IMPROVED FLOOD DETECTION AND SUSCEPTIBILITY MODELLING USING REMOTE SENSING AND GEOGRAPHIC INFORMATION SYSTEM

By

MAHYAT SHAFAPOURTEHRANY

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

August 2015
COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia
Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

IMPROVED FLOOD DETECTION AND SUSCEPTIBILITY MODELLING USING REMOTE SENSING AND GEOGRAPHIC INFORMATION SYSTEM

By

MAHYAT SHAFAPOURTEHRANY

August 2015

Chair: Associate Professor Biswajeet Pradhan, PhD
Faculty: Engineering

Natural hazards such as floods, landslides, and land subsidence are destructive events which cause catastrophic damages to both human lives and properties. Accurate and easy to implement prediction models are needed to forecast these hazards and delineate the susceptible areas. Although several methods and techniques have been proposed and examined by researchers to map the flood susceptible areas and to provide flood inventory maps, however optimized approaches for flood susceptibility mapping and modeling could not be encountered in the international literature. In traditional way of flood mapping, multiple field works are generally performed to map and monitor floods which is often time consuming and not economically viable. In the last few decades remote sensing based mapping has become hugely popular among the research fraternity. However, optical remote sensing (RS) data and the available classification schemes may not be appropriate for flood extent mapping. This is mainly attributed due to the severe presence of cloud cover especially during the flood seasons. For that reason, flood modelers and remote sensing scientists has to rely on the use of active remote sensing data such as space-borne radar data for flood area mapping. In this regard, a combination of optical and radar data is highly sought after in flood mapping.

In disaster management flood susceptibility mapping is one of the basic steps. There are various types of methods exist in flood susceptibility mapping e.g. traditional based hydrological methods, statistical, probabilistic and data mining based approaches. Traditional hydrological methods are based on linear assumption and require extensive field work. The most popular statistical methods in flood susceptibility assessment are frequency ratio (FR), weights-of-evidence (WoE), and logistic regression (LR). Similarly, the most commonly used data mining approaches in flood susceptibility assessment are artificial neural networks (ANN), fuzzy logic and many more models. However, each of the above mentioned techniques has certain pros and cons. For example, LR is not able to assess the impact of each class of flood conditioning factor on flood occurrence. On the other hand, FR and WoE are capable of evaluating the correlation between them, but they neglect such correlation among the conditioning
factors themselves. Consequently, ANN method is well known for over-training of the dataset.

This study adopted several approaches to investigate and analyze flood occurrence in terms of detection, modeling and optimization of the flood conditioning factors. The current research is divided into two general aspects. The first aspect mainly explored the use of RS technology to detect the flooded areas in Kuala Terengganu, Malaysia using TerraSAR-X image. A TerraSAR-X satellite image was captured during the flood occurrence and Landsat image was captured before the flood occurrence. Both images were classified using object-based method and flooded locations were extracted by subtracting two classes of water bodies. Subsequently, confusion matrix was used to evaluate the results. The second aspect of the current research is related to the use of geographic information system (GIS) in flood susceptibility mapping. A Decision tree (DT) method was implemented for the first time in flood susceptibility mapping. The efficiency of DT to map the flood prone areas in Kelantan, Malaysia was evaluated using the well-known area under the curve (AUC) validation technique. Validation results showed 87% and 82% for success rate and prediction rate respectively.

In order to improve the prediction accuracy of the individual methods such as FR, LR, WoE, and a data-mining based support vector machine (SVM) model, the current research proposed three novel ensemble methods in GIS environment. The overall theory of the ensemble method includes combining the statistical and data-mining methods by integrating the outputs of multiple classifiers to decrease the generalization error. It started with the model development by ensembling FR and LR methods which was then tested in two study areas: Busan, South Korea and Kelantan, Malaysia. In the case study of Busan, the results of the accuracy assessment showed a success rate of 92.7% and a prediction rate of 82.3%. Similarly, the ensemble result of FR and LR models in the Kelantan achieved 90% and 83% for success rate and prediction rate respectively. Next, the second ensemble method was realized by integrating FR and SVM models and was applied for flood susceptibility mapping in Kelantan, Malaysia. The validation results showed 88.71% and 85.21% for success rate and prediction rate respectively.

Next, a new ensemble method was proposed by utilizing WoE and SVM models to produce flood susceptibility map and was applied in Kuala Terengganu area, Malaysia. Validation results of the WoE-SVM ensemble model showed 96.48% (success rate) and 95.67% (prediction rate) accuracy. Another objective of this research was to implement SVM model individually and to evaluate the performance of all its kernel types in flood susceptibility mapping. The validation results for SVM using different kernel types showed that the highest achieved prediction rate (82.16%) was for SVM-RBF. The last goal of the current research was to perform the optimization of the flood conditioning factors using the SVM model aided with Cohen’s kappa index. The result demonstrated that the most influential factors were altitude and slope for all kernel types. Overall, this thesis proposed several new methodologies for flood area mapping and flood susceptibility assessment. The outcome of the current research may assist researchers and local government agencies in flood mitigation strategies and planning.
Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PENAMBAHBAIKAN PENGESANAN DAN PEMETAAN KECENDERUNGAN BANJIR MENGGUNAkan PENDERIAAN JARAK JAUH DAN TEKNOLOGI SISTEM MAKLUMAT GEOGRAFI

Oleh

MAHYAT SHAFAPOURTEHRANY

Ogos 2015

Pengerusi: Profesor Madya Biswajeet Pradhan, PhD
Fakulti: kejuruteraan


Dalam pengurusan bencana banjir, pemetaan kecenderungan adalah salah satu daripada langkah-langkah asas. Terdapat pelbagai jenis kaedah wujud dalam banjir pemetaan kecenderungan, sebagai contoh; kaedah berasaskan hidrologi tradisional, statistik dan data berasaskan pendekatan perlombongan kebarangkalian. Kaedah hidrologi tradisional adalah berdasarkan kepada andaian linear dan memerlukan kerja lapangan yang luas. Kaedah statistik yang paling popular dalam penilaian kecenderungan banjir adalah berasaskan nisbah frekuensi (FR), berat bukti (WoE), dan regresi logistik (LR). Antara
kaedah lain yang biasa digunakan dalam pendekatan perlombongan data dalam penilaian kecenderungan banjir termasuklah: rangkaian neural tiruan (ANN), logik kabur dan banyak lagi model. Walau bagaimanapun, setiap teknik dinyatakan mempunyai kebaikan dan kekurangan tersendiri. Contohnya, LR tidak dapat menilai kesan setiap kelas untuk faktor pendeterminan banjir semasa berlakunya banjir. Sebaliknya, FR dan WoE mampu menilai korelasi antara mereka, namun mereka mengabaikan hubungan itu antara faktor pendeterminan diri mereka sendiri. Oleh yang demikian, kaedah ANN terkenal dengan lebih - latihan dalam memproses data.


