

**FLOOD RISK ASSESSMENT USING MULTI-SENSOR REMOTE
SENSING, GEOGRAPHIC INFORMATION SYSTEM, 2D
HYDRAULIC AND MACHINE LEARNING BASED MODELS**

By

HOSSEIN MOJADDADI RIZEEI

**A Thesis Submitted in Fulfilment of the degree of
Doctor of Philosophy**



**Faculty of Engineering and Information Technology
University of Technology Sydney (UTS),
New South Wales, Australia**

October 2018

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I, Hossein Mojaddadi Rizeei declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctoral of Philosophy, in the FEIT at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:

Signature removed prior to publication.

Signature:

Hossein Mojaddadi Rizeei

Date: 29/01/2019

DEDICATION

This thesis is dedicated to my Father.

In memory of Mohammad Reza Mojaddadi Rizeei. You left fingerprints of grace on my life. You shan't be forgotten. May God almighty bless you.

ACKNOWLEDGEMENT

Praise belongs to God, the Lord of the world who inspires me everywhere.

There are a number of people without whom this thesis might not have been written, and to whom I am greatly indebted.

I would like to express the deepest appreciation to my supervisor distinguished professor 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 support. Without his supervision and constant help this dissertation would not have been possible.

In addition, I would like to thank my wife Maryam Adel Saharkhiz. She is being always there cheering me up and stood beside me through the good times and bad. I deeply appreciate her belief in me.

Last but not the least, I would like to thank my family, especially my dear mother who never gives up to supporting me spiritually throughout my life.

LIST OF PAPERS/PUBLICATIONS

Published journal articles

1. **Rizeei, H. M.**, Pradhan, B., Saharkhiz, M. A. (2018). Urban object extraction using Dempster Shafer Feature Based Image Analysis from Worldview-3 satellite imagery. *International journal of remote sensing*. <https://doi.org/10.1080/01431161.2018.1524173>
2. **Rizeei, H. M.**, Pradhan, B., Saharkhiz, M. A. (2018). Surface runoff prediction regarding LULC and climate dynamics using coupled LTM, optimized ARIMA, and GIS-based SCS-CN models in tropical region. *Arabian Journal of Geosciences*, 11(3), 53. <https://doi.org/10.1007/s12517-018-3397-6>
3. **Rizeei, H. M.**, Pradhan, B., Saharkhiz, M. A. (2018). An integrated fluvial and flash pluvial model using 2D high-resolution sub-grid and particle swarm optimization-based random forest approaches in GIS. *Complex & Intelligent Systems*, 1-20. <https://doi.org/10.1007/s40747-018-0078-8>
4. **Mojaddadi, H.**, Pradhan, B., Nampak, H., Ahmad, N., & Ghazali, A. H. B. (2017). Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS. *Geomatics, Natural Hazards and Risk*, 1-23. <http://dx.doi.org/10.1080/19475705.2017.1294113>.

Published book chapter

1. **Rizeei, H. M.**, Pradhan, B., Saharkhiz, M. A. (2017). Surface Runoff Estimation and Prediction Regarding LULC and Climate Dynamics Using Coupled LTM, Optimized ARIMA and Distributed-GIS-Based SCS-CN Models at Tropical Region. In *Global Civil Engineering Conference (pp. 1103-1126)*. Springer, Singapore.

Published conference paper

1. **Mojaddadi Rizeei, H.**, Pradhan, B (2016) Quantitative assessment of Random Forest flood probability model with 2D hydraulic High Resolution Sub-grid flood inundation model at Damansara River catchment. *IEEE Workshop and conference on Geoscience and Remote Sensing (IWGRS2016)*. **Awarded the best presented paper.**

All the aforementioned papers have been published during my PhD candidature.

