational Reliability Assessment Model
   2.5.1 Data Matching
3 RESEARCH METHODOLOGY

3.1 Introduction 62

3.2 Problem Situation and Solution Concept 63

3.3 Research Framework 64

3.4 Data Sampling and Analysis 67

3.4.1 ILI Data Preparation 68

3.4.2 Data Sampling 70

3.4.3 Data Analysis 72

3.4.3.1 Statistical and Probability Analysis 73

3.4.3.2 Probability Distribution of Corrosion 76

3.4.3.3 Correctional Methods 78
### 3.5 Development of Computational Corrosion Growth Prediction Modelling

- 3.5.1 The Development of ANN Model 81
- 3.5.2 The Development of SVM Model 82

### 3.6 Computational Reliability Modelling

- 3.6.1 Development of Dimensionless Limit State Function 85
- 3.6.2 Development of Computational Model 86

### 3.7 Instrumentation and Result Analysis

- 3.7.1 Hardware and Software Requirement 88
- 3.7.2 Testing and Analysis 88
- 3.7.3 Evaluation Metrics 89

### 3.8 Summary 92

### 4 AUTOMATED SYSTEM AND ILI DATA QUANTIFICATION 93

- 4.1 Introduction 93
- 4.2 Overview of the Investigation 95
- 4.3 Data Sampling 96
  - 4.3.1 Design of Matching Algorithm 97
  - 4.3.2 Analysis of Matched Data 102
- 4.4 Data Analysis 103
  - 4.4.1 Statistical Analysis 104
    - 4.4.1.1 Sampling Quality Analysis 104
    - 4.4.1.2 Corrosion Dimension Analysis 107
    - 4.4.1.3 Corrosion Growth Analysis 111
  - 4.4.2 Probability Analysis 114
    - 4.4.2.1 Construction of the Histogram 115
4.4.2.2 Estimation of the Parameter Values 126
4.4.2.3 Verification of the Parameter Values 128
4.4.3 Modified Corrosion Rate 134
4.4.4 Linear Prediction of Future Corrosion Defect Sizes 135

4.5 Summary 137

5 COMPUTATIONAL CORROSION GROWTH PREDICTION MODEL 138

5.1 Introduction 138
5.2 The Proposed Computational Prediction 139
5.3 Artificial Neural Network Modelling (ANN-CGM) 142
  5.3.1 Determination of Input Parameters 144
  5.3.2 Group of Dataset 146
  5.3.3 Optimization of Network Parameters 147
  5.3.4 Results of Testing Dataset 149
5.4 Support Vector Machine Modelling (SVM-CGM) 155
  5.4.1 Determination of Input Parameters 156
  5.4.2 Preprocessing of Dataset 158
  5.4.3 Group of Dataset 159
  5.4.4 Optimization of Network Parameters 159
    5.4.4.1 Kernel Type 159
    5.4.4.2 Determination of Parameter Subset 160
  5.4.5 Prediction on Testing Dataset 162
5.5 Experimental Results and Discussion 168
  5.5.1 Comparative Discussion 168
  5.5.2 Evaluation of Model’s Generalization Performance 174
5.6 Summary