Aspek kedua di dalam kajian ini adalah berkaitan dengan permodelan spatial GIS. Pepohon keputusan (DT) telah digunakan buat kali pertama dalam pemetaan kecenderungan banjir dan ketepatannya telah dinilai. Kelantan, Malaysia telah dipilih sebagai kawasan ujian untuk aplikasi teknik DT. Keputusan pengesahsahian yang diperoleh menunjukkan bahawa Luas Bawah Kawasan (AUC) menggunakan teknik DT adalah 87% untuk kadar kejayaan dan 82% untuk kadar ramalan. Oleh itu, penyelidikan ini telah menambah baik keberkesanan model kecenderungan banjir yang sedia ada dengan menambah kaedah ensembel yang baru dalam persekitaran GIS. Kaedah tersebut bermula dengan pembangunan nisbah kekerapan ensembel (FR) dan regresi logistik (LR) yang telah diuji di dua kawasan kajian di Busan, Korea Selatan dan Kelantan, Malaysia. Di kawasan kajian pertama iaitu Busan, keputusan penilaian ketepatkan menunjukkan kadar kejayaan 92.7% dan kadar ramalan 82.3%. Keputusan yang lebih kurang sama telah diperoleh menggunakan kaedah ensembel di kawasan kajian kedua dengan keputusan 90% dan 83% masing-masing untuk kadar kejayaan dan kadar ramalan. Kaedah ensembel kedua telah dicapai dengan menggunakan integrasi FR dan algoritma pembelajaran mesin daripada ‘support vector machine’ (SVM). Bahagian hulu kawasan tadahan lembangan di Kelantan, Malaysia telah dipilih sebagai kawasan kajian untuk menguji kaedah ensembel yang dicadangkan. Keputusan pengesahsahian menunjukkan kadar kejayaan sebanyak 88.71% dan kadar ramalan 85.21% telah diperoleh.

Seterusnya, kaedah ensembel baru telah dicadangkan menggunakan WoE dan SVM model untuk menghasilkan peta banjir kecenderungan dan telah dilaksanakan di kawasan Kuala Terengganu, Malaysia. Keputusan pengesahan model ensembel WoE-SVM menunjukkan 96,48% (kadar berjaya) dan 95,67% (kadar ramalan) ketepatan. Satu lagi objektif kajian ini adalah untuk melaksanakan model SVM secara individu dan untuk menilai prestasi semua jenis kernel dalam pemetaan kecenderungan banjir. Keputusan pengesahan untuk SVM menggunakan kernel yang berbeza menunjukkan bahawa kadar
ramalan yang paling tinggi dicapai (82.16%) adalah SVM-RBF. Matlamat terakhir penyelidikan semasa adalah untuk melaksanakan pengoptimuman faktor pendingin banjir menggunakan model SVM yang dibantu dengan indeks kappa Cohen. Keputusan penyelidikan menunjukkan bahawa faktor yang paling berpengaruh adalah ketinggian dan kecerunan untuk semua jenis kernel. Keseluruhannya, tesis ini mencadangkan beberapa kaedah baru bagi pemetaan kawasan banjir dan penilaian kecenderungan banjir. Hasil daripada kajian semasa dapat membantu penyelidik dan agensi-agensi kerajaan tempatan dalam strategi tebatan banjir dan perancangan.
ACKNOWLEDGEMENTS

Alhamdulillahi robbil 'alamin, praise belongs to Allah, the Lord of the Worlds Who in every century inspires. Nothing I can do without His Will, Guidance, and Permission.

I would like to express the deepest appreciation to my supervisor Assoc. Prof. Dr. Biswajeet Pradhan, who has shown the attitude and the substance of a genius: he continually and persuasively conveyed a spirit of adventure in regard to research and scholarship, and an excitement in regard to teaching. Without his supervision and constant help this dissertation would not have been possible.

I would like to thank my committee members, Prof. Shattri Mansor and Dr. Noordin Bin Ahmad, whose work demonstrated to me that concern for global affairs supported by an “engagement” in comparative literature and modern technology, should always transcend academia and provide a quest for our times. I offer my sincere appreciation for the learning opportunities provided by my committee.

In addition, I would like to thank Mustafa Neamah Jebur. He was always there cheering me up and stood by me through the good times and bad. I greatly value his friendship and I deeply appreciate his belief in me.

Last but not the least, I would like to thank my family: my parents, for giving birth to me at the first place and supporting me spiritually throughout my life. I also thank my sister who encouraged me and prayed for me throughout the time of my research.
I certify that a Thesis Examination Committee has met on 28 August 2015 to conduct the final examination of Mahyat Shafapourtehrany on her thesis entitled “Improved Flood Detection and Susceptibility Modelling Using Remote Sensing and Geographic Information System” in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

**Hussain bin Hamid, PhD**  
Associate Professor  
Faculty of Engineering  
Universiti Putra Malaysia  
(Chairman)

**Abd. Rahman bin Ramli, PhD**  
Associate Professor  
Faculty of Engineering  
Universiti Putra Malaysia  
(Internal Examiner)

**Helmi Zulhaidi bin Mohd Shafri, PhD**  
Associate Professor  
Faculty of Engineering  
Universiti Putra Malaysia  
(Internal Examiner)

**Shiuan Wan, PhD**  
Professor  
Ling Tung University  
China  
(External Examiner)

---

**ZULKARNAIN ZAINAL, PhD**  
Professor and Deputy Dean  
School of Graduate Studies  
Universiti Putra Malaysia  

Date: 22 September 2015
This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

