Investigation of landslide based on high performance and cloud-enabled geocomputation

Jun Liu[1], Shuguang Liu[2,*], Qiming Zhou[3], Jing Qian[1]

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen Guangdong, 518055, China
Telephone: (+86) 18290480019
Email: jun.liu@siat.ac.cn, jing.qian@siat.ac.cn

[1] Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, 400714, China
Telephone: (+86) 13809890997
* Corresponding author, Email: liushuguang@cigit.ac.cn

[2] Department of Geography, Hong Kong Baptist University, Hong Kong, China
Telephone: (852) 34115048
Email: qiming@hkbu.edu.hk

1. Introduction

Landslide, defined as the mass movement of rock, debris or earth down a slope results in a geomorphic makeover of the Earth's surface, can be triggered by various external stimuli. In recent years, landslide hazard assessment has played an important role in developing land utilization regulations aimed at minimizing the loss of lives and damage to property.

An objective procedure is often desired to quantitatively support the landslide studies. According to the literature, many attempts have been made to predict the landslides and to prepare susceptibility maps using different methods (Lee and Pradhan, 2007). Most of these studies were based on the establishment of the relationships between the categories of causative factors and the incidences of landslides in a given region through spatial data analyses (Chauhan, et al. 2010).

There are qualitative and quantitative predictive methods investigating the relationship between environmental variables and landslide occurrence. Qualitative methods depend on experts' opinion and are often useful for regional assessment (Aleotti and Chowdhury 1999; Van Westen, et al. 2003). On the other hand, extensive literatures reported various quantitative methods. Numerous approaches have been proposed, which can be categorized into qualitative factor overlay, statistical models, geotechnical processing models, etc. (Wang, et al. 2005). Predictive models of regional landslide are generally used to identify where landslides may occur based on a set of relevant environmental characteristic, assuming that slope failures will be more likely to occur with the similar conditions of past slope movement (Carrara, et al. 1995). Zhou, et al. (2002) presented a statistical approach to study the spatial relationship between landslides and their causative factors at regional level. Many of the recent studies have applied probabilistic models, e.g. Pradhan, et al. (2006, 2009), and Lee and Pradhan (2007). On the other hand, Pradhan, et al. (2011) applied frequency ratio-based statistical techniques for landslide hazard analysis and achieved an accuracy of 89%.

In addition, it is well recognized that the various factors contributing to landslides in a region are complexly interrelated and the relationships between these factors and the landslides are nonlinear in nature. Therefore, AI (Artificial Intelligence) techniques such as fuzzy logic, artificial neural networks, and genetic algorithms were also adopted. All the methods proposed involve an increasing degree of analysis, and rigor, not necessarily an increasing accuracy in the assessment of probability. Many consider that a landslide hazard map obtained by systematic data manipulation with the support of GIS is more objective than manual operation. However, to date, there is no general agreement on the methods or even the scope of these investigations (Wang, et al. 2005). In addition, use of GIS has been emphasized in almost every landslide study published in recent years. Therefore, the general trend related to landslide assessments seems to be the utilization of quantitative methods and especially, GIS based ones (Ercanoglu, 2008).

Landslide susceptibility analysis and its early warning system require not only the understanding of its controlling factors, but also its triggering factors. Rainfall is an obvious one in the rainfall-induced landslides, and rainfall-runoff process significantly affects the slope stability. Understanding and monitoring the rainfall-runoff process will provide a strong basis for investigating slope stability and hence landslide susceptibility. Beven and Kirkby (1979) proposed a DEM-based hydrologic model (TOPMODEL) to derive terrain parameters which control the distribution of hydrologic/soil properties. Another well-known attempt is the Soil and Water Assessment Tool (SWAT) developed by Arnold, et al. (1994). The SWAT model is a semi-distributed model, which integrates GIS and remotely sensed data on a grid-based data structure, to provide spatial information for extracting hydrological response units (HRUs) with homogeneous hydrologic characteristics. However, most of these GIS-based hydrological models were designed for the investigations of water resources, so that they generally deal with the hydrological characteristics in long term with coarser spatial and temporal scales. The models often need to be calibrated using in-situ hydrological data before being applied into a geographical region. For real-time rainfall-runoff processes, a dynamic model, which can calculate key hydrologic parameters at real-time (e.g. soil moisture, generation of surface flow, flow volume and velocity, and spatio-temporal distribution of accumulated flow), is required.