LIST OF ABBREVIATIONS

Symbol	Description
AHP	Analytical Hierarchy Process
LTM	Land Transformation Model
ANN	Artificial Neural Network
HRS	High Resolution Sub-grid
AUC	Area Under Curve
BSA	Bivariate Statistical Analysis
DEM	Digital Elevation Model
DT	Decision Tree
EBF	Evidential Belief Function
FIS	Fuzzy Interface System
FR	Frequency Ratio
GIS	Geographic Information System
LN	Linear
LR	Logistic Regression
LULC	Land use/cover

MSA	Multivariate Statistical Analysis
NDVI	Normalized Difference Vegetation Index
PL	Polynomial
POF	Plateau Objective Function
RBF	Radial Basis Function
RS	Remote Sensing
InSAR/IFSAR	Interferometric Synthetic Aperture Radar
SIG	Sigmoid
SPI	Stream Power Index
SVM	Support Vector Machine
TRI	Topographic Roughness Index
TWI	Topographic Wetness Index
PSO	Particle Swarm Optimization
ARIMA	Autoregressive Integrated Moving Average
SCS	Soil Conservation Service
CN	Curve Number
FBIA	Feature Based Image Analysis

PBIA	Pixel Based Image Analysis
OBIA	Object Based Image Analysis
FPP	Flash Pluvial Flood
FF	Fluvial Flood
RF	Random Forest
DSM	Digital Surface Model
ROC	Receiver Operating Characteristic
RMSE	Root Mean Square Error

TABLE OF CONTENTS

CERTIFICATE OF AUTHORSHIP/ORIGINALITY	ii
DEDICATION	iii
ACKNOWLEDGEMENT	iv
LIST OF PAPERS/PUBLICATIONS	v
LIST OF ABBREVIATIONS	vi
TABLE OF CONTENTS	1
LIST OF TABLES	7
LIST OF FIGURES	8
ABSTRACT	12
INTRODUCTION	15
1.1 General	15
1.2 Problem statement	17
1.2 Motivation behind the thesis	17
1.5 Research objectives	22
1.6 Scope of the Study	23
1.2 Research Questions	24
1.7 Thesis organization.....	26
LITERATURE REVIEW.....	28
2.1 Introduction	28
2.2 Important definitions in flood risk study.....	28
2.3 GIS and remote sensing techniques.....	31
2.3.1 Optical imagery.....	32
2.3.2 Active imagery.....	33
2.3.3 Interferometry SAR (InSAR).....	34

2.4 Change detection and prediction methods.....	35
2.5 LULC extraction methods	40
2.6 Flood contributing factors	43
2.7 Flood modelling approaches	47
2.7.1 Hydrological and hydraulic methods.....	51
2.7.2 GIS-based methods	58
2.7.3 Qualitative methods	59
2.7.3.1 Simple additive weighting	59
2.7.3.2 Elimination and choice expressing reality	60
2.7.3.3 Technique for order preference by similarity to ideal solutions	62
2.7.3.4 Analytical hierarchy process	64
2.7.4 Quantitative methods	65
2.7.4.1 Statistical data-driven methods	65
a) Logistic regression	66
b) Frequency ratio	68
c) Fuzzy logic model	69
2.7.5 Machine learning methods.....	69
a) Artificial neural network	70
b) Decision tree	71
c) Support vector machine.....	76
2.7.6 Ensemble methods	80
2.8 Flood inventory	82
2.9 Validation	85
2.9.1 LULC accuracy assessment	85

2.9.2 Flood probability accuracy assessment.....	86
2.10 Summary	87
MATERIALS AND METHODOLOGY	91
3.1 Overall Methodology	91
3.2 Study Areas	92
3.2.1 Damansara, Malaysia.....	93
3.2.2 Semenyih, Malaysia.....	95
3.3 Datasets	97
3.4 Monitoring and simulating surface runoff (First objective).....	99
3.4.1 Satellite image preprocessing	101
3.4.2 Satellite image processing	103
3.4.3 LULC modeling by LTM method	104
3.4.4 Precipitation forecasting by ARIMA model.....	106
3.4.4.1 Statistical Taguchi technique	108
3.4.5 Runoff simulation by SCS-CN model	110
3.4.5.1 Curve number calculation	111
3.5 LULC extraction for physical flood vulnerability analysis (Second objective)...	113
3.5.1 Training and testing sites	114
3.5.2 Satellite image pre-processing.....	115
3.5.3 Classification methods.....	117
3.5.3.1 Pixel based image analysis.....	117
3.5.3.2 Object based image analysis	118
a) Image Segmentation	119
b) Object Feature Extraction	121