**Biswajeet Pradhan, PhD**
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

**Shattri Mansor, PhD**
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

**Noordin Bin Ahmad, PhD**
Associate Professor
National Space Agency
(Member)

**BUJANG BIN KIM HUAT, PhD**
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:
Declaration by graduate student

I hereby confirm that:
- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ________________________      Date: __________________

Name and Matric No.: _________________________________________
Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: 

______________________________

Name of Chairman of Supervisory Committee:

______________________________

Signature: 

______________________________

Name of Member of Supervisory Committee:

______________________________

Signature: 

______________________________

Name of Member of Supervisory Committee:
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>ABSTRAK</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>APPROVAL</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>DECLARATION</td>
<td>ix</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>xiv</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xv</td>
</tr>
<tr>
<td></td>
<td>LIST OF ABBREVIATIONS</td>
<td>xvii</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>General</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Research Questions</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Problem Statement</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>Motivation Behind the Thesis</td>
<td>5</td>
</tr>
<tr>
<td>1.5</td>
<td>Research Objectives</td>
<td>7</td>
</tr>
<tr>
<td>1.6</td>
<td>Scope of the Study</td>
<td>8</td>
</tr>
<tr>
<td>1.7</td>
<td>Thesis Organization</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE REVIEW</td>
<td>12</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>Important Definitions in Flood Studies</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>Flood Inventory Map</td>
<td>14</td>
</tr>
<tr>
<td>2.4</td>
<td>Flood Detection</td>
<td>15</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Traditional Methods</td>
<td>16</td>
</tr>
<tr>
<td>2.4.2</td>
<td>RS and GIS Techniques</td>
<td>16</td>
</tr>
<tr>
<td>2.5</td>
<td>Flood Conditioning Factors</td>
<td>23</td>
</tr>
<tr>
<td>2.6</td>
<td>Flood Modeling</td>
<td>25</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Traditional Hydrological and Hydrodynamic</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Methods</td>
<td></td>
</tr>
<tr>
<td>2.6.2</td>
<td>RS and GIS Based Methods</td>
<td>26</td>
</tr>
<tr>
<td>2.6.3</td>
<td>Qualitative Methods</td>
<td>27</td>
</tr>
<tr>
<td>2.6.4</td>
<td>Quantitative Methods</td>
<td>27</td>
</tr>
<tr>
<td>2.6.5</td>
<td>Machine Learning Methods</td>
<td>33</td>
</tr>
<tr>
<td>2.6.6</td>
<td>Hybrid Methods</td>
<td>39</td>
</tr>
<tr>
<td>2.6.7</td>
<td>Ensemble Methods</td>
<td>41</td>
</tr>
<tr>
<td>2.7</td>
<td>Validation</td>
<td>41</td>
</tr>
<tr>
<td>2.7.1</td>
<td>Flood Location Map</td>
<td>41</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Flood Susceptibility Maps</td>
<td>41</td>
</tr>
<tr>
<td>2.8</td>
<td>Summary</td>
<td>42</td>
</tr>
<tr>
<td>3</td>
<td>MATERIALS AND METHODS</td>
<td>44</td>
</tr>
<tr>
<td>3.1</td>
<td>Overall Methodology</td>
<td>44</td>
</tr>
</tbody>
</table>
3.2 Study Areas
   3.2.1 Kuala Terengganu, Malaysia 46
   3.2.2 Kelantan, Malaysia 47
   3.2.3 Busan, Korea 48
3.3 Flood Detection
   3.3.1 Study Area and Data Used 50
   3.3.2 Preprocessing 52
   3.3.3 Rule-Based Classification 53
3.4 Rule-based Flood Modeling Using DT 56
   3.4.1 Study Area and Flood Inventory 56
   3.4.2 Flood Conditioning Factors 57
   3.4.3 Rule-Based DT 63
3.5 Ensemble Flood Modeling 64
   3.5.1 Ensemble FR and LR 64
   3.5.2 Ensemble FR and SVM 74
   3.5.3 Ensemble WoE and SVM 77
3.6 Comparisons of the SVM Kernel Types 86
   3.6.1 Study Area 86
   3.6.2 Flood Inventory Map 86
   3.6.3 Flood Conditioning Factors 87
   3.6.4 SVM Performance 88
3.7 Optimization of Conditioning Factors in Flood Mapping 89
3.8 Validation 90
3.9 Summary 91

4 RESULTS AND DISCUSSION 93
4.1 Flood Location Map 93
   4.1.1 Optimization and Segmentation Results 93
   4.1.2 Classified Maps 96
4.2 Flood Susceptibility Map Derived from DT (Rule-Based) Classification 99
4.3 Flood Susceptibility Map Derived From Ensemble Methods 103
   4.3.1 Flood Susceptibility Map Derived from Ensemble FR and LR 103
   4.3.2 Flood Susceptibility Map Derived from Ensemble FR and SVM 113
   4.3.3 Flood Susceptibility Map Derived from Ensemble WoE and SVM 116
4.4 Flood Susceptibility Maps Derived from Four Kernel Types of SVM 125
4.5 Assessment of Significance of Conditioning Factors 126
4.6 Validation 128
   4.6.1 Validation of Flood Location Map 128
   4.6.2 Validation of Flood Susceptibility Map Using DT 129
   4.6.3 Validation of Flood Susceptibility 130
4.6.4 Validation of Flood Susceptibility Map Using Ensemble FR and SVM