In this paper, we describe the framework and platform of our study on the investigation of landslide based on high performance and cloud-enabled geocomputation. The mail objective of our study is to construct cloud-computation platform for verifying and refining the models based on high-performance computing technology, improving the algorithms and models by parallelization and optimization, and achieving the real-time dynamic landslide early warning services. Through the open cloud-service platform, this platform is expected to facilitate the sharing and collaboration of landslide research and contribute to land use planning decision making and disaster mitigation effort.

2. The structure of our platform

In this section, our cloud-enabled landslide early warning platform is described. This platform takes an integrated approach to address two research issues, namely, how the environmental variables collectively influence the occurrence of rainfall-induced landslides by altering the rainfall-runoff process, and how to implement a real-time, regional specific platform that allows early warning of potential disastrous consequence in a timely manner. The integration of scale-adaptive DTA method and the flow-path network model, for rainfall-runoff modelling will contribute to uncover the mechanism behind landslide occurrence and surface flow path in a regional scale. To create an experiment platform for real-time landslide risk assessment, it is desirable to integrate the two components together. The methodologies and individual objectives to be addressed are described below.

No.	*Methodologies*
1	Study site selection and reconnaissance field investigation
2	Spatial data acquisition and specification
3	Hydrological ground data collection and rainfall/runoff analysis
4	Surface/sub-surface water discharge analysis
5	The development of landslide susceptibility and risk analysis model
6	Field tests and rainfall-runoff simulation experiment
7	Computer platform implementation
8	System calibration and evaluation

Table 1.Objectives of our platform.

The whole platform can be divided into the following three layers:

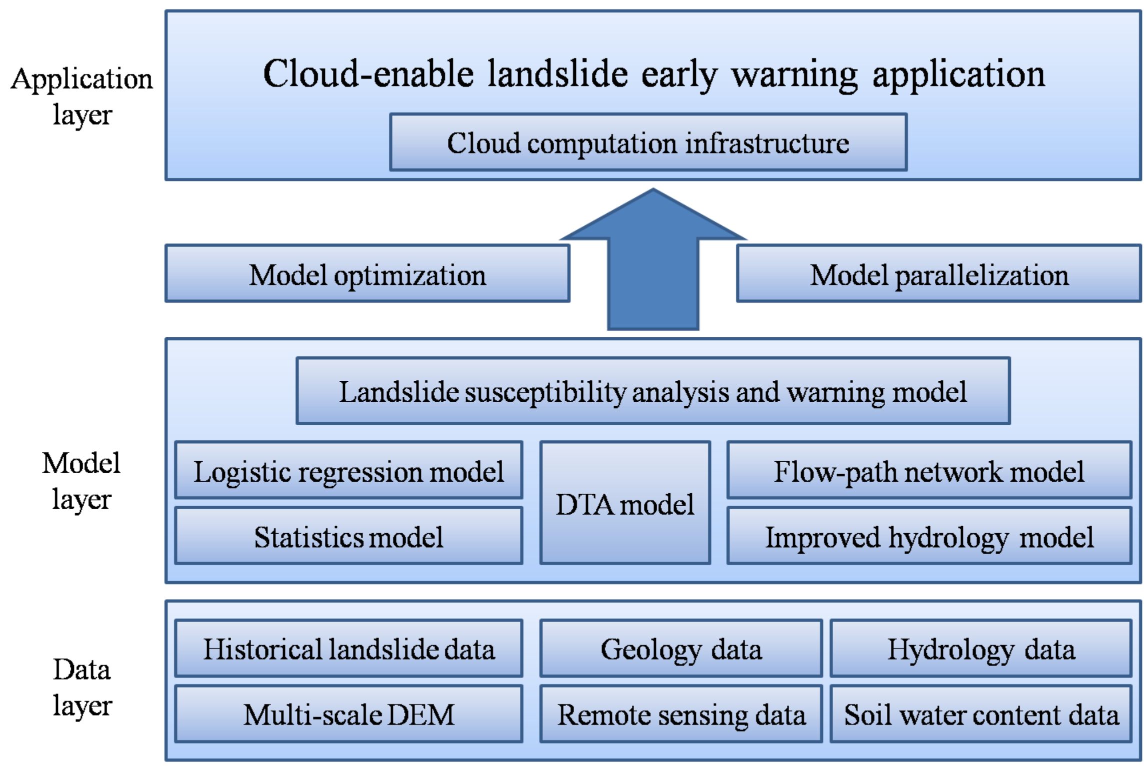


Figure 1. The structure of the platform.

(1) Data layer

All the input data are provided in this layer, including historical landslide data, multi-scale DEM, geology data, remote sensing data, hydrology data, soil water content data and other possible and available data. These data are the foundation of models and application in this platform.

(2) Model layer

After the collection of all data, some models will be carried out to produce more results that benefit to build landslide susceptibility analysis and warning model. The logistic regression model and statistics model will be firstly used to find out the relationship between environmental factors and landslide; then some hydrology models will be tested and compared to proposed an improved hydrology model; finally on the basis of our previous work, including DTA model, flow-path network model and so on, and the studies on the implementation and parallelization of the above models, the landslide susceptibility analysis and warning model can be achieved.

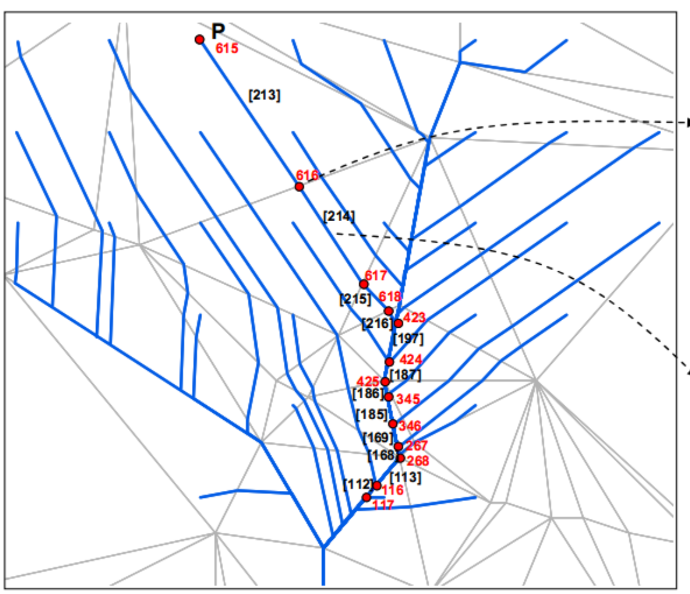
(3) Application Layer

The main objective of this platform is to dynamically monitor and early warn the development tendency of landslide using DTA and hydrology approaches. Therefore, in this layer, an application will be constructed. Since the input data vary in sources, types, structures and volumes, the cloud computation technology will be used to provide high-performance geocomputation services.

3. High-performance experiments and discussions

In our platform, the cloud-enable high-performance geocomputation is used to parallelize the models. For a specific scale DEM, the flow-path network can be produced using our adaptive DTA algorithm; and then the particle system is used to simulate the track of raindrops that flow along with the flow-path network; and finally the volume and velocity can be calculated by sum up all the particles that go through some specific flow-path network. This process is parallelized using high-performance geocomputation approach.