3.5.3.3 Feature based image analysis	123
a) Dempster-Shafer theory	123
b) Feature-based fusion of classified maps	126
c) Statistical evaluation of classified land use analysis	126
3.5.4 Vulnerability indices for element at risk	127
3.6 Probability and hazard assessment for two types of flooding (Third objective) ..	128
3.6.1 Preprocessing of geo-statistical GIS-based approach	130
3.6.2 Flood conditioning factors	131
3.6.2.1 Surface elevation	133
3.6.2.2 Surface slope	133
3.6.2.3 Curvature	133
3.6.2.4 Hydrological indices	134
3.6.2.5 Surface runoff	136
3.6.2.6 LULC	136
3.6.2.7 Rainfall intensity	137
3.6.2.8 Distance from river	137
3.6.2.9 Lithological and soil type	138
3.6.3 Inventory of historical pluvial flood events	145
3.6.4 PFF probability assessment using GIS and physical-based model	147
3.6.4.1 Optimizer method of PSO	148
3.6.4.2 Ensemble of PSO-RF model	149
3.6.5 FF probability assessment using 2D-HRS inundation analysis	150
3.6.5.1 Sub-grid resolution principals	152
3.6.5.2 Flow hydrograph analysis	155

3.6.7 Combined FF and PFF probability analysis	156
3.7 Flood risk and hazard assessment using geospatial models (Forth objective)	159
3.7.1 Inventory of historical fluvial flood events.....	160
3.7.2 Flood Probability models.....	163
3.7.2.1 Multivariate GIS data-driven models.....	164
a) Frequency Ratio Model	164
b) Logistic Regression model	165
3.7.2.2 Machine Learning algorithm.....	167
a) Support Vector Machine model	167
3.7.2.3 Ensemble Models	169
3.7.2.4 Qualitative Multi Criteria Decision Making (MCDM) Method	170
Analytic Hierarchy Process (AHP)	170
3.7.3 Flood hazard and risk evaluation	171
3.7.4 Accuracy assessment, calibration, and validation	174
3.8 Summary	176
RESULTS AND DISCUSSION	180
4.1 Simulating surface runoff.....	180
4.1.1 LULC Change detection and prediction forecasting	180
4.1.2 Precipitation change analysis.....	183
4.1.3 GIS-based SCS-CN model results	187
4.2 LULC map extraction results for physical vulnerability assessment	191
4.2.1 PBIA results.....	192
4.2.2 OBIA results	195
4.2.3 FBIA Results.....	200

4.2.4 Visual examination of thematic maps.....	202
4.2.5 Accuracy assessment results.....	205
4.2.5.1 McNemar evaluation results	206
4.2.5.2 Computational running time	207
4.2.6 Assigning the vulnerable rate to LULC classes.....	209
4.3 Fluvial and pluvial flash flood probabilistic assessment.....	210
4.3.1 Simulated probability for PFF	210
4.3.2 Simulated probability for FF by hydrological method	215
4.3.3 Combined PFF and FF probabilistic result	217
4.3.3.1 Sensitivity analysis.....	219
4.3.3.2 Calibration and validation assessment	220
4.4 Hazard and risk assessment for fluvial flooding	224
4.4.1 Flood probability analysis using the GIS-based models.....	224
4.4.1.1 FR results	224
4.4.1.2 LR results	230
4.4.1.3 SVM results.....	234
4.4.1.4 Ensemble FR-SVM results.....	237
4.4.1.5 AHP results	240
4.4.2 Models validation and verification	244
4.4.2 Flood hazard simulation.....	247
4.4.3 Flood risk simulation	251
4.4.3.1 Field verification	254
4.7 Summary	256
CONCLUSION.....	260

