

1982年7月長崎災害におけるラフ集合を用いた 土石流発生・非発生の評価ルール設定に 関する研究

佐藤 文晴¹・河野 泰典²・荒木 義則³・中山 弘隆⁴・水山 高久⁵・古川 浩平⁶

¹正会員 修(理) 株式会社エイトコンサルタント事業開発本部(〒700-8617 岡山県岡山市津島京町 3-1-21)

²宮崎市役所 教育委員会事務局(〒880-8505 宮崎県宮崎市橘通西 1-1-1)

³正会員 博(工) 中電技術コンサル株式会社 河川部(〒734-8510 広島県広島市南区出汐 2-3-30)

⁴工博 甲南大学教授 理工学部情報システム工学科(〒658-0072 兵庫県神戸市東灘区岡本 8-9-1)

⁵正会員 農博 京都大学大学院教授 農学研究科(〒606-8502 京都府京都市左京区北白川追分町)

⁶正会員 工博 山口大学教授 工学部社会建設工学科(〒755-8611 山口県宇部市常盤台 2-16-1)

1982年7月長崎災害は、日本観測史上最も激しい集中豪雨により、多数の土石流が発生した事例である。しかしながら隣接する溪流で土石流が発生していない場合もある等、土石流の発生は誘因となる降雨に加えて素因となる地形の影響を受けている。本研究では、ラフ集合を用いて土石流危険溪流の地形要因から、土石流発生に関する重要要因を抽出した。また、ラフ集合におけるルールの領域を拡張させることで、土石流ルールを新しく定義し、その有用性を検討した。

Key Words: rough sets, data mining, debris flow, rule

1. はじめに

我々の国土は、変化に富む豊かな自然と共に様々な自然災害をもたらしてきた。特に豪雨による土石流やがけ崩れ等に代表される土砂災害は未だに後を絶たず、甚大な被害を与えることも少なくない。土砂災害による人的被害防止は重要な課題であり、そのためにはハード対策の整備と共に、ソフト対策として「いつ」と「どこで」発生するかをできるだけ高い精度で予測することが求められる。

「いつ」という点(誘因)に関しては、豪雨・地震・火山活動等が挙げられるが、豪雨に伴う土砂災害の発生が多く、降雨観測体制の充実が必要となる。現在、豪雨時における土石流発生予測手法の研究では、時間雨量及び実効雨量を用いた土石流発生危険基準線の設定^{1), 2), 3)}について、種々の提案がなされ、一部実用化されている。

「どこで」という点(素因)に関しては、地形と地質の要因を考慮する必要がある。既往の研究では、土石流と降雨に関する研究^{4), 5), 6)}、統計論やAI(人工生命)技術による地形と降雨を組合せた研究等^{7), 8)}、

様々な見解が示されているが、土石流発生のメカニズムは複雑であり、どのような地形を有する溪流で、土石流が発生しやすいかは未だに明確になっていない。

そこで、本研究では1982年7月に日本観測史上最も激しい集中豪雨により長崎市周辺で激甚な災害をもたらした土石流災害を対象とし、土石流発生・非発生の事象の境界を明確にするためには、どのような地形要因の組合せが最も重要であるかを、ラフ集合を用いたデータマイニングを適用して分析を行った。さらに、ラフ集合で定義されるルールの領域を拡張させることで、矛盾のない土石流発生・非発生の評価ルール(以下土石流ルール)を定義し、その有用性を検討した。

2. ラフ集合の概要

(1) ラフ集合による重要要因抽出の概要

ラフ集合は1982年にZ. Pawlak⁹⁾によって提唱されたもので、データ分析等において、データを条件ご

とに「類別」し、不要な属性を取り除くことによって、できる限り簡単なルールを抽出する手法である。この手法を地形要因に適用し、土石流発生重要要因の抽出を行った。

図-1 にラフ集合の概念図を示す。土石流が発生した溪流を●、非発生の溪流を○とすると、発生のみ集合は図-1 に示される部分集合となる。全体集合を2要因で計9個のカテゴリーに分類しても、発生溪流の部分集合は、完全に分類されない。このように、カテゴリー分類によって完全に分離されない集合をラフ集合 (rough sets) という。これらの要因によって分類されたカテゴリーのうちデータがすべて同種 (全て発生または非発生) であればそのカテゴリー内のデータを整合データ (図中影の部分)、混在しているようであれば矛盾データと定義する。要因を増加させると、細かい部分集合が多数になるため、整合データが増加する。ここで、全データ中に存在する整合データの割合を整合度として式(1)のように定義する。

$$\text{整合度} = \frac{\text{整合データ数}}{\text{全データ数}} \quad (1)$$

図-2 に要因数と整合度の関係を示す。同じ要因数の組合せにおける整合度の取り得る範囲を概念的に示している。整合度を高くするためには、以下の2つの手法が考えられる。

