Understanding of Important Factors Pertaining to Occurrence of Shallow Landslide in Nachi River Basin on September 4, 2011 Using the Data Mining Method

Takehiro SAKAGUCHI[1*], Tsuneshi NISHIOKA[1], Kazuo TSUTSUI[1], Kazuhisa FUKUDA[1], Atsuhiko KINOSHITA[2], Yasutaka TANAKA[3], Noriko TADAKUMA[4] and Soichi KAIHARA[4]

[1] Wakayama Sabo Research and Education Institute, Japan
[2] National Institute for Land and Infrastructure Management, MLIT, Japan
[3] Sediment Disaster Prevention Technology Center, Kinki Regional Development Bureau, MLIT, Japan
[4] Eight-Japan Engineering Consultants Inc., Japan
*corresponding author;E-mail:sakaguchi_t0013@pref.wakayama.lg.jp

INTRODUCTION

In September, 2011, Typhoon Talas caused a lot of sediment disasters in the Nachi River Basin in the area of Nachi-Katsuura Town, Wakayama Prefecture with total rainfall and hourly rainfall exceeding 1,000 mm and 120 mm respectively. Featuring shallow landslide caused by Typhoon Talas in the Nachi River Basin and using rough set which is a data mining method, this study extracted the geographical, geological and rainfall factors closely related to the occurrence of a sediment disaster and analyzed the conditions for disaster occurrence or non-occurrence.

ANALYSIS METHOD

The concept of "rough set"[1] is shown in **Fig. 1**. Let us consider the case where both the slope gradient and slope height are classified into four categories each and then disaster-hit basins (o) and disaster-free basins (•) are plotted on the chart. In rough set, the data in those cells with a white background where all data is classified as either disaster-hit or disaster-free is called coherent data. In contrast, the data in those cells with a grey background where both disaster-hit and disaster-free basins are plotted is called incoherent data. The ratio of coherent data to the total number of data is called coherency and any factor increasing coherency can be considered to be an important factor. Meanwhile, when the elimination of geography condition 2 in **Tab. 1** does not affect the occurrence or non-occurrence of a disaster, this geography condition 2 can be eliminated from the database. This process is called reduction. The remaining geography conditions 1 and 3 are, therefore,

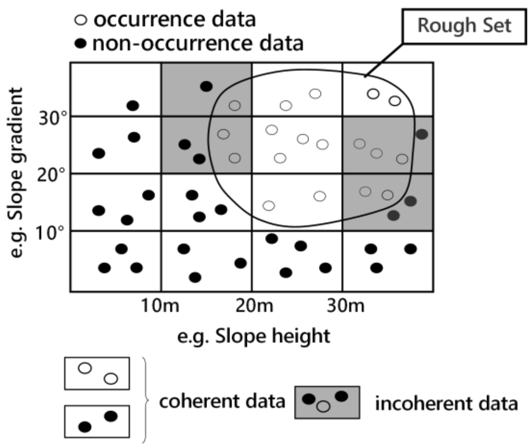

Fig.1 The concept of rough set

Tab.1 Image of the reduction of the database using rough set

drainage basin ID	geography condition 1	geography condition 2	geography condition 3	disasters results
A	large	low	large	occurrence
B	small	high	small	non-occurrence
C	small	low	large	non-occurrence
D	large	high	large	occurrence
E	large	low	small	non-occurrence

eliminate geography condition 2

drainage basin ID	geography condition 1	geography condition 3	disasters results
A	large	large	occurrence
B	small	small	non-occurrence
C	small	large	non-occurrence
D	large	large	occurrence
E	large	small	non-occurrence

considered to be important factors affecting the occurrence or non-occurrence of a disaster.

In this study, each evaluation unit of 100 x 100 m mesh was classified as either disaster-hit mesh or disaster-free mesh to develop a database based on the interpretation results of collapsed sites in the subject basin in 2011. From this database, factors considered to be highly relevant to either the occurrence or non-occurrence of a sediment disaster were extracted using rough set from 28 possible factors (**Tab.2**) and the conditions for the occurrence or non-occurrence were examined. Each of these 28 possible factors, including short and long-term rainfall data based on mesh-specific rainfall data collected by radar rain gauges and corresponding return period of exceedance and such geographical factors as the slope direction and gradient, geological category, aerial electromagnetic wave prospecting results, etc. were classified into 5 – 8 categories depending on the subject item. Aerial electromagnetic wave prospecting results were included for examination as an index for the ground water content and points for variation of the geological structure, etc.

DISCUSSION AND CONCLUSIONS

As a result of analysis using rough set, it was possible to determine either occurrence or non-occurrence with a coherency of approximately 96% with combinations of seven factors out of the originally considered 28 factors. Such factors as “slope direction”, “difference in river level” and “valley density” are present in every combination, followed by “cumulative rainfall analyzed”, “difference in summit level”, “First Derivation of Electromagnetic Wave Prospecting Results (Depth)” and “Return Period of 24 Hours” .

Tab.2 List of Combination between possible factors and identified causative factors

a	b	c	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
7	7-1	96.718	●		●		●	●						●			●									●			
	7-2	96.718	●			●	●	●						●	●											●			
	7-3	96.718	●			●	●	●						●		●												●	
	7-4	96.718	●			●	●	●						●			●									●			
	7-5	96.783	●		●		●	●						●		●										●			
	7-6	96.783	●			●	●	●		●						●										●			
	7-7	96.783	●			●	●	●						●		●										●			
	7-8	96.783	●			●	●	●						●														●	
d		-	8	0	2	6	8	8	0	1	0	0	0	7	1	4	2	0	0	0	0	0	0	0	0	6	0	2	0

Explanatory notes:

a:Number of Factors b: Combination Number C:Coherency d:Frequency of Selection

A: Slope Direction B:Geological Category C:Slope Gradient D:Difference between Summit Level and Grand Level E:Difference between River Level and Grand Level F:Valley Density G:Rainfall Analysed for 10 Minutes H:Rainfall Analysed for One Hour I:Rainfall Analysed for 24 Hours J:Rainfall Analysed for 36 Hours K:Rainfall Analysed for 48 Hours L:Cumulative Rainfall Analysed M:Return Period of One Hour N:Return Period of 24 Hours O:Return Period of 36 Hours P:Return Period of 48 Hours Q:Cumulative Return Period R:Return Period of One Hour (Nachi observatory) S:Sediment Disaster Warning Information T:Soil Water Index U:Specific Resistance Value V: Surface Specific Resistance Value W: First Derivation of Electromagnetic Wave Prospecting Results X: First Derivation of Electromagnetic Wave Prospecting Results (Depth) Y: Secondary Differentiation of Electromagnetic Wave Prospecting Results Z: Secondary Differentiation of Electromagnetic Wave Prospecting Results (Depth) AA:Neighbouring Difference in First Derivation

It is inferred that the reason for the selection of “slope direction” in this study is that the Nachi River Basin lies to the east of the path of Typhoon Talas, causing wet air to directly hit the southern slope to produce a high level of rainfall. The selection of “difference in river level” and “valley density” as important factors is believed to not pose any contradiction as these are indexes of historic erosion to create the current topography.

REFERENCES

1) Pawlak Z: Rough Sets, International Journal of Computer and Information Sciences, Vol.11, pp.341-356, 1982

Keywords: Shallow landslide, Geographical factors, Geographical conditions, Rainfall conditions, Rough set