5.1 General	260
5.2 Conclusion.....	261
5.3 Implications and Implementation	268
5.4 Limitations.....	269
5.5 Recommendation for future works	269
REFERENCES.....	271

LIST OF TABLES

Table 3.1. Hydro-geomorphological details for Damansara River catchment.....	94
Table 3.2 Geomorphological details for Semenyih catchment.	97
Table 3.3. The geometrical information for two subset study areas	97
Table 3.4. Streamflow and rainfall stations	99
Table 3.5. Optimized segmentation parameters.....	103
Table 3.6. The LTM driver variables applied for LULC forecasting	105
Table 3.7. CN indices under AMC II Conditions based on TR-55 table	111
Table 3.8. Adjustments to select CN for soil moisture conditions (Ward, and Trimble, 2003).	112
Table 3.9. The geometrical information for two subset study areas	114
Table 3.10. The optimized parameters in segmentations for both subsets.....	120
Table 3.10. The object feature characteristics in detail.....	121

Table 3.11. Step by step PSO-RF technique	149
Table 3.12. Optimal values of Manning's n.....	155
Table 4.1. ARIMA model optimization results.....	185
Table 4.2. Calibrated and Adapted CN look-up table.....	187
Table 4.3. Results of confusion matrix accuracy assessment for both sites	205
Table 4.4. Results of McNemar examination for each paired assessment.....	206
Table 4.5. Measured computational operating time of each image analysis approach..	208
Table 4.6. Calibration and validation results for FF and PFF probabilities with observed data at metrological stations.....	222
Table 4.7. FR spatial correlation between flooded area and all conditioning factors	226
Table 4.8 Spatial relationship between each conditioning factor and flooding extracted by LR method.....	230
Table 4.9. AHP spatial correlation between flooded area and each conditioning factor	240
Table 4.10. Area percentage of the risk map of the Damansara River catchment.....	253

LIST OF FIGURES

Figure 1.1 Thesis conceptual framework.	26
Figure 3.1. Overall methodology flowchart.....	92
Figure 3.2. Location of the Damansara River Catchment.....	94
Figure 3.3. Location of the second study area, Semenyih, Malaysia. (Left map was adopted from National Geographic World Map)	96

Figure 3.4. The first objective computational methodology flowchart.....	100
Figure 3.5. The second objective computational methodology flowchart.....	115
Figure 3.6. The third objective computational methodology flowchart.....	130
Figure 3.7. Flood Contributing factors of Damansara: (a) surface elevation, (b) surface slope, (c) curvature, (d) SPI, (e) TWI, and (f) TRI,	139
Figure 3.7. (Continued) Flood Contributing factors of Damansara: (g) STI, (h) Surface runoff, (i) LULC, (j) rainfall intensity, (k) distance from river, and (l) soil type.	140
Figure 3.7. (Continued) Flood Contributing factors of Damansara: (m) lithological geology. .	141
Figure 3.8. Flood contributing factors of Semenyih: (a) surface elevation and (b) surface slope,	141
Figure 3.8. (Continued) Flood Contributing factors of Semenyih: (c) curvature, (d) SPI, (e) TWI, and (f) TRI,	142
Figure 3.8. (Continued) Flood Contributing factors of Semenyih: (g) STI, (h) Surface runoff, (i) LULC, and (j) Rainfall intensity,	143
Figure 3.8. (Continued) Flood Contributing factors for Semenyih: (k) distance from river, (l) soil type, and (m) lithological geology.	144
Figure 3.9. PFF inventory of historical events.....	146
Figure 3.10. 2D-HRS modelling computational mesh terminology	154
Figure 3.11. The forth objective computational methodology flowchart	160
Figure 3.12. Historical inventory events of fluvial flooding nearby the NKVE highway	162
Figure 3.13. Historical inventory events of fluvial flooding in Semenyih catchment	163
Figure 4.1. Comparison of LULC dynamics in 2000, 2010 and predicted 2020	181
Figure 4.2. LULC maps for 2000, 2010 and 2020	181
Figure 4.3. Trend analysis plot for the Kg. Pasir gauging station.....	184
Figure 4.4. ACF of residuals (with 5% significance limits for autocorrelations)	185
Figure 4.5. Optimized ARIMA time series forecasting results till 2020 (95% confidence limits).	186