- 1) 要因数を増加させる方法 (図-2 中 A の方法)。
- 2) 要因の組合せを変化させ、その最適な組合せを発見する方法 (図-2 中 B の方法)。

このうち、A については、土石流発生の原因 (どの要因が発生に効いているのか) となる重要要因の抽出という目的から外れる。いかに少ない要因数で高い整合度が得られるかが重要であるため、本研究では整合度の要求水準を設定し、方法 B について検討を行う。

(2) ラフ集合論におけるルールの抽出

重要要因の組合せによって分割されたカテゴリーのうち整合データのみで構成されるカテゴリーのことをルールという。ルールは「IF (条件部) THEN (結論部)」という形で簡単に表記できる。本研究での条件部は、地形要因で構成され、結論部には土石流発生あるいは非発生となる。この節ではラフ集合における一般的なルール抽出法を述べる。ルール抽出のフローを図-3 に、またルール抽出の一例を表-1 から表-4 までに示す。

まず要因に注目し、取り除いたときに矛盾が生じ

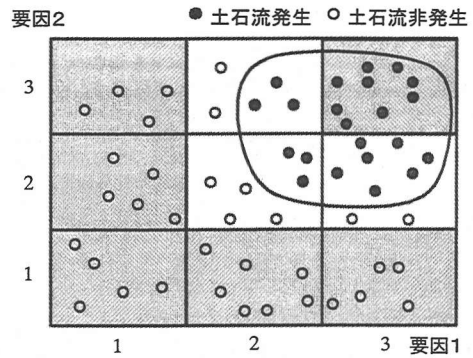


図-1 ラフ集合の概念

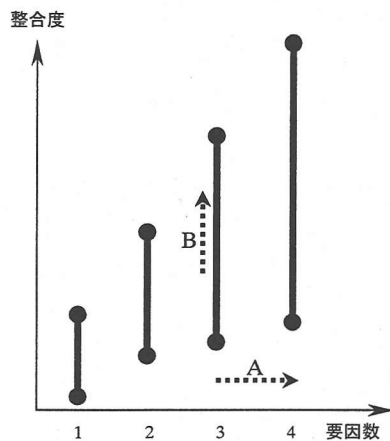


図-2 要因数と整合度の関係

るかどうかが確認する。矛盾が生じない場合はその要因が削除される。例えば表-1 で要因 1 を削除した場合を考えると、溪流 2 と溪流 7 は共に要因 2 のカテゴリーが 2、要因 3 のカテゴリーが 1、要因 4 のカテゴリーが 2 と全く同じであるにも関わらず、判定が異なっており、矛盾が生じている。つまり、要因 1 は重要な要因として考えられる。同様にして各要因で行うと、要因 3 を削除しても、矛盾が生じない。要因 3 を削除した表が表-2 である。

次に各要因のカテゴリーを削除する。削除したときに判定に矛盾が生じないカテゴリーを削除する。例えば表-2 で溪流 1 の要因 1 を削除しても要因 2 のカテゴリーが 1、要因 4 のカテゴリーが 2 であれば判定は 0 であり、これに矛盾するデータはないので削除できる。これに対し、溪流 2 の要因 1 を削除すると、要因 2 のカテゴリーが 2、要因 4 のカテゴリーが 2 となり溪流 7 と同じである。この 2 つの溪流は判定が異なることから、矛盾が生じている。つ

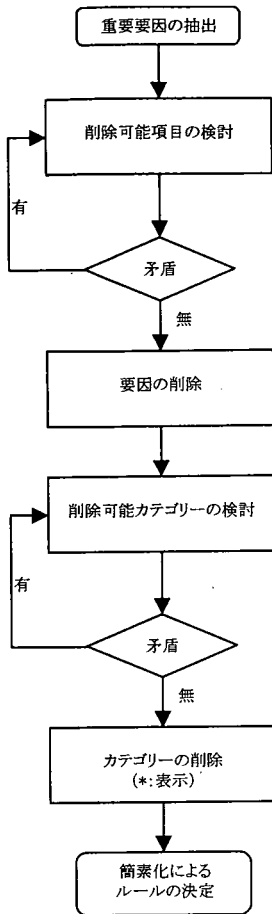


図-3 ルール抽出フロー

まり、溪流2の要因1は削除できない。このようにして必要のないカテゴリーをすべて削除したのが表-3である。アスタリスク(*)はどんな値でもよいという意味である。表-3を簡潔にまとめたのがルールであり、表-4に示す。