6 COMPUTATIONAL RELIABILITY ASSESSMENT MODEL

6.1 Introduction

6.2 Requirement of the Proposed Computational Models

6.2.1 Statistical Parameter

6.2.1.1 Material Properties

6.2.1.2 Defect Properties

6.2.2 Criterion for Model Evaluation

6.2.2.1 Failure Model

6.2.2.2 Limit State Function (LSF)

6.2.2.3 Calculation of Probability of Failure (POF)

6.2.2.4 Target Reliability

6.2.3 Development of Computational Reliability Assessment Model

6.2.3.1 Development of ACROA Model

6.2.3.2 Development of PSO Model

6.2.3.3 Development of DE Model

6.3 Experimental Results and Analysis

6.3.1 Estimation of Fitness Values

6.3.2 Structural Reliability Analysis

6.3.2.1 Reliability Index Calculation

6.3.2.2 Calculation of POF

6.3.3 Statistical analysis

6.4 Summary
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Summary of related works based on reliability assessment issues and problems</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>Matching methods for multiple ILI data</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>ILI data analysis methods</td>
<td>35</td>
</tr>
<tr>
<td>2.4</td>
<td>Corrosion growth prediction model</td>
<td>37</td>
</tr>
<tr>
<td>2.5</td>
<td>Limit state function model</td>
<td>39</td>
</tr>
<tr>
<td>2.6</td>
<td>Summary of existing reliability model</td>
<td>42</td>
</tr>
<tr>
<td>3.1</td>
<td>Description on phases involved in the research framework</td>
<td>66</td>
</tr>
<tr>
<td>3.2</td>
<td>Summary of recorded pigging data</td>
<td>68</td>
</tr>
<tr>
<td>3.3</td>
<td>Number of recorded defects for each set</td>
<td>68</td>
</tr>
<tr>
<td>3.4</td>
<td>A typical presentation of pigging data</td>
<td>69</td>
</tr>
<tr>
<td>3.5</td>
<td>Parameters used to reduce the variation of corrosion depth</td>
<td>79</td>
</tr>
<tr>
<td>3.6</td>
<td>The software or application tools used in this research</td>
<td>87</td>
</tr>
<tr>
<td>3.7</td>
<td>Testing analysis</td>
<td>89</td>
</tr>
<tr>
<td>4.1</td>
<td>An overview of the experimental procedures for the selection of parameters</td>
<td>96</td>
</tr>
<tr>
<td>4.2</td>
<td>Presentation of pigging data from Pipeline B</td>
<td>99</td>
</tr>
<tr>
<td>4.3</td>
<td>Example of matched data result from Pipeline B for doublet matching</td>
<td>102</td>
</tr>
<tr>
<td>4.4</td>
<td>Automated matching results (Pipeline B)</td>
<td>103</td>
</tr>
<tr>
<td>4.5</td>
<td>Difference in the relative distance for matched data</td>
<td>106</td>
</tr>
<tr>
<td>4.6</td>
<td>Difference in the orientation for matched data</td>
<td>106</td>
</tr>
<tr>
<td>4.7</td>
<td>Analysis of automated matching sampling</td>
<td>107</td>
</tr>
</tbody>
</table>
4.8 Average and standard deviation sample of corrosion depth 108
4.9 Average and standard deviation sample of corrosion length 108
4.10 Correlation coefficient ($R^2$) between corrosion depth and length (Noor, 2006) 110
4.11 Correlation coefficient ($R^2$) between corrosion depth and length using Regression Analysis 111
4.12 Corrosion growth rate for defect depth 112
4.13 Corrosion growth rate for defect length 112
4.14 Example of matched data with difference of relative distance more than 1 meter (Pipeline B) 114
4.15 Frequency table of corrosion depth, $d_{392}$ (%wt) (Pipeline B) Sample A 116
4.16 Frequency table of corrosion depth, $d_{392}$ (%wt) (Pipeline B) Sample A 117
4.17 Frequency table of corrosion depth, $d_{395}$ (%wt) (Pipeline B) Sample A 118
4.18 Frequency table of corrosion depth, $d_{390}$ (%wt) (Pipeline B) Sample B 119
4.19 Frequency table of corrosion depth, $d_{392}$ (%wt) (Pipeline B) Sample B 120
4.20 Frequency table of corrosion depth, $d_{395}$ (%wt) for Sample B 121
4.21 Frequency table of corrosion rate depth, $CR_{d,90-95}$ (Pipeline B) Sample A 122
4.22 Frequency table of corrosion rate depth, $CR_{d,90-95}$ (Pipeline B) Sample B 123
4.23 Frequency table of corrosion rate length, $CR_{l,90-95}$ (Pipeline B) Sample A 124
4.24 Frequency table of corrosion rate length, $CR_{l,90-95}$ (Pipeline B) Sample B 125
4.25 Estimated Weibull parameters for corrosion depth 127
4.26 Estimated Exponential parameters for corrosion length 127
4.27 Estimated Normal parameter for corrosion rate of depth growth 127
4.28 Estimated Normal parameters for corrosion rate of length growth 127
4.29 Goodness of fit test (Anderson-Darling) for various probability distribution functions

4.30 Estimation of chi-square value for corrosion depth, $d_{95}$.

4.31 Parameters used to reduce the variation of corrosion depth taken from verified distribution (Pipeline B – match type1)

4.32 Parameters used to reduce the variation of corrosion depth taken from verified distribution (Pipeline B – match type3)