Figure 4.6. Distributed-maximum surface runoff in three considered time periods.	189
Figure 4.7. Compared Distributed-maximum surface runoff in three considered time periods within sub-basins.	190
Figure 4.8. LULC extraction results using PBIA classifiers in site A, a) Bayes, b) KNN, and c) SVM.....	193
Figure 4.9. LULC extraction results of PBIA classifiers over Site B; a) Bayes, b) KNN, and c) SVM.....	194
Figure 4.10. Range of spectral indices extracted from different LULC classes by the OBIA-DT approach.....	195
Figure 4.11. Range of.....	196
Figure 4.12. Range of textural values extracted from different LULC classes by the OBIA-DT approach.....	198
Figure 4.13. LULC extraction results using OBIA- DT for subsets (a) A and (b) B.....	199
Figure 4.14. Compared the accuracy of different Masses of Belief using ROC curve	200
Figure 4.15. LULC results from DS-FBIA fusion method for subsets A (a) and B (b).....	201
Figure 4.16. Machine learning results in finer scale for visual comparison: a) Bayes, b) KNN, c) SVM, d) DT, e) fusion DS, and f) WV-3 satellite image.....	203
Figure 4.17. Extracted flood vulnerable index for Damansara basin.....	209
Figure 4.18. Optimized weightage of parameters extracted by PSO-RF model	211
Figure 4.19. PFF probability map using GIS-based PSO-RF model	213
Figure 4.20. PFF inundation depth hazard map using GIS-based PSO-RF model	214
Figure 4.21. a) Maximum FF probability map and b) maximum FF inundation depth map	216
Figure 4.22. Distribution of combined FF and PFF probability classes	218
Figure 4.23. Combined PFF and FF probabilistic map	218
Figure 4.24. Compared simulated FF inundation depths with observed water level depths over three gauged stations.....	221

Figure 4.25. Compared simulated PFF inundation depth with observed precipitation depth over three rainfall stations.....	222
Figure 4.26. ROC accuracy assessment of GIS-based PFF probability map.....	223
Figure 4.27. Flood probability map obtained by the FR model in Damansara.....	228
Figure 4.28. Flood probability map obtained by the FR model in Semenyih.....	229
Figure 4.29. Flood probability map obtained by the LR model in Damansara.....	232
Figure 4.30. Flood probability map obtained by the LR model in Semenyih.....	233
Figure 4.31. The assigned probability weightage for each parameter obtained by SVM	234
Figure 4.32. Flood probability map obtained by the SVM model in Damansara	235
Figure 4.33. Flood probability map obtained by the SVM model in Semenyih	236
Figure 4.34. The assigned weightage for each parameter obtained by the ensemble FR-SVM model	237
Figure 4.35. Flood probability map obtained by the ensemble FR-SVM model in Damansara	238
Figure 4.36. Flood probability map obtained by the ensemble FR-SVM model in Semenyih	239
Figure 4.37. Flood probability map obtained by AHP model in Damansara.....	242
Figure 4.38. Flood probability map obtained by AHP model in Semenyih.....	243
Figure 4.39. The ROC success rate represents the accuracy of applied flood probability models	245
Figure 4.40. The ROC prediction rate represents the accuracy of applied flood probability models.....	246
Figure 4.41. Hazardous triggering index of Damansara catchment.....	249
Figure 4.42. Flood hazard index map of Damansara catchment.....	250
Figure 4.43. Flood risk levels of Damansara	252
Figure 4.44. Field verification for Damansara catchment.....	255

ABSTRACT

Flooding events threaten the population, economy and environment worldwide. In recent years, several spatial methods have been developed to map flood susceptibility, hazard and risk for predicting and modelling flooding events. However, this research proposes multiple state-of-the-art approaches to assess, simulate and forecast flooding from recent satellite imagery.