4.6.5 Validation of Flood Susceptibility Map Using Ensemble WoE and SVM

4.6.6 Validation of Flood Susceptibility Maps Derived from Four Kernel Types of SVM

4.7 Summary

5 CONCLUSION AND RECOMMENDATIONS

5.1 Introduction
5.2 Summary
5.3 Limitations
5.4 Recommendation for Future Work

REFERENCES

BIODATA OF STUDENT

LIST OF PUBLICATIONS
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Kernel Types and Their Required Parameters</td>
</tr>
<tr>
<td>3.1</td>
<td>Defined Rules for Terra SAR X</td>
</tr>
<tr>
<td>3.2</td>
<td>Defined Rules for Landsat</td>
</tr>
<tr>
<td>4.1</td>
<td>L25 Orthogonal Array and POI for Segmentation Process</td>
</tr>
<tr>
<td>4.2</td>
<td>Results of FR and LR in the Case of Each Factor</td>
</tr>
<tr>
<td>4.3</td>
<td>Results of Ensemble FR and LR in Each Case</td>
</tr>
<tr>
<td>4.4</td>
<td>Results of Ensemble FR and SVM in the Case of Each Factor</td>
</tr>
<tr>
<td>4.5</td>
<td>The Optimal SVM Parameters for Each Testing Model and Their Accuracies</td>
</tr>
<tr>
<td>4.6</td>
<td>Spatial Relationship between Each Factor and Flooding Extracted by WoE Method</td>
</tr>
<tr>
<td>4.7</td>
<td>The Optimal RBF-SVM, SIG-SVM, LN-SVM and PL-SVM Parameters for Each Testing Model and Their Accuracies</td>
</tr>
<tr>
<td>4.8</td>
<td>Accuracy Achieved from Cohen’s Kappa Index for the Trained SVM Models of Flood Susceptibility without One of the Conditioning Factors</td>
</tr>
<tr>
<td>4.9</td>
<td>Confusion Matric for TerraSAR-X</td>
</tr>
<tr>
<td>4.10</td>
<td>Confusion Matric for Landsat</td>
</tr>
<tr>
<td>4.11</td>
<td>The Results of AUC for Each Derived Model</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Structure of the Thesis</td>
<td>10</td>
</tr>
<tr>
<td>2.1</td>
<td>Venn Diagram to Illustrate the Illustrate the FR</td>
<td>30</td>
</tr>
<tr>
<td>3.1</td>
<td>Overall Methodology Flowchart</td>
<td>45</td>
</tr>
<tr>
<td>3.2</td>
<td>Location of the First Study Area, Kuala Terengganu</td>
<td>47</td>
</tr>
<tr>
<td>3.3</td>
<td>Location of the Second Study Area, Kelantan, Malaysia</td>
<td>48</td>
</tr>
<tr>
<td>3.4</td>
<td>Location of the Third Study Area, Busan, Korea</td>
<td>49</td>
</tr>
<tr>
<td>3.5</td>
<td>Methodology Flowchart</td>
<td>51</td>
</tr>
<tr>
<td>3.6</td>
<td>Study Area: Malaysian States (Left Image) and Landsat Imagery (Right Image)</td>
<td>52</td>
</tr>
<tr>
<td>3.7</td>
<td>Flood Location Map with Hill-Shaded Map Of Kelantan State, Malaysia</td>
<td>57</td>
</tr>
<tr>
<td>3.8</td>
<td>Input Conditioning Factors: (a) Altitude</td>
<td>58</td>
</tr>
<tr>
<td>3.8</td>
<td>Input Conditioning Factors: (b) Slope, and (c) Curvature</td>
<td>59</td>
</tr>
<tr>
<td>3.8</td>
<td>Input Conditioning Factors: (d) SPI, and (e) TWI</td>
<td>60</td>
</tr>
<tr>
<td>3.8</td>
<td>Input Conditioning Factors: (f) River, and (g) Geology</td>
<td>61</td>
</tr>
<tr>
<td>3.8</td>
<td>Input Conditioning Factors: (h) Rainfall, and (i) LULC</td>
<td>62</td>
</tr>
<tr>
<td>3.8</td>
<td>Input Conditioning Factors: (j) Soil Types</td>
<td>63</td>
</tr>
<tr>
<td>3.9</td>
<td>Flood Location Map with Hill-Shaded Map of Busan City, South Korea</td>
<td>65</td>
</tr>
<tr>
<td>3.10</td>
<td>List of All the Conditioning Factors: (a) Curvature, and (b) Altitude</td>
<td>66</td>
</tr>
<tr>
<td>3.10</td>
<td>List of All the Conditioning Factors: (c) Geology, (d) Greenfarm, (e) Rainfall, and (f) River</td>
<td>67</td>
</tr>
<tr>
<td>3.10</td>
<td>List of All the Conditioning Factors: (g) Slope, (h) Soil Drain, (i) Soil Effect, and (j) Soil Texture</td>
<td>68</td>
</tr>
<tr>
<td>3.10</td>
<td>List of All the Conditioning Factors: (k) SPI, (l) Timber Drain, (m) Timber Density, and (n) Timber Diameter</td>
<td>69</td>
</tr>
<tr>
<td>3.10</td>
<td>List of All the Conditioning Factors: (o) Timber Type</td>
<td>70</td>
</tr>
<tr>
<td>3.11</td>
<td>Input Conditioning Factors: (a) DEM, (b) Curvature, (c) Geology, and (d) River</td>
<td>72</td>
</tr>
<tr>
<td>3.11</td>
<td>Input Conditioning Factors: (e) SPI, (f) Rainfall, (g) LULC, and (h) Soil</td>
<td>73</td>
</tr>
<tr>
<td>3.11</td>
<td>Input Conditioning Factors: (i) TWI, and (j) Slope</td>
<td>74</td>
</tr>
<tr>
<td>3.12</td>
<td>Flow Chart Employed in This Study for Ensemble FR and SVM Method</td>
<td>75</td>
</tr>
<tr>
<td>3.13</td>
<td>SPSS Clementine Data Mining Environment</td>
<td>76</td>
</tr>
<tr>
<td>3.14</td>
<td>Conditioning Factors: (a) Slope, and (b) SPI</td>
<td>79</td>
</tr>
<tr>
<td>3.14</td>
<td>Conditioning Factors: (c) TWI, and (d) Altitude</td>
<td>80</td>
</tr>
<tr>
<td>3.14</td>
<td>Conditioning Factors: (e) Curvature, and (f) Distance from River</td>
<td>81</td>