4.33 Comparison between measured and modified data (Pipeline B-match type1)

4.34 Comparison between measured and modified data (Pipeline B-match type3)

4.35 Comparison between uncorrected and corrected corrosion growth rate distribution parameters ($CR_{B(1)92-95}$)

4.36 Comparison between uncorrected and corrected corrosion growth rate distribution parameters ($CR_{B(3)92-95}$)

4.37 Comparison between measured and modified data (Pipeline B-match type3)

4.38 Comparison between uncorrected and corrected corrosion growth rate distribution parameters ($CR_{B(1)92-95}$)

4.39 Comparison between uncorrected and corrected corrosion growth rate distribution parameters ($CR_{B(3)92-95}$)

5.1 Training parameters and its values

5.2 Comparison of various input parameter and error performance

5.3 The groups of samples

5.4 Comparison of network parameters

5.5 Actual rates vs predicted rates for first samples group (90-92)

5.6 Prediction results for first samples group (90-92)

5.7 Actual rates vs predicted rates for second samples group (92-95)

5.8 Prediction results for second samples group (92-95)

5.9 Actual rates vs estimated rates for third samples group (90-95)

5.10 Prediction results for third samples group (90-95)
5.11 Actual rates vs prediction rates for fourth samples group (96-01) 154
5.12 Prediction results for fourth samples group (96-01) 154
5.13 Training parameters and its values 157
5.14 Comparison of input parameters 157
5.15 Comparison of network parameters for SVM 161
5.16 Actual rates vs predicted rates for first samples group (90-92) 164
5.17 Prediction results for first samples group (90-92) 164
5.18 Actual rates and estimated rates for second samples group (92-95) 165
5.19 Prediction results for second samples group (92-95) 165
5.20 Actual rates and predicted rates for third samples group (90-95) 166
5.21 Prediction results for third samples group (90-95) 166
5.22 Actual rates vs predicted rates for fourth samples group (96-01) 167
5.23 Prediction results for fourth samples group (96-01) 167
5.24 Actual values vs predicted results using ANN and SVM methods 169
5.25 Evaluation on the prediction performance of four different test groups estimated by ANN and SVM 174
6.1 Statistical parameters of Pipeline B 182
6.2 Input variables of corrosion defect 183
6.3 Random value equation for probability distribution 188
6.4 Parameters for computational reliability modelling 205
6.5 Average Best Fitness (ABF) (Cost) with 417 samples 206
6.6 Fitness value for year 90 207
6.7 Fitness value for year 92 207
6.8 Fitness value for year 95 208
6.9 Estimation of reliability index by ACROA 209
6.10 POF (%) calculation for different years of defect 210
6.11 F-Test two sample for variances (Pipeline B: 417 samples) 211
6.12 F-Test two sample for variances (Pipeline B: 917 samples) 212
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Taxonomy on research motivation</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>Structure of literature review</td>
<td>14</td>
</tr>
<tr>
<td>2.2</td>
<td>Internal corrosion in submarine pipelines</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>Corroded pipelines (external corrosion)</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>The irregular <em>length, width and depth</em> of a typical corrosion defect. <em>(Adapted from Cosham <em>et al.</em>, 2007)</em></td>
<td>15</td>
</tr>
<tr>
<td>2.5</td>
<td>Five intrinsic mode of corrosion <em>(Adapted from Freeman, 2002)</em></td>
<td>16</td>
</tr>
<tr>
<td>2.6</td>
<td>Relationship between erroneous data and poor decision making on IRM strategies.</td>
<td>18</td>
</tr>
<tr>
<td>2.7</td>
<td>Example of ILI tools <em>(Pigging Products &amp; Service Association, 2014)</em></td>
<td>19</td>
</tr>
<tr>
<td>2.8</td>
<td>Advances in pipeline inspection tools and pigging data acquisition</td>
<td>19</td>
</tr>
<tr>
<td>2.9</td>
<td>Assessing Intelligent Pig Data <em>(Jones, 2006)</em></td>
<td>20</td>
</tr>
<tr>
<td>2.10</td>
<td>Research area in RB-CMS</td>
<td>28</td>
</tr>
<tr>
<td>2.11</td>
<td>Schematic feature matching.</td>
<td>32</td>
</tr>
<tr>
<td>2.12</td>
<td>Computation at a node</td>
<td>48</td>
</tr>
<tr>
<td>2.13</td>
<td>A Multi-Layer Perceptron</td>
<td>49</td>
</tr>
<tr>
<td>2.14</td>
<td>The Logistic and Hyperbolic Tangent Transfer Functions</td>