Firstly, a model was proposed to monitor changes in surface runoff and forecast future surface runoff on the basis of land use/land cover (LULC) and precipitation factors because the effects of precipitation and LULC dynamics have directly affected surface runoff and flooding events. Land transformation model (LTM) was used to detect the LULC changes. Moreover, an autoregressive integrated moving average (ARIMA) model was applied to analyse and forecast rainfall trends. The parameters of the ARIMA time series model were calibrated and fitted statistically to minimise prediction uncertainty through modern Taguchi method. Then, a GIS -based soil conservation service-curve number (SCS-CN) model was developed to simulate the maximum probable surface runoff. Results showed that deforestation and urbanisation have occurred upon a given time and have been predicted to increase. Furthermore, given negative changes in LULC, surface runoff increased and was forecasted to exceed gradually by 2020. In accordance with the implemented model calibration and accuracy assessment, the GIS-based SCS-CN combined with the LTM and ARIMA model is an efficient and accurate approach to detecting, monitoring and forecasting surface runoff.

Secondly, a physical vulnerability assessment of flood was conducted by extracting detailed urban features from Worldview-3. Panchromatic sharpening in conjunction with atmospheric and topographic corrections was initially implemented to increase spatial resolution and reduce atmospheric distortion from satellite images. Dempster–Shafer (DS) fusion classifier was proposed in this part as a feature-based image analysis (FBIA) to extract urban complex objects. The DS-FBIA was investigated among two sites to examine the transferability of the proposed method. In addition, the DS-FBIA was compared with other common image analysis approaches (pixel- and object-based image analyses) to discover its accuracy and computational operating time. k-nearest neighbour, Bayes and support vector machine (SVM) classifiers were tested as pixel-based image analysis approaches, while decision tree classifier was examined as an object-based image analysis approach. The results showed improvements in detailed urban extraction obtained using the proposed FBIA with 92.2% overall accuracy and with high transferability from one site to another.

Thirdly, an integrated model was developed for probability analysis of different types of flood using fully distributed GIS-based algorithms. These methods were applicable, particularly where annual monsoon rains trigger fluvial floods (FF) with pluvial flash flood (PFF) events occur simultaneously. A hydraulic 2D high-resolution sub-grid model of hydrologic engineering centre river analysis system was performed to simulate FF probability and hazard. Moreover, machine learning random forest (RF) method was used to model PFF probability and hazard. The RF was optimised by particle swarm optimisation (PSO) algorithm. Both models were verified and calibrated by cross

validation and sensitivity analysis to create a coupled PFF– FF probability mapping. The results showed high accuracy in generating a coupled PFF–FF probability model that can discover the impact and contribution of each type to urban flood hazard. Furthermore, the results provided detailed flood information for urban managers to equip infrastructures, such as highways, roads and sewage network, actively.

Fourthly, the risk of a flood can be assessed through different stages of flood probability, hazard and vulnerability. A total of 13 flood conditioning parameters were created to construct a geospatial database for flood probability estimation in two study areas. To estimate flood probability, five approaches, namely, logistic regression, frequency ratio (FR), SVM, analytical hierarchy process and combined FR–SVM, were adopted. Then, a flood risk map was generated by integrating flood hazard and vulnerability. The accuracy of flood probability indices indicated that the combined FR–SVM method achieved the highest accuracy among the other approaches. The reliability of the results obtained from this research was also verified in the field. The most effective parameters that would trigger flood occurrence were rainfall and flood inundation depth.

In this research, transferable residency from one study area to another was verified through all the implemented methods. Therefore, the proposed approaches would be effectively and easily replicated in other regions with a similar climate condition, that condition that is, having a sufficient amount of flooding inventory events. Moreover, the results of the proposed approaches provided solid-detailed information that would be used for making favourable decisions to reduce and control future flood risks.