UNIVERSITI PUTRA MALAYSIA

SPATIAL PREDICTION OF LANDSLIDE HAZARDS AND RISK AREAS USING INTEGRATED STATISTICAL AND DATA MINING APPROACHES

OMAR FAISAL S. ALTHUWAYNEE

FK 2014 21

SPATIAL PREDICTION OF LANDSLIDE HAZARDS AND RISK AREAS USING INTEGRATED STATISTICAL AND DATA MINING APPROACHES

By

OMAR FAISAL S. ALTHUWAYNEE

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

September 2014

DEDICATION

To my parents, loving family, and friends, whose genuine love and support are behind my success.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in Fulfillment of the Requirements for the Degree of Doctor of Philosophy

SPATIAL PREDICTION OF LANDSLIDE HAZARDS AND RISK AREAS USING INTEGRATED STATISTICAL AND DATA MINING APPROACHES

By

OMAR FAISAL S. ALTHUWAYNEE

September 2014

Supervisor: Assoc. Prof. Biswajeet Pradhan, PhD
Faculty: Engineering

Landslides are one of the many forms of natural hazards that often cause severe property damages, economic loss, and high maintenance costs. Slope failures are a result of multiple triggering factors, including anthropogenic activities, earthquakes, and intense rainfall, and of reactions of a host of unstable surface materials related to geology, land cover, slope geometry, moisture content, and vegetation. This thesis presents a set of novel GIS-based statistical approaches developed for the hazard mapping of rainfall-induced landslides. These approaches were tested in two study areas: (1) Kuala Lumpur (KL) and surrounding areas in Malaysia and (2) Pohang–Gyeongju area in South Korea.

This research has four objectives; the objective number one focuses on developing an alternative technique for landslide inventory modeling based on spatial pattern characterization. The cluster patterns of inventory were extracted and used as training data in constructing the landslide susceptibility models. The resultant map showed that the prediction results of the proposed approach were more accurate than those of the random selection method. The results also exhibited a noticeable reduction in uncertainty, particularly in landslide-prone areas with high to moderate hazard.

The objective number two focuses on the statistical correlation between landslide conditioning factors and the location of occurrence. The ensemble methodology was employed for pairwise relationships generated among the spatial factors by integrating the weights of the evidential belief function (EBF) into the analytical hierarchy process (AHP). The results successfully determined the conditioning factors most relevant to landslide occurrence. The interrelationship between the conditioning factors with respect to landslide occurrence was also considered in the analysis. The objective number two also seeks to develop an ensemble methodology that uses chi-squared automatic interaction detection (CHAID) for the auto-classification of the landslide conditioning factors and whose results are integrated into the logistic regression (LR) model. The most significant landslide conditioning factors were prioritized, and the developed models were validated using success and prediction rate curves. The research findings confirmed the research hypothesis by testing the proposed hybrid ensemble models with landslide locations which were

not used during model building, and comparing it with existing individual models. The prediction maps yielded higher prediction accuracy and achieved better discrimination of susceptible zones than the other models did.

The objective number three focuses on conducting rainfall threshold analysis on Kuala Lumpur (KL) and the surrounding areas. Four different antecedent periods: 5-, 10-, 15-, and 30-day relationships, were used in defining the rainfall threshold of each rainfall station in the study area. The results fill a gap in the literature through the formation of a medium-scale hazard map that was developed based on the multiplied results of the spatial and temporal landslide susceptibility maps of KL and the surrounding areas using available information from 2000 to 2012.

The objective number four focuses on the semi-quantitative risk assessment of landslide hazards in the western and northern regions of KL; only medium-scale data were used because of data scarcity. A valid integration between the elements at risk and the hazard map of 2017 was then accomplished to predict the number of elements that are likely to be affected by direct risks. The resultant methodology employs an exposure-based analysis method to calculate the number of elements at risk and prioritizes land use, population, and road networks. Results showed the approximate percentage of loss in the following areas: about 50% for residential dwellings, 35% for commercial buildings, 17% for industrial buildings, 31% for utilities, 18% for the population, and 27% for road networks. The results prove the capacity of the proposed method to make valid predictions under landslide risk conditions in a data-scarce environment. Missing attributes from damaged records rendered the validation of current findings an impossible task.