The following figure illustrates the data structure of the flow-path network which is our previous work. The original DEM is used to detect the key point and then build the flow-path network. Every node and line of the network is stored topologically as shown in tables. Through this network, the raster DEM has been transformed into the sequential tables which is helpful for parallelization.



Node table

Node ID	X (m)	Y (m)	Z (m)	...
615	402306	4072762	1169.52	...
616	402338	4072715	1129.89	...
617	402359	4072683	1115.94	...
...	...	...	...	...

Line table

Line ID	Start node	End node	Slope	length (m)	velocity (m/s)	...
213	615	616	20	217.5	=v(...)	
214	616	617	15	135.1	=v(...)	
215	617	618	10	32.4	=v(...)	
...	...	...	...	...	...	...

$$v = f(r, s, n)$$

Figure 2. The structure of the flow-path network.

In our platform, the simulation of rainfall is essential for the landslide model. Hence, in this section, only the parallelization simulation of rainfall is tested. We assume that many raindrops drop at the source of every line, and then track the movement of every raindrop using particle system. Due to the structure of flow-path network, we only have to track the raindrops along with every line.

We set up a small cluster to complete this simulation. The cluster consists of three normal personal computers. The configuration of cluster is:

ID	CPU type	CPU core	CPU frequency	RAM
1	Intel® Core™ i5-2500	4	3.30GHz	4G
2	Intel® Core™ i5-2500	4	3.30GHz	4G
3	Intel® Core 2 E7400	2	2.80GHz	4G
Software environment		Ubuntu 12.04 LTS		
Runtime environment		Message Passage Interface (MPI)		

Table 2. The configuration of cluster.

The test consists of two parts:

(1) The influence of CPU cores on computing time

The main purpose of this test is to determine how the CPU cores influence the computing time. If this influence is linear, we can use more cores to decrease the consuming time. The results are shown as follows:

Figure 3. The influence of CPU cores on computing time

(2) The influence of raindrop number on computing time

The main purpose of this test is to determine how the raindrop (particle) number influence the computing time. The results are shown as follows:

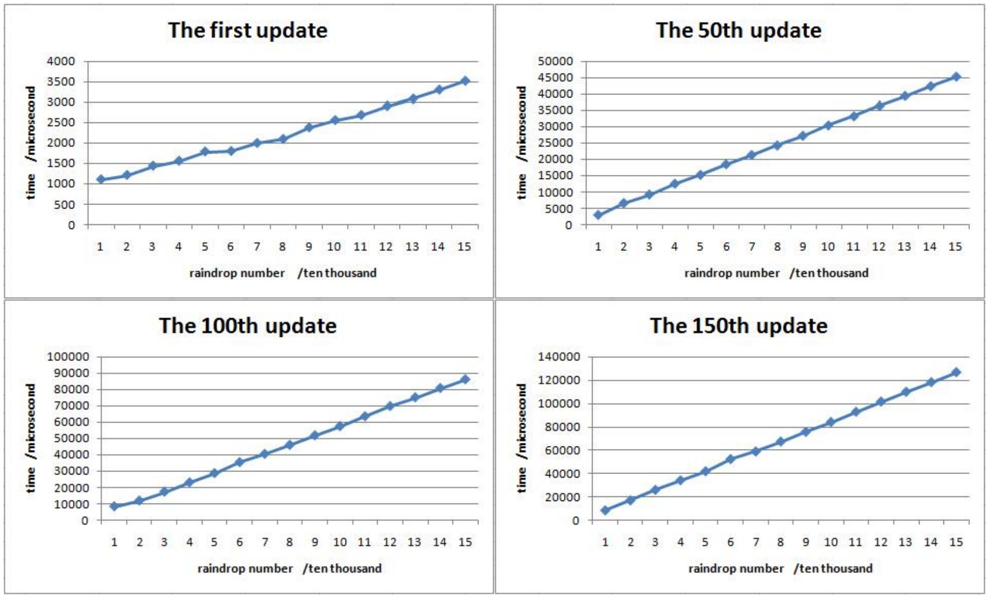


Figure 4. The influence of raindrop number on computing time

From figure 3 and 4 we can see that as the increase of CPU cores, the computing time decrease approximately linearly; meanwhile as the increase of raindrop number, the computing time increase linearly. These results indicate that the platform shows good linear characteristics in parallel high-performance computation and real-time computation.