<td>50</td>
</tr>
<tr>
<td>2.15</td>
<td>Cross validation for termination</td>
<td>52</td>
</tr>
<tr>
<td>2.16</td>
<td>Mapping from original nonlinear separating region to a linear one <em>(Rocco and Moreno, 2002)</em></td>
<td>54</td>
</tr>
<tr>
<td>2.17</td>
<td>Overall procedure of ACROA</td>
<td>55</td>
</tr>
<tr>
<td>3.1</td>
<td>Mapping of problems and solutions</td>
<td>63</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.2</td>
<td>Flow of research implementation</td>
<td>65</td>
</tr>
<tr>
<td>3.3</td>
<td>Flow of data sampling and analysis</td>
<td>67</td>
</tr>
<tr>
<td>3.4</td>
<td>The flow chart of data sampling process</td>
<td>71</td>
</tr>
<tr>
<td>3.5</td>
<td>Overall statistical analysis for the ILI data</td>
<td>72</td>
</tr>
<tr>
<td>3.6</td>
<td>The flow chart of statistical analysis on matched defects</td>
<td>75</td>
</tr>
<tr>
<td>3.7</td>
<td>The flow chart of probability distribution construction</td>
<td>77</td>
</tr>
<tr>
<td>4.1</td>
<td>Overall matching processes</td>
<td>98</td>
</tr>
<tr>
<td>4.2</td>
<td>Program snippet for matching sizing</td>
<td>100</td>
</tr>
<tr>
<td>4.3</td>
<td>Coding for data match console</td>
<td>101</td>
</tr>
<tr>
<td>4.4</td>
<td>Console that shown a doublet matching</td>
<td>101</td>
</tr>
<tr>
<td>4.5</td>
<td>Defect length plotted against defect depth of year 1990 data</td>
<td>109</td>
</tr>
<tr>
<td>4.6</td>
<td>Defect length plotted against defect depth of year 1992 data</td>
<td>109</td>
</tr>
<tr>
<td>4.7</td>
<td>Defect length plotted against defect depth of year 1995 data</td>
<td>110</td>
</tr>
<tr>
<td>4.8</td>
<td>Difference of relative distance upon corrosion rate for depth growth (Pipeline B)</td>
<td>113</td>
</tr>
<tr>
<td>4.9</td>
<td>PDF versus corrosion defect depth for dB90 (Pipeline B) for Sample A</td>
<td>116</td>
</tr>
<tr>
<td>4.10</td>
<td>PDF versus corrosion defect depth for dB92 (Pipeline B) for Sample A</td>
<td>117</td>
</tr>
<tr>
<td>4.11</td>
<td>PDF versus corrosion defect depth for dB95 (Pipeline B) for Sample A</td>
<td>118</td>
</tr>
<tr>
<td>4.12</td>
<td>PDF versus corrosion defect depth for dB90 (Pipeline B) for Sample B</td>
<td>119</td>
</tr>
<tr>
<td>4.13</td>
<td>PDF versus corrosion defect depth for dB92 (Pipeline B) for Sample B</td>
<td>120</td>
</tr>
<tr>
<td>4.14</td>
<td>PDF versus corrosion defect depth for dB95 (Pipeline B) for Sample B</td>
<td>121</td>
</tr>
<tr>
<td>4.15</td>
<td>PDF versus corrosion rate depth, CR_{90-95} (Pipeline B) sample A</td>
<td>122</td>
</tr>
<tr>
<td>4.16</td>
<td>PDF versus corrosion rate depth, CR_{90-95} (Pipeline B) sample B</td>
<td>123</td>
</tr>
<tr>
<td>4.17</td>
<td>PDF versus corrosion rate length, CR_{190-95} (Pipeline B) for Sample A</td>
<td>124</td>
</tr>
<tr>
<td>4.18</td>
<td>PDF versus corrosion rate length, CR_{190-95} (Pipeline B) for Sample B</td>
<td>125</td>
</tr>
</tbody>
</table>
4.19 Exponential Probability plot for corrosion length, $L_{B90}$, $L_{B92}$, $L_{B95}$ (Sample A and Sample B)
4.20 Weibull Probability plot for corrosion depth, $d_{B90}$, $d_{B92}$, $d_{B95}$ (Sample A and Sample B)
4.21 Normal Probability plot corrosion rate for depth, $CR_{dB90}$, $CR_{dB92}$, $CR_{dB95}$, $CRL_{B90}$, $CRL_{B92}$, and $CRL_{B95}$ (Sample A).
4.22 Comparison of prediction data from 1992 to 1995 using corrected corrosion rate and uncorrected corrosion rate ($Pipeline B$ – match type 1)
4.23 Comparison of prediction data from 1992 to 1995 using corrected corrosion rate and uncorrected corrosion rate ($Pipeline B$ – match type 3)
5.1 Computational corrosion growth model development framework
5.2 MLP algorithm
5.3 Structure of the ANN-CGM
5.4 The SVM algorithm
5.5 Comparison of the pattern of actual values against predicted values for ANN model
5.6 Comparison of the pattern of actual values against predicted values for SVM model
6.1 Computational reliability assessment model framework
6.2 The Flowchart of Artificial Chemical Reaction Optimization Algorithm (ACROA)
6.3 Pseudo-code for ACROA
6.4 Initialization function for ACROA
6.5 Flowchart of the Particle Swarm Optimization (PSO)
6.6 Pseudo-code for PSO
6.7 Flowchart of the Differential Equation (DE)
6.8 Pseudo-code for DE
6.9 Number of iteration for all model
7.1 Problem overview and solutions
LIST OF ABBREVIATIONS