The results are expected to not only provide a quick yet comprehensive assessment of future landslide hazards and risks but also serve as a guide for land use planners. The presented methods and information will add a valuable contribution to the landslide hazard and risk assessment of medium scale data analysis.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

RAMALAN SPATIAL BERBAHAYA TANAH RUNTUH DAN KAWASAN RISIKO MENGGUNAKAN PENDEKATAN BERSEPADU PERLOMBONGAN DAN STATISTIK DATA

Oleh

OMAR FAISAL S. ALTHUWAYNEE

September 2014

Pengerusi: Professor Madya Biswajeet Pradhan, PhD
Fakulti: Kejuruteraan

Tanah runtuh ialah satu daripada pelbagai jenis bencana alam yang sering membawa kerosakan teruk harta benda, kerugian ekonomi dan menanggung kos pemulihan yang besar. Kegagalan cerun terhasil daripada berbilang faktor , termasuk aktiviti-aktiviti antropogen, gempa bumi , hujan lebat dan reaksi dari sejumlah bahan permukaan tidak stabil berkaitan dengan geologi, litupan tanah, geometri cerun, kandungan lembapan dan tumbuhan. Tesis ini membentangkan satu pembaharuan dalam GIS berasaskan pendekatan statistik yang dibangunkan untuk pemetaan bahaya tanah runtuh yang disebabkan oleh hujan. Pendekatan-pendekatan ini telah diuji di dua kawasan kajian: (1) Kuala Lumpur (KL) dan kawasan persekitaran Malaysia dan Pohang Gyeongju di Korea Selatan.

Penyelidikan ini mempunyai empat objektif utama; Objektif pertama tertumpu kepada pembangunan satu teknik alternatif untuk pemodelan inventori tanah runtuh berdasarkan pencirian corak ruang. Corak-corak kelompok inventori telah diekstrak dan digunakan sebagai data latihan dalam membina model-model kerentanan tanah runtuh. Peta menunjukkan bahawa keputusan ramalan pendekatan yang dicadangkan lebih tepat berbanding dengan kaedah pilihan rawak. Keputusan juga menunjukkan penurunan ketara yang tidak dikenalpasti, terutamanya di kawasan sering terjadi tanah runtuh sederhana bahaya. Matlamat kedua menumpukan pada korelasi statistik antara faktor terjadinya tanah runtuh dan lokasi kejadian. Kaedah kelompok digunapakai untuk hubungan berpasangan yang dihasilkan oleh faktor ruang dengan mengintegrasikan berat fungsi kepercayaan keterangan (EBF) ke dalam proses analisis hierarki (AHP). Hasil kajian ini telah berjaya mengenalpasti faktor utama kejadian tanah runtuh. Hubungan antara faktor ini diambil kira dalam menganalisis kajian ini.

Objektif kedua juga bertujuan untuk membangunkan satu kaedah keseluruhan yang menggunakan khi- kuasa dua pengesanan interaksi automatik (CHAID) bagi auto-klasifikasi faktor penyebab tanah runtuh dan keputusan ini diintegrasikan ke dalam model regresi logistik. aktor penyebab tanah runtuh yang paling ketara diutamakan dan model yang dibangunkan telah disahkan menggunakan kadar lengkung ramalan. hasil dapatan kajian mengesahkan hipotesis kajian dengan menguji model hibrid

keseluruhan dengan lokasi tanah runtuh dimana ia tidak digunakan semasa membangunkan model dan membandingkannya dengan model individu sedia ada. Peta-peta ramalan menghasilkan ketepatan ramalan lebih tinggi dan mencapai diskriminasi zon-zon rentan lebih baik berbanding model yang lain.

Objektif ketiga memfokuskan kepada analisis nilai had hujan di kuala lumpur (KL) dan kawasan sekitar. Empat tempoh hubungan yang berbeza: 5-, 10-, 15-, dan 30-hari telah digunakan untuk mentakrifkan had hujan di setiap stesen hujan di kawasan kajian. Keputusan kajian ini telah mengisi ruang dalam kesusasteraan melalui penghasilan peta bencana skala medium yang dibangunkan melalui beberapa keputusan ruang dan waktu berasaskan peta kecenderungan tanah runtuh di Kuala Lumpur dan kawasan sekelilingnya meggunakan maklumat sejak dari tahun 2000 hingga 2012.