4. Conclusion

In this paper, an high-performance and cloud-enable geocomputation-based landslide monitoring platform is introduced, and a testing of parallelization simulation indicates that the platform shows linear characteristics and is suitable to simulate the rainfall using the DEM in a region. In the future work, the interaction between raindrop and environmental factors such as gradient, slop, aspect and other auxiliary data will be taken into account to simulate rainfall more precisely.

5. Acknowledgements

This work is jointly supported by the International Science and Technology Collaboration Project of China (2010DFA92720-24), the National Natural Science Foundation program (No. 41301403 and No. 41471340) and Chongqing Basic and Advanced Research General Project (No. cstc2013jcyjA40010), the Research Grants Council (RGC) of Hong Kong General Research Fund (GRF) (Project No. 203913), and Hong Kong Baptist University Faculty Research Grant (FRG1/12-13/070 and FRG2/11-12/030).

6. References

Aleotti, P., Chowdhury, R., 1999. Landslide hazard assessment: summary review and perspectives, *Bulletin of Engineering Geology and the Environment*, 58: 21-44.

Arnold, J.G., Williams, J. R., Srinivasan, R., King, K.W., Griggs, R.H., 1994. SWAT - Soil and Water Assessment Tool - User Manual. Agricultural Research Service, Grassland, Soil and Water Research Laboratory, US Department of Agriculture.

Beven, K.J. and Kirkby, M. J., 1979. A physically based, variable contributing area model of basin hydrology, *Hydrological Sciences Bulletin*, 24: 43-69.

Carrara, A., Cardinali, M., Guzzetti, F., Reichenbach, P., 1995. GIS technology in mapping landslide hazard. In Carrara, A. and Guzzetti, F., (Eds.), Geographical Information Systems in Assessing Natural Hazards. Dordrecht: Kluwer, 135-75.

Chauhan, S., Sharma, M., Arora, M. K., Gupta, N. K., 2010. Landslide Susceptibility Zonation through ratings derived from Artificial Neural Network, *International Journal of Applied Earth Observation and Geoinformation*, 12: 340-350.

Ercanoglu, M., 2008. An Overview on the Landslide Susceptibility Assessment Techniques, *in The 1st WSEAS International Conference on Environmental and Geological Science and Engineering (EG '08)*, 131-134.

Lee, S., Pradhan, B., 2007, Landslide hazard mapping at Selangor, using frequency ratio and logistic regression models, *Landslides*, 4: 33-41.

Pradhan, B., Lee, S., Bucheroithner, M. F., 2009. Use of geospatial data for the development of fuzzy algebraic operators to landslide hazard mapping: a case study in Malaysia, *Applied Geomatics*, 1: 3-15.

Pradhan, B., Lee, S., 2010. Regional landslide susceptibility analysis using backpropaga-tion neural network model at Cameron Highland, Malaysia, *Landslides*, 7: 13-30.

van Westen, C. J., Rengers, H., Soeters, R., 2003. Use of geomorphological information in indirect landslide susceptibility assessment, *Natural Hazards*, 30: 399-419.

Wang, H., Liu, G., Xu, W., Wang, G. H., 2005. GIS-based landslide hazard assessment: an overview, *Progress in Physical Geography*, 29(4): 548-567.

Zhou, Q., Liu, X., 2002. Error assessment of grid-based flow routing algorithms used in hydrological models, *International Journal of Geographical Information Science*, 16 (8): 819-842.