ACROA - Artificial Chemical Reaction Optimization Algorithm
AFV - Average Fitness Value
ANN - Artificial Neural Networks
ANN-CGM - Artificial Neural Network Corrosion Growth Model
AR - Accuracy Rate
BFA - Bacterial Foraging Algorithm
BPANN - Backpropagation Artificial Neural Networks
CDA - Corrosion Defect Assessment
CDF - Cumulative Distribution Function
CGM - Computational Growth Model
CompRAM - Computational Reliability Assessment Model
COV - Correlation coefficients
Cr - Corrosion rate
d - depth
DE - Differential Evolution
ER - Error rate
F - F-measure
GA - Genetic Algorithm
ILI - In-line Inspection
IUR - Improved Unit Range
l - length
LSF - Limit State Function
MAE - Mean Absolute Error
MLP - Multi Layer Perceptron
MoD - Mitigation of Defect
MSE - Mean Squared Error
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF</td>
<td>Probability Distribution Function</td>
</tr>
<tr>
<td>POF</td>
<td>Probability of Failure</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>$R^2$</td>
<td>Correlation coefficient</td>
</tr>
<tr>
<td>RAE</td>
<td>Relative Absolute Error</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>RRSE</td>
<td>Root Relative Squared Error</td>
</tr>
<tr>
<td>SVM-CGM</td>
<td>Support Vector Machine Corrosion Growth Model</td>
</tr>
<tr>
<td>$w$</td>
<td>width</td>
</tr>
<tr>
<td>$wt$</td>
<td>wall thickness</td>
</tr>
</tbody>
</table>
## LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>List of related publications</td>
<td>240</td>
</tr>
<tr>
<td>B</td>
<td>Probabilistic Estimation and Verification Approach</td>
<td>241</td>
</tr>
<tr>
<td>C</td>
<td>Example of Dataset</td>
<td>244</td>
</tr>
<tr>
<td>D</td>
<td>Example of Data Tolerance</td>
<td>245</td>
</tr>
<tr>
<td>E</td>
<td>Example of Cr Calculation</td>
<td>246</td>
</tr>
<tr>
<td>F</td>
<td>Example of Corrosion Rate Data</td>
<td>256</td>
</tr>
<tr>
<td>G</td>
<td>Example of Calculation for Chi Square Test Analysis</td>
<td>257</td>
</tr>
</tbody>
</table>
1.1 Overview

Oil and gas industry utilized pipelines as their main infrastructure to transport their goods. Millions of kilometres of pipelines are laid out across the globe either onshore or offshore cannot escape from deterioration over their lifetime of service. However, the number of accidents has also dramatically increased with the increasing number of operating pipelines (Hopkins, 1995; Paik, et al., 2004; Noor, 2006; Chae et al., 2001; Dawson, 2004; Mohd and Paik, 2013; Mohd et al., 2014). Thus, a Pipeline Integrity Management (PIM) becomes an important research field in pipeline lifetime starting from its design, operation, maintenance and replacement. Pipeline can fail due to many factors including construction errors, material defects, operational errors, control system malfunctions, third parties excavations and corrosion. Data on pipelines accidents and their causes compiled by the U.S Department and Transportation’s Research and Special Program Administration, Office of Pipeline Safety (RSPA/OPS) shows that corrosion either external or internal is the most common cause of pipeline accidents with total percentage of 36.6 percent (Li et al., 2009). Cost incur based on corrosion interpreted via repair, lost and contaminated product, environmental damage and possible human safety and health. Corrosion is a complex process influenced by surrounding environment and operational systems which cannot be interpreted by deterministic approach as in the...
industry codes and standards (Mustaffa, 2011). Hence, the Corrosion Management System (CMS) need to be reviewed with an alternative solution on assessing its condition (Zhang, 2014). The main focus of this study is to identify, apply and judge the suitability of the computational methods in evaluating the pipeline reliability of offshore pipeline subjected to internal corrosion. The analysis involves in every stage of assessment will entirely based on the in-line inspection (ILI) data collected at different time interval of the pipelines.

1.2 Research Motivation

Previous studies show the incapability of the deterministic or industry code methods in dealing with ILI data are infeasible economically and practically. The limitation was mostly hindered by the uncertainties that occur in every stage of CDA. Eventhough exist a standards design and codes to provide guidance on the design, standards, constructions and operations of pipelines, the use of codes need to be customized to suit the operation of different environment and conditions (Alkazraji, 2008). Moreover, previous reliability model develop is based on experimental works using a controlled parameters which not the case of real applications. Therefore, motivation of this research is to study and model the reliability of the pipeline from the inspection data (metal loss or ILI data) including the uncertainties govern by it.