</tr>
<tr>
<td>3.14</td>
<td>Conditioning Factors: (g) Geology, and (h) Rainfall</td>
<td>82</td>
</tr>
<tr>
<td>3.14</td>
<td>Conditioning Factors: (i) LULC, and (j) Soil Type</td>
<td>83</td>
</tr>
<tr>
<td>3.15</td>
<td>Methodological Flow Chart for Ensemble WoE and SVM Method</td>
<td>85</td>
</tr>
<tr>
<td>3.16</td>
<td>Surface Runoff Map</td>
<td>88</td>
</tr>
<tr>
<td>3.17</td>
<td>Methodology Flowchart Used In SVM Analysis</td>
<td>89</td>
</tr>
<tr>
<td>4.1</td>
<td>(a) Segmented TerraSAR-X and (b) Landsat Imagery</td>
<td>95</td>
</tr>
<tr>
<td>4.2</td>
<td>(a) Classified Map from TerraSAR-X</td>
<td>96</td>
</tr>
<tr>
<td>4.2</td>
<td>(b) Classified Map from Landsat</td>
<td>97</td>
</tr>
</tbody>
</table>
4.3 (a) Hill-shaded Map Of The Study Area With Flood Locations 98
4.3 (b) Flooded Areas 99
4.4 DT For Classifying Flood Susceptibility 101
4.5 (a) Flood Probability Map Derived from DT 102
4.5 (b) Flood Susceptibility Maps Derived from DT 103
4.6 The Probability Map Obtained after the LR Analyses 107
4.7 Flood Susceptibility Map Produced from the LR Model 108
4.8 The Probability Map Obtained from: (a) Ensemble FR and LR Analyses and (b) DT 112
4.9 Flood Susceptibility Map Produced from: (a) Ensemble FR and LR Model and (b) DT 112
4.10 ROC: (a) Success Rate (b) Prediction Rate Represents the Most Accurate Results of Each Cross-Validation Model 115
4.11 (a) The Flood Probability Map, and; (b) The Flood Susceptibility Map Produced By an Ensemble FR and SVM Method 116
4.12 Flood Probability Maps Derived from Ensemble, (a) WoE and RBF-SVM, and (b) WoE and SIG-SVM 121
4.12 Flood Probability Maps Derived from Ensemble, (c) WoE and LN-SVM, and (d) WoE and PL-SVM 122
4.13 Flood Susceptibility Maps Derived from Ensemble (a) WoE and RBF-SVM, and (b) WoE and SIG-SVM 123
4.13 Flood Susceptibility Maps Derived from Ensemble (c) WoE and LN-SVM, and (d) WoE and PL-SVM 124
4.14 Flood Probability Maps Derived from (a) SVM-LN, (b) SVM-PL, (c) SVM-RBF, and (d) SVM-SIG 125
4.15 Flood Susceptibility Maps Derived from (a) SVM-LN, (b) SVM-PL, (c) SVM-RBF, and (d) SVM-SIG 126
4.16 Conditioning Factors Importance Graphics Obtained from Ensemble FR and SVM Method 127
4.17 AUC 129
4.18 The AUC 131
4.19 AUC (a) the Ensemble FR and LR Method (b) for DT Method 132
4.20 The AUC Curve for the Ensemble FR and SVM Method 134
4.21 (a) Illustration of Cumulative Frequency Diagram Showing Cumulative Flood Occurrence (%; Y-Axis) Occurring In Flood Probability Index Rank (%; X-Axis); Success Rate, and (B) Prediction Rate 136
4.21 (b) Illustration of Cumulative Frequency Diagram Showing Cumulative Flood Occurrence (%; Y-Axis) Occurring In Flood Probability Index Rank (%; X-Axis); Prediction Rate 137
4.22 AUC: (a) Success Rate and (b) Prediction Rate 138
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHP</td>
<td>Analytical Hierarchy Process</td>
</tr>
<tr>
<td>ANFIS</td>
<td>Adaptive Neuro-Fuzzy Interface System</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>ANN-GA</td>
<td>Genetic Algorithm-Based Artificial Neural Network</td>
</tr>
<tr>
<td>AUC</td>
<td>Area Under Curve</td>
</tr>
<tr>
<td>BSA</td>
<td>Bivariate Statistical Analysis</td>
</tr>
<tr>
<td>DEM</td>
<td>Digital Elevation Model</td>
</tr>
<tr>
<td>DT</td>
<td>Decision Tree</td>
</tr>
<tr>
<td>EBF</td>
<td>Evidential Belief Function</td>
</tr>
<tr>
<td>FIS</td>
<td>Fuzzy Interface System</td>
</tr>
<tr>
<td>FR</td>
<td>Frequency Ratio</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>HH</td>
<td>Horizontal Transmit And Horizontal Receive</td>
</tr>
<tr>
<td>LN</td>
<td>Linear</td>
</tr>
<tr>
<td>LR</td>
<td>Logistic Regression</td>
</tr>
<tr>
<td>Acronym</td>
<td>Term</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>LULC</td>
<td>Landuse/cover</td>
</tr>
<tr>
<td>MSA</td>
<td>Multivariate Statistical Analysis</td>
</tr>
<tr>
<td>NDVI</td>
<td>Normalized Difference Vegetation Index</td>
</tr>
<tr>
<td>PL</td>
<td>Polynomial</td>
</tr>
<tr>
<td>POF</td>
<td>Plateau Objective Function</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial Basis Function</td>
</tr>
<tr>
<td>RS</td>
<td>Remote Sensing</td>
</tr>
<tr>
<td>SAR</td>
<td>Synthetic Aperture Radar</td>
</tr>
<tr>
<td>SIG</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>SPI</td>
<td>Stream Power Index</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>TRI</td>
<td>Topographic Roughness Index</td>
</tr>
<tr>
<td>TWI</td>
<td>Topographic Wetness Index</td>
</tr>
<tr>
<td>WoE</td>
<td>Weights-of-Evidence</td>
</tr>
<tr>
<td>VV</td>
<td>Vertical Transmit and Vertical Receive</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 General