Objektif keempat memberi tumpuan kepada penilaian kuantitatif separa risiko bahaya tanah runtuh di kawasan barat dan utara Kuala Lumpur: data skala sederhana digunakan kerana terdapat masalah kekurangan data. integrasi yang sah antara unsur-unsur risiko dan peta bencana 2017 terhasil untuk meramalkan bilangan elemen yang mungkin terjejas oleh risiko langsung. Metodologi paduan menggunakan kaedah analisis berdasarkan pendedahan untuk mengira bilangan elemen risiko dan mengutamakan rangkaian penggunaan tanah, penduduk, dan jalan. Keputusan menunjukkan peratusan anggaran kerugian dalam bidang-bidang berikut: kira-kira 50% bagi rumah kediaman, 27% untuk bangunan komersial, 17% untuk bangunan industri, 31% untuk utiliti, 18% bagi penduduk, dan 27% bagi rangkaian jalan raya. Keputusan ini membuktikan bahawa walaupun terdapat kekangan kekurangan data, namun, kaedah yang dicadangkan mampu menghasilkan ramalan yang sah untuk kondisi risiko tanah runtuh. kehilangan atribut daripada rekod kerosakan menjadikan pengesahan untuk penemuan semasa adalah mustahil.

Keputusan yang dijangkakan bukan sahaja menyediakan penilaian yang cepat dan komprehensif malahan berfungsi sebagai panduan untuk perancang guna tanah. Kaedah-kaedah dan maklumat yang dikemukakan akan memberikan sumbangan yang berharga kepada penilaian bahaya tanah runtuh dan penilaian risiko oleh analisis data skala sederhana.

ACKNOWLEDGEMENTS

All praise to ALLAH, Most Gracious, and Most Merciful, Who alone brings forgiveness and light and new life to those who call upon Him, and to Him I dedicate this work.

"Read! In the Name of your Lord who has created (all that exists).
He has Created man from a clot.
Read! And your Lord as the Most Generous.
Who has taught (the writing) by the pen.
He has taught man that which he knew not."
Qur'an (Alaq) 96: 1-5.

I wish to thank my parents, who deserve my sincerest appreciation, for their unselfish love and care as well as for the support and motivation they have always given me. I am grateful for the countless sacrifices they have endured to ensure that I was able to continue pursuing my dreams and for always being there for me. May ALLAH always protect them and bless them with long and healthy lives. Words will not be enough to express all my praise and thanks to them.

I also thank my brothers and sisters for their support, affectionate encouragement, and constant prayers.

I extend my sincerest appreciation to my supervisor, Assoc. Prof. Dr. Biswajeet Pradhan, who supported me throughout my thesis with his thoughtful guidance and insightful suggestions. I am deeply grateful to the worthy members of my supervisory committee, Dr. Noordin Bin Ahmad and Assoc. Prof. Dr. Ahmad Rodzi Bin Mahmoud, for their active involvement and sound advice.

I also wish to express my warmest gratitude to all my laboratory colleagues, technicians, and friends at the Geospatial Information Science Research Centre, formerly the Spatial and Numerical Modelling Laboratory.

Finally, I thank my Uncle (Khazael Salih Althuwaynee) for his endless support and encouragement.

My journey as a student ends with the completion of this thesis. Many people have shared my best and worst moments during the past few years. I thank them all.

I certify that a Thesis Examination Committee has met on 4th September 2014 to conduct the final examination of Omar Faisal S. Althuwaynee on his PhD thesis entitled "Spatial Prediction Of Landslide Hazards And Risk Areas Using Integrated Statistical And Data Mining Approaches" 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:

Mohammad Firuz Ramli, PhD
Associate Professor
Faculty of Environmental Studies
Universiti Putra Malaysia
(Chairman)

Shattri Mansor 1, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Helmi Zulhaidi Mohd Shafri 2, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Yilmaz Isik, PhD
Professor
Faculty of Engineering
Cumhuriyet University, Sivas, Turkey
(External Examiner)

_____

NORITAH OMAR, PhD
Associate Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 19 September 2014

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment on the requirement for the degree of Doctor of Philosophy. The members of the supervisory committee are as follows:

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

Ahmad Rodzi bin Mahmood, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Noordin Ahmad, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(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 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 Vic-Chancellor (Research and Innovation) before thesis is published (in the form written, printed or in electronic form) including books, journals, modules, proceeding, 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 writing of this thesis was under our supervision;
- Supervision responsibility as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____	Signature: _____
Name of Chairman of Supervisory Committee : _____	Name of Member of Supervisory Committee : _____
Signature: _____	Signature: _____
Name of Member of Supervisory Committee : _____	Name of Member of Supervisory Committee : _____

TABLE OF CONTENTS