The specific motivation leads to this research is simplified as follows:

1) Identification of internal corrosion as one of the major factors that leads to pipeline failure. This triggers an extensive inspection process that generates a huge number of ILI data (metal loss) that is still under utilized. This fact has been proved by Mustaffa (2011), Yahaya (2000), Noor (2006), and Mohd and Paik (2013). Furthermore, using ILI data from repeated inspection on a single pipeline can determine the corrosion rate of it (Desjardin, 2002).
2) *The complexity and time consuming data analysis process* tends to overburden the operators involved and may result in poor planning and maintenance scheduling. Often the operators focused the research on reliability assessment rather than the preceding data modelling and analysis which tend to affect the overall result of pipeline condition prediction.

3) *Traditional analysis* process provides insufficient information to be use for reliability assessment which leads to inaccurate result due to insignificant variables (Noor, 2006; Mustaffa, 2011).

4) *Pipeline codes and standards*: Confusion on adoption of different codes and standards by different countries for guidance in design, construction and operation of pipelines (Alkazraji, 2008). Most of the early design standards were prepared via experimental and/or numerical works, which might differ for different condition and operating practice. Further, the variables and parameters in the laboratory works are manipulated depending on the needs of studies that not represent a real application. Therefore, discrepancies aspects remain unsolved issues among pipeline operators.

5) *Implementation of new computational reliability methods vs deterministic methods for structure assessment*: The use of reliability based computational methods is not to replace the current assessment (deterministic methods), rather it will provide an alternative benchmarks for IRM process. It is less favourable when knowledge about it is still not well understood among industries.

It is important to notice that the new computational method in CDA is by means of complimentary or alternative rather than replacing the current practice. The proposed model is hope to provide a more variation and solution towards IRM management and pipeline integrity preservation.
1.3 Problem Background

In Reliability Based Corrosion Management Systems (RB-CMS) three main parts related to reliability studies is necessary to complete the CMS cycle namely; inspection process, assessment process, and mitigation process (Zhang, 2014; Desjardin, 2002; Noor, 2006). The inspection process of the oil and gas pipeline related to corrosion will produce defect data which known as in-line inspection (ILI) data. Meanwhile in assessment process, a defect will go through an analysis process or known as Corrosion Defect Assessment (CDA). Result from this process is used for Mitigation of Defect (MoD) by means of coating, inhibitors, or even replacement towards pipeline sustainable and effective inspection, repair, and maintenance scheme (IRM). The execution of RB-CMS sequential process is repeated several times dependings on the results from the engineering process until end of the pipeline lifetime. The challenge is how to build a system capable of processing a data and turn it into knowledge in the context of managing pipeline integrity (Wiegele et al., 2004). The importance of CDA in producing an acceptable result was governed by the uncertainties inherits from the interpretation of the ILI, modelling of the corrosion progress and the simulation of its reliability. Thus, the problems in this study centered its discussion on two major problems.

First, the ILI data are in low quality due to uncertainties and use of simplistic approaches in interpreting the corrosion growth (Mustaffa, 2011; Kariyawasam and Wang, 2012). Due to advancement of pipeline inspection technology, abundance of ILI data was available. Unfortunately, it is still under-utilize and this was agreed by Lecchi (2011), Perich et al. (2003), Kamrunnahar et al. (2005), Clouston and Smith (2004), Clausard (2006), Noor (2006), Det Norske Veritas (1999), B31G (1991), and Chouchaoui and Pick (1994). It has been acknowledged that the current practice of pipeline integrity assessment is lack proper guidelines focusing on issues related to data quantification (sampling and data analysis), as well as the intelligent reliability analysis due to the abovementioned research problems (Zio, 2009; Niu et al., 2010; Kuniewski et al., 2008; Noor, 2006; Mustaffa, 2011). This problem occurred due to:
1) *Uncertainties in ILI data:* Particularly for corrosion inspection, the ILI tools such as Magnetic flux leakage (MFL) has also been considered as source of uncertainties (Maes and Salama, 2008; Zhang, 2014, Kariyawasam and Wang, 2011; Mustaffa, 2011).

2) Based on (Kuniewski *et al.*, 2008; Kamrunnahar *et al.*, 2005), *imprecise corrosion data sampling* was due to the limited resolution of inspection tools, imperfect measurement of defect dimension, pipeline material properties operational load and the rate of corrosion growth result in uncertain description of the pipeline condition. As been suggested by Kuniewski *et al.*, 2008 and Noor, 2006, besides the manual procedure on processing the sample data, the sampling size is not accurately fit a current analysis. For example the manual feature matching process is a time consuming, inconsistent and might be vulnerable to human error. Since the diagnosis and interpretation of the corrosion effects depends solely on the experience and the capability of the engineers and inspection personnel.