Natural hazards, such as landslide, earthquake, flood and etc., cause huge loss of lives and properties worldwide every year (Tierney et al., 2001). Natural disasters are the main cause of irrecoverable damages worldwide (Vorogushyn et al., 2012). Flood is considered as a severe natural hazard and the coverage of its damages is not measurable (Rozalís et al., 2010). Floods are of mainly three types: flash flood, river flood and coastal flood. They occur at different intervals with varying durations. Flood causes serious damages to the transportation, cultural heritage, environmental ecosystem, economy, and humans' lives, etc. (Yu et al., 2013). Kron (2005) describes flooding as a result of heavy precipitation and snow melting that makes the rivers overflow from their normal border and temporarily covers the land which was not used to be covered by water. This type of flooding is classified as river flood. While there are two other types of flash flood and coastal flood exist, but river flood can be predicated through proper methods.

Coastal floods have been defined as floods that happen beside the coasts. This type of floods is triggered by wind storms such as cyclones and low atmospheric pressure that finally result to the set-up of water levels on the coast (Jonkman, 2005). In the case that this set-up of water levels coincides with astronomical high tide at the coast, coastal floods can lead to high water levels and thus flooding of the coastal area. Flash floods are defined by their rapid speed occurrence mostly after a heavy and high intensity localized precipitation. This in turn leads to an unexpected and fast increase of water levels producing a threat to lives and belongings of the citizens. Other triggering factors for flash flooding contain steep slopes, impervious ground surfaces and soils with low permeability (Jonkman, 2005).

Many studies have been done in order to measure and classify the flood impacts from various perspectives. Generally, damages can be direct and indirect, or tangible and intangible which all should be considered in flood damage assessment (Merz et al., 2004; Smith & Ward, 1998). As well as the huge economic cost, floods can bring pathogens into urban environments and cause lingering damp and microbial development in buildings and infrastructure (Dawod et al., 2012; Taylor et al., 2011). Opolot (2013) stated that between 2000 and 2008 almost 99 million people per year were affected by flood alone worldwide. For instance, high frequency of the flood occurrence in Malaysia made this disaster as the most important natural hazard causing many deaths, loss of properties and damages to the ecosystem (Pradhan & Youssef, 2011). Since the 1920s many reports have been recorded about the flood occurrences in Malaysia. Department of Irrigation and Drainage (DID) stated that 9% of land area (29,800 km²) in Malaysia is susceptible to flood and also 22% of the population (4.82
The frequent increase of flood events are mainly due to rapid urbanization and civilization along the rivers, and also cutting the forests (Bronstert, 2003; Christensen & Christensen, 2003). Human activities such as interference in natural cycle by landuse/cover (LULC) changes, unplanned urban expansion near to the bank of the rivers, and uncontrolled construction of buildings can influence the spatial and temporal pattern of flood hazards. Therefore, an assessment of the basin structure, climate condition, and susceptible areas, may assist to prevent the damages which threat the human lives and properties. Because of the tremendous and irreversible potential damages to agriculture, transportation, bridges, and many other aspects of urban infrastructure, flood control and prevention measures are urgently required (Billa et al., 2006). Susceptible areas to the flood should be detected in order to avoid more development in these areas and also to be able to have fast emergency response in various circumstances.

Usually, flood management can be done through few stages: prediction, prevention and damage assessment (Konadu & Fosu, 2009). GIS is known as a powerful set of tools that facilitates gathering, storing, retrieval, analysis and exhibition of spatial information (Opolot, 2013). RS on the other hand is commonly described as the technique of obtaining information about the earth's surface without really being in physical interaction with it. The transmission of information is done using electromagnetic radiation with the help of sensors. RS has been reported to have played a part in the development of GIS, both as a source of technology and as a source of data. The efficiency of remote sensing (RS) and geographic information system (GIS) made the revolution in hydrology and specially flood management which could fulfill all the requirements at each stage. Different types of analyses can be done prior to the flood occurrence, during and after its event. Traditional flood models are increasingly improved or replaced by rule-based and automated methods which are more robust in hazard analyses (Hostache et al., 2013).

Information on predicted future flood locations is essential for the government, the public and emergency department in order to ease early arrangements and planning well in advance beforehand the actual flooding. Through susceptibility analysis the areas which have high potential to the flooding can be recognized and therefore; early warning and emergency response can be performed in order to facilitate early preparations and decrease the effects of this disaster (Kia et al., 2012). The basis of susceptibility mapping is to detect the flood locations and producing flood inventory map with high precision. Although it is not possible to prevent flooding, it can be predicted and controlled through proper analysis and forecasting methods (Cloke & Pappenberger, 2009). In natural hazard management especially in flood management time is one of the most important factors i.e. the employed model should be accurate in order to assist the early warning and prevention measures. Therefore, this study aims to perform flood extent and susceptibility mapping using optimized techniques. Moreover, optimization of flood conditioning factors will be done in order to recognize the significant factors in flood studies.
1.2 Research Questions

In order to fulfill the research objectives of flood detection and susceptibility mapping, the following research questions are addressed in this dissertation:

- Is it possible to extract flood inundated regions from satellite images?

- Is it applicable to enhance the flood susceptibility mapping using ensemble methods?

- How can the concepts of bivariate and multivariate statistical analysis be linked to each other to map the flood susceptible areas?

- What are the optimum conditioning factors that contribute in flood occurrences in each study area?

- Which flood conditioning factors are most relevant to the mapping of flood-prone areas? What weights should be given to each factor?

- What types of assessment can be used to select the input of flood conditioning factors for flood susceptibility mapping?

- Is there a good agreement between the results of different approaches for flood susceptibility mapping and actual flooded areas in the model validation process?

- Which Data-driven GIS modelling technique is the most suitable for delineation of flood prone areas?

- How can the quality and reliability of temporal and spatial probability models be determined, and how can their prediction capability and performance be measured?
1.3 Problem Statement

1- In recent years, due to the growth of urbanization and land sprawl, there has been a significant increase in flood occurrences. Lack of the inventory maps, absence of proper flood analyses, and interpretational difficulties are the main limitations in flood studies and subsequently urban planning (Chau et al., 2005).

2- Flood detection is an initial step for flood susceptibility mapping which should be rapid and accurate as much as possible. However, due to the presence of speckle noise in synthetic aperture radar (SAR) imageries (Pradhan et al., 2014), specular reflectance from other objects (Schlaffer et al., 2015), and spatial heterogeneity of urban areas, classification methods developed for optical images are often not adaptable for flood recognition and mapping (Martinez & Le Toan, 2007). Visual interpretation is another method for flood detection which is based on expert’s knowledge and can be biased (Chambenoit et al., 2003). Threshold segmentation algorithm which is another method for flood extent mapping is very sensitive to low contrast images and it is also based on expert’s opinion. The generated segments should be defined separately and individually for each imagery which make this method not optimized for flood extent extraction (Pulvirenti et al., 2011).