3) The *complexity of statistical analysis* often views as a too academic by plant engineers and inspection personnel distance themselves from this kind of method. Although a standard exists for the statistical analysis of laboratory corrosion test data, no such standard exists for the analysis of inspection data relating to corrosion measurement (HSE, 2002; Mohd and Paik, 2013).

Secondly, a reliability assessment for both offshore and land based structures becoming important especially in risk-based inspection and maintenance planning (Lecchi, 2011; Zio, 2009; Faber and Straub, 2001; Nakken and Valrsgaard, 1995). For the assessment of structural condition, much attention is focus on the conventional method or industrial practice being tested by a number of authors (Shu et. al., 2009; Melchers and Jeffrey, 2007). Their results show that these approaches
are too rigid in estimating the current and future states of an existing structure. This was due to factors such as:

1) The *simulation-based statistical analysis tends to be time consuming and requires a high level of expertise* to complete the task. Typically a much higher level of accuracy is required both for predictions of structural safety and for predictions of likely future corrosion (Lecchi, 2011). Thus, a model to speed up the performance of simulation is much needed. With that, computational models for reliability assessment come into the picture.

2) *Uncertainties in modelling*, whereby the current implementation used a predefined safety factor or limit states that might differ from one pipeline from the others thus the modelling did not present the real condition of the assess pipeline (Mustaffa, 2011). Moreover, a deterministic and statistical model is a model-driven method compared to computational which is a data-driven method.

The above discussion is summarized and illustrated in Figure 1.1. The flow of CDA research problem and their causes is outline. The successful implementation of RB-CMS depends on CDA to give an insight of the condition of current operating pipeline. The decision from this would benefit the whole process of IRM and at the same time help the pipeline operator preserving their resources and hinder from catastrophics event.
Domain:
Pipeline Integrity Management (PIM)

Reliability-Based Corrosion Management Systems (RB-CMS)

Focus:
Inspection Process (ILI Data)  
Corrosion Defect Assessment (CDA)  
Mitigation of Defects (MoD)

This research

Issues & Problems:

1. Tedious task, inconsistent sampling
2. Uncertainties in corrosion growth modeling
3. Negative corrosion rate

Multiple ILI data

1) Low quality of ILI data
2) Modelling uncertainties

Data Sampling and Analysis
i. Manual matching process
ii. Unstandardize analysis
iii. Expert verification

Corrosion Growth Prediction Modelling
i. Use of simple linear model and deterministic model
ii. Need predefined rules
iii. The absence of large and consistent ILI data.

Computational Reliability Modelling
i. Using a high computational simulation model.
ii. Deterministic parameters setting
iii. No parameters correlation

Previous work:

Data Sampling and Analysis
i. Manual matching process
ii. Unstandardize analysis
iii. Expert verification

Corrosion Growth Prediction Modelling
i. Use of simple linear model and deterministic model
ii. Need predefined rules
iii. The absence of large and consistent ILI data.

Computational Reliability Modelling
i. Using a high computational simulation model.
ii. Deterministic parameters setting
iii. No parameters correlation

Proposed work
(Computational Based Automated Pipeline Corrosion Data Assessment):

i. Perform an automated matching for data sampling.
ii. Performs a structured ILI data quantification.
iii. Computational corrosion growth prediction modelling.
iv. Computational Reliability Modelling.

Figure 1.1: Taxonomy on research motivation
To compensate the shortcomings of the sampling and matching methods an automated matching procedure and a structured statistical method is used to handle the timeliness and accuracies of the task involved. Instead of relying on experimental data, a large amount of inspection data from real structures will give a better insight and accurate information in corrosion assessment. The source of uncertainty inherent in the in-line inspection data and its significance in the context of corrosion reliability analysis was discussed. Implementation of computational model gives significance result for corrosion prediction as compared to the strategy of deterministic techniques. Therefore, prediction based on computational models supported by the available ILI data for comparison provides alternative measures in pipeline maintenance decision.

1.4 Problem Statement

The absence of inspection data quantification standard and predictive corrosion modelling for maintenance of offshore pipeline may cause some difficulties (Lechhi, 2011; M. Kamrunnahar et al., 2005; Clouston and Smith, 2004; Yahaya, 1999; Clausard, 2006; Perich et al., 2003). In the context of corrosion management, the essence of this approach is to combine important pipeline parameter based on in-line inspection data within a computational reliability assessment model for probability of failure estimation. A key element in this analysis approach is explicit consideration of all significant forms of uncertainty, including the uncertainties inherent in the data obtained from in-line inspection. It is hope that this alternative reliability-based process can provide the basis for an industry-accepted approach and an assessment method to manage pipeline integrity with respect to corrosion.