3- Susceptibility maps are the basis of further researches such as hazard and risk analysis (Pradhan, 2010). Governments spend huge budget to prevent flooding, but the lack of accurate flood forecasting and mapping still remains. Based on the literature, most of the existing methods for flood analysis have few drawbacks which should be overcome (Liu & De Smedt, 2005). Traditional hydrological methods such as WetSpa and SWAT require various internal parameters and in order to define those parameters calibration and sensitivity analysis should be done (Lin et al., 2015; Liu & De Smedt, 2005). They require field survey which is significantly time consuming and is not proper for real-time studies (Pavelsky et al., 2014). Bivariate statistical methods such as FR neglects the impact of whole conditioning factor on flood occurrence (Lee et al., 2012). On the other hand, multivariate statistical analysis methods such as LR assess the influence of conditioning factors on flood occurrence while it neglects the impact of each class on flood (Demir et al., 2015; Kavzoglu et al., 2014). Machine learning technique of artificial neural network (ANN) for instance, is considered as black box due to its complex procedure and its high capacity computer requirement (Kia et al., 2012). In addition, qualitative methods such as analytical hierarchy process (AHP) are based on expert’s knowledge which can be affected by person’s situation (Lawal et al., 2012). Hence, based on the aforementioned gaps in the flood studies, it’s necessary to establish more advanced and precise method in order to overcome the existing drawbacks.

4- Numerous conditioning factors such as altitude, slope, aspect and etc. can be used in generation of flood susceptibility maps. Each factor has specific
influence in analysis. However, some factors may have similar impact or may have no significant impact on the final results. Therefore, optimized conditioning factors should be recognized to reduce the time and budget of data collection and consequently, decrease the computation time for analysis of non-significant factors.

The current research aims to cover all the requirements in order to have efficient flood detection and susceptibility modeling by improving the available methods. It is expected that the achievements of the proposed susceptibility mapping be able to enhance the results of the previous studies.

1.4 Motivation Behind the Thesis

Nowadays, natural hazards are common in today’s life. Increasing amounts of natural catastrophes have proved to the human the vital importance of the natural hazards issues for the safety of the environment, and the populations. Rapid urbanization and climate change are expected to raise the amount of rainstorms and river floods. The floods that occur in tropical countries, especially in Malaysia, emphasize the extreme in climatic variations. That is why, the topic of flood monitoring, mapping, modeling and mitigation are among priority tasks in governments schedule (Kussul et al., 2008). These phenomena occur due to the unexpected variation in state of natural features due to natural forces. In most of the cases human is not capable to control and predict these disasters precisely. Main natural catastrophes such as floods, earthquakes, landslides and land subsidence when they occur, they lead to affect the human lives, belongings, infrastructure, farming and environment. The influence of natural hazards is varying based on its amount and coverage region.

Floods are the most common occurring natural catastrophes that influence human and its adjacent environment. It is more vulnerable to Asia and the Pacific regions which affects social and economic stability of those countries. As stated by Pradhan (2010), approximately 90 percent of the destructions related to natural catastrophes in Malaysia are produced by flood. Furthermore, average annual flood damage is as high as US$100 millions. The attention for providing proper flood management has rose over the last centuries. The recent reasons for recurrent flooding of some regions are mostly due to un-planned urbanization, construction and deforestation. If proper management is not available it leads to tragic, the dams can fail, the highways can be flooded and bridge can be breakdown thus increasing the risk for flood. In spite of all this its again human involvement to control flood disaster by immense use of various technology. The use of technology can facilitate flood prevention actions to detect the flood prone areas and to have an early warning for this catastrophe.

In early days field works were used to map and monitor floods with restriction of time and weather circumstances. Nowadays, by invention of GIS and RS technologies those limitations were overcome and flood studies have improved day by day. Those technologies made revolution in hazard and especially flood researches that led to mitigate that phenomenon. Especially use of GIS and RS technologies has really
brought a revolution in mitigation of flood disaster. With improvement of technology in today’s world, it is easier to predict and mitigate damages due to flooding that was not possible in early days. Although several methods and techniques were proposed and examined to map the flood susceptible areas and to provide flood inventory maps, most of them have considerable drawbacks that need to be solved. On the other hand, there are some methods such as rule-based methods which have not been tested in flood studies.

Flood detection analysis should be rapid (Brakenridge et al., 2003) because floods can subside quickly in an inundated area. Thus, researchers have limited time with which to map all of the locations. Fieldwork and traditional methods are unsuitable for such analysis given on-site challenges and difficulties, as well as the long required duration. Furthermore, traditional hydrological methods, such as gauge and discharge measurements, cannot be used to monitor and map flood locations because of the temporal and spatial heterogeneity of large wetlands (Martinez & Le Toan, 2007). Visual interpretation of satellite images is another technique which is a time-consuming, inaccurate, and costly method. It is based on expert knowledge; therefore, it can be erroneous (Chambenoit et al., 2003).

The threshold segmentation algorithm or histogram thresholding is a simple but widely used and effective method to generate a binary image (Pulvirenti et al., 2011). The effectiveness of thresholding procedures for floodplain recognition with SAR sensors depends on the contrast between flooded and non-flooded regions. Therefore, thresholding is sensitive to low-contrast images. However, this method is limited because it is tailored to each satellite scene; that is, it is usually based on visual interpretation. Moreover, its procedure is manual and time-consuming (Pulvirenti et al., 2011). The extent of flooding in an area can also be mapped by active contour modeling. This method can only be used by a researcher with a priori knowledge of the statistical properties of images. Moreover, the method is hindered by local minima and is inaccurate when the initial selected contour is simple or is far from the object boundary. Synthetic aperture radar (SAR) interferometry as another available method, should produce a coherence map; however, this technique is often difficult to be understood (Jebur et al., 2013b). The generation of a coherence map is also complex and disadvantageous; for instance, it requires ground data and two precisely co-registered SAR images (Brisco et al., 2013).

All of the optical images are unsuitable for flood detection applications (Pradhan et al., 2014; Sanyal & Lu, 2004) because clouds usually cover the sky during a flood event, thereby limiting the observational capability of these optical sensors. However, SAR signals can penetrate vegetation and forest (Karjalainen et al., 2012). These sensors can operate both day and night and can highlight different aspects of a single terrain because of their single- or multi-polarized capability. Therefore, it’s the aim of this research to overcome the weak points of optical data using active TerraSAR-X imagery. Regarding the susceptibility mapping, in some methods such as LR, the impact of classes of each conditioning factor on flood occurrence is not considered (Pradhan, 2010). Other statistical methods such as FR method, consider the relationship between flood occurrence and each conditioning factor separately, while not considering the relationships among all the conditioning factors themselves (Lee et al.,

© COPYRIGHT UPM
This thesis aims to propose optimized techniques to map the flood locations and map the flood susceptible areas using ensemble methods. Combination of some methods might increase and enhance the efficiency of the available techniques. 

The key motivation of this research is to use the generated maps in order to avoid more urbanization in inundated areas and have sustainable environment. To reduce the damage and victims in case of a flood occurrence, it is critical to locate the susceptible areas. To recognize those susceptible regions flood inventory map should be generated as a basis of flood susceptibility mapping. Besides the flood inventory and susceptibility mapping, optimization of conditioning factors is of great interest as well. Governments and planners can utilize the produced results by this study to recognize safe regions for citizens, support first responders in emergencies, and update the urban planning strategies. Such data can decrease the requirement to perform field surveys by agencies such as departments of surveying.

1.5 Research Objectives

The general objective of this research is to improve the available techniques for flood mapping and modeling in order to produce more reliable flood inventory map and susceptibility maps.

More specific objectives are:

1. To produce an easy and accurate RS technique for flood extent mapping using active TerraSAR-X data.

2. To perform rule based decision tree (DT) method for flood modeling.

3. To generate novel ensemble techniques of 1) frequency ratio (FR) and logistic regression (LR), 2) FR and support vector machine (SVM), and 3) weights-of-evidence (WoE) and SVM for flood susceptibility mapping in GIS environment.

4. To determine optimized conditioning factors in flood susceptibility mapping using Cohen’s kappa index method.

5. To compare the impact of four SVM kernel types of linear (LN), polynomial (PL), radial basis function (RBF), and sigmoid (SIG) in SVM performance to map the flood susceptible areas.
1.6 Scope of the Study

Flood hazards are the most common and damaging of all natural catastrophes. Each year, flood disasters cause considerable losses and social troubles worldwide. Flood management plans address all aspects of flood management focusing on prevention, protection, preparedness, including flood forecasts and early warning systems (Plate, 2002). During the pre-disaster stage of the flood management, many studies can be done such as flood detection, flood susceptibility, hazard, vulnerability and risk mapping (Kron, 2005). The main scope of this research is related to the improvement of flood detection and susceptibility mapping methods. As it has been mentioned in problem statement, there are some weak points in existing techniques regarding the flood extent mapping and flood susceptibility mapping (Liu & De Smedt, 2005).

In the scope of flood detection, optical data and most of the available classifications techniques for them are not applicable (Sanyal & Lu, 2004). The main problem is cloud cover and the limitation of optical sensors to penetrate the clouds. Traditional gauge and discharge measurements are based on very simple assumptions and they have linear structure (Pavelsky et al., 2014). However, flood and river structures are very complex and non-linear. Other available methods of visual interpretation and threshold segmentation algorithm are based on expert’s knowledge which can be biased (Chambenoit et al., 2003). Change detection method using interferometric technique is very complex and it requires two precisely co-registered SAR images (Hostache et al., 2013). Hence, this research aims to overcome the available drawbacks and difficulties in flood detection by proposing an optimized technique using TerraSAR-X data. Rule-based classification and Taguchi optimization techniques will be used to support the optimized technique.

Regarding the susceptibility mapping, requirement of field work, hydrological expert, and long analysis time for traditional hydrological methods, makes these techniques not optimized enough for flood studies (Lin et al., 2015). FR and LR are bivariate and multivariate statistical methods respectively. Bivariate statistical methods such as FR and WoE assess the correlation between the classes of each conditioning factor on flood occurrence while they neglect the impact of whole factor on that phenomenon (Lee et al., 2012). On the other hand, LR evaluates the effect of each factor on flooding while it doesn’t assess the influence of each class of conditioning factor on flood occurrence (Kavzoglu et al., 2014).

One of the main scopes of this research is to develop novel ensemble techniques to combine the advantages of FR, WoE and LR in order to eliminate their weak points. Some of the machines learning methods such as DT have not been used for flood susceptibility mapping. Hence, this study attempts to implement and evaluate the efficiency of DT in flood mapping. In addition, SVM as another machine learning method can be enhanced and accelerated by its integration with bivariate statistical methods such as FR and WoE. The current research attempts to cover all the requirements in order to have efficient flood susceptibility modeling. It is expected that the achievements of the proposed susceptibility mapping be able to enhance the results of the previous studies.
Three study areas are used in this research which two of them are located in tropical country of Malaysia (Kuala Terengganu and Kelantan) and the third one is located in non-tropical country of South Korea. The reason is to test the efficiency of the proposed methods in different geographical regions. These two countries have some differences in their spatial dataset which might affect the methods performance.
1.7 Thesis Organization

The structure of this thesis consists of five chapters as follows:

![Diagram of thesis structure]

Figure 1.1. Structure of the Thesis
The first chapter contains the introduction, which produces a brief description of the flooding, its negative impacts, its detection, and its modeling. Furthermore, the motivation behind this study, problem statement, objectives, and scope of the study are discussed in this chapter. The second chapter is related to the literature review which explains about the available flood mapping technique, flood susceptibility methods and conditioning factors. Moreover, the available optical and active sensors which can be used in flood studies will be discussed. The third chapter presents the study area, data used, and models employed in analysis and mapping. For each method, detailed information, equations and methodology flowchart will be presented. The fourth chapter covers the results and discussions of the research. Methods will be compared and the most accurate method will be selected. The fifth chapter concludes this thesis with a summary of the work and suggestions for future research.
REFERENCES


Donati, L., & Turrini, M. (2002). An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: application to an area of the Apennines (Valnerina; Perugia, Italy). *Engineering Geology*, 63(3), 277-289.


water level observations into flood modelling. *Journal of Hydrology*, 495(1), 252-266.


Zare, M., Pourghasemi, H. R., Vafakhah, M., & Pradhan, B. (2013). Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arabian Journal of Geosciences*, 6(8), 2873-2888.