Thus, the following issues will be considered in order to solve the problem:
1) How to design an automated application for matching a repeated ILI data in a timely manner and consistency?

2) How to measure the statistical relationship among the defect parameter?

3) How to predict the corrosion growth variable before proceeding to its reliability assessment?

4) How to design and model an explicit LSF for reliability based model in order to predict the pipeline probability of failure base on ILI data?

5) How to model the computational method to enhance the reliability computational performance?

1.5 Research Objectives

Providing the above problem statement, the research objectives are:

1) To develop an automated matching system and ILI data quantification analysis to improve the data quality for reliability assessment.

2) To develop a corrosion rate model using computational methods for improving the uncertainties in corrosion rate prediction.

3) To develop computational model for improving the simulation based reliability performance of ILI data.
1.6 Research Scopes

The following scopes and limitations have been made mainly due to lack of data in developing deterioration models in this study:

1) The development of the corrosion related models are totally based on the physical evidence from metal loss volume.

2) The effects of material properties, operational condition, and environmental parameters upon corrosion growth are not considered.

3) The data involved a repeated and random inspection data detailing the volume of metal loss.

4) ANN and SVM are used as non-linear model to predict the corrosion growth.

5) Three types of engineering structures transporting crude oil pipelines is chosen involving three different sample set of metal loss data are used to validate the quality and performance of proposed application and model.

6) An optimization of reliability simulation adopting an ACROA, PSO, and DE are used to enhance the performance of reliability assessment process.

7) The inspection data for internal pipeline inspection provided by various inspection vendors such as Petronas, Exxon Mobile, BP Amoco and Rosen from Year 1990 until Year 2001.
1.7 Research Significance

The significance of this study is two-folds: computational and structural aspects. From computational aspect, the proposed method is intended to improve the precision of pipeline reliability assessment from ILI data with inherent inspections uncertainties. It serves as an automated system for tedious and time consuming task of experimental prediction. Thus minimizing the variants and correcting the negative rates from the ILI data. Furthermore, the computational reliability simulation improved the simulation performance in terms of simulation time as compared to the previous works using Monte Carlo simulation. From structural assessment aspects, the integrity prediction embodies reliability assessment information that provide details insight into the states of the structure such as prediction of corrosion rates (Cosham, 2001; Valor, 2003), deriving an explicit LSF (Mustaffa, 2011), and prediction of the failure probabilities (Noor, 2006; Mustaffa, 2011). In assessing structure integrity, combination of this knowledge provides an option to improve the procedure of the assessment as well as optimizing the large volume of inspection data available. Furthermore, the proposed statistical analysis and computational modelling will allow the pipeline operator to design a proper inspection programs and maintenance. For example, in maintenance planning and decision making, a reliability and integrity assessment contributes to minimize the operating structure cost. List of publication produced by this study is listed in Appendix A.

1.8 Summary

This chapter gives an overview of the research conducted in this study. The explanations include overview of the research area, research motivation, problem background, problem statement, objectives, limitations, and contributions of the study. This thesis is organized into seven chapters. A brief description on the content of each chapter as follows: Chapter 1 defines the challenges, problems, objectives, scopes and significance of the study. Chapter 2 reviews the main subjects of interest,
which are automated matching system and ILI data quantification, computational based model for corrosion rate prediction, rigidity of current code practices, limit states functions concepts, and reliability assessment model. Chapter 3 presents the design of the computational reliability assessment model that support the objectives of the study; this includes data sources instrumentations and analyses. Chapter 4 details the sampling and analysis of ILI data, and development of that is resilient towards uncertainties parameters. The analysis results is validated using chi square method, regression analysis and comparison against real ILI data obtain from inspection. Chapter 5 describes the prediction of corrosion growth variables for selected pipeline that addresses the problem of negative corrosion growth as well as uncertainties inherent in inspection data. The ANN-CGM and SVM-CGM is used to model the corrosion growth rate and a performance comparison is made. Chapter 6 simulates a reliability of pipeline conditions represented by computational optimization methods ACROA, PSO and DE to overcome simulation performance problem face by the current method. Chapter 7 draws a general conclusion of the accomplished results and presents the findings of the study as well as recommendations for future study.


Anghel, C. I. Risk assessment for pipelines with active defects based on artificial


Chae, M. J. (2001). Automated Interpretation and Assessment of Sewer Pipeline Infrastructure. Doctor Philosophy, Purdue University.


Sadowski, L. (2013). Non-destructive investigation of corrosion current density in
steel reinforced concrete by artificial neural network. Archive of Civil and Mechanical, 13, 104-121.


Tran, H. D. (2007). Investigation of Deterioration Models for Stormwater Pipe...
Systems. Doctor Philosophy, Victoria University.