ここで抽出されたルールの精度と汎用性を表す尺度としてそれぞれ確信度(式(2)), サポート(式(3))を定義する。以下ルールの重要度は、確信度とサポートで評価する。

$$\text{確信度} = \frac{\text{ルールに含まれる整合データの溪流数}}{\text{ルールに含まれる全溪流数}} \quad (2)$$

$$\text{サポート} = \frac{\text{ルールに含まれる溪流数}}{\text{全溪流数}} \quad (3)$$

表-1 削除可能項目の検討

	要因1	要因2	要因3	要因4	判定
溪流1	2	1	1	2	0
溪流2	1	2	1	2	0
溪流3	1	1	2	1	0
溪流4	2	2	2	1	0
溪流5	2	1	2	1	1
溪流6	2	1	1	1	1
溪流7	2	2	1	2	1

表-2 不要な要因の削除(要因3)

	要因1	要因2	要因4	判定
溪流1	2	1	2	0
溪流2	1	2	2	0
溪流3	1	1	1	0
溪流4	2	2	1	0
溪流5	2	1	1	1
溪流6	2	1	1	1
溪流7	2	2	2	1

表-3 不要なカテゴリーを削除

	要因1	要因2	要因4	判定
溪流1	*	1	2	0
溪流2	1	*	*	0
溪流3	1	*	*	0
溪流4	*	2	1	0
溪流5	2	1	1	1
溪流6	2	1	1	1
溪流7	2	2	2	1

表-4 ルール決定

ルール番号	要因1	要因2	要因4	判定
Rule1	*	1	2	0
Rule2	1	*	*	0
Rule3	*	2	1	0
Rule4	2	1	1	1
Rule5	2	2	2	1

3. 使用データの詳細

長崎市では、土石流危険溪流に関して詳細な調査が行われている。そこで、本研究では長崎災害時に土石流が発生した溪流を調査するため、長崎市内における土石流危険溪流(612溪流)について地形要因の計測^{10), 11)}を行った。そのうち、土石流が発生した溪流は150溪流であり、非発生溪流は残りの462溪流である。

地形要因の抽出は、地形図¹⁰⁾からの計測^{7), 8)}によるもので、1/5,000の森林基本図を基にする。また土砂移動現象は土石流発生のみを用い、山腹斜面崩壊等の小規模な土砂移動は土石流非発生として扱った。地形要因の抽出は、一次谷に分割した土石流危

表-5 地形要因一覧表

地形要因	番号	地形要因の説明
水系模様	1	流水の流下形態。(最短溪流長 / 主溪流長)の比。主溪流長の始点から終点までの直線最短距離を主溪流長で除した値。
流域平均勾配(°)	2	流下のしやすさ。(流域の上流端と原点の高低差) / (流域の上流端と原点の距離)で求める。このときの上流端とは1次谷の最上流点とし、原点とは流域の最下流点とする。
最急溪床勾配(°)	3	溪床の移動のしやすさ。溪床において高低差30mとなる等高線の間隔の最小値を計測し、勾配(角度)を逆算して求める。
主溪流長(km)	4	流水の流下距離。1次谷の長さ。流域の原点付近は谷地形を示していない場合があるが、中流域及び上流域で1次谷が150m以上連続していれば、谷の出口から原点にかけても谷(溪流)が連続しているとみなす。
流域面積(km ²)	5	集水面積の大小。氾濫開始点、及び原点より上流の流域面積。
流域長(km)	6	流域の内溪流方向の長さ。
流域幅(km)	7	流域の内溪流方向の幅。
流域形状比	8	流域形状(流域幅 / 流域長)で求められる。
谷深比	9	谷の発達程度。(主溪流長 / 流域長)で求められる。
溪流方位	10	溪流の向き。東西南北の四方向のうちどの方向に流下しているか。
0次谷の数	11	土石流発生場の数。等高線間隔10mの地形図で6本(高低差50m)以上の等高線において凹地形が連続していれば、0次谷として扱う。
流域最大傾斜(°)	12	斜面崩壊のしやすさ。最急溪床勾配と同様に、斜面上で、高低差30mとなる等高線の間隔の最小値を計測し、逆算により、勾配(角度)を求める。
源頭部面積(km ²)	13	土石流発生源の面積。主溪流の終点より上流の流域面積。土石流危険溪流が0次谷であれば流域面積と同じになる。終点とは1次谷と0次谷の境界点とする。
源頭部最急谷勾配(°)	14	土石流発生源の傾斜。源頭部内で、最急溪床勾配と同様に、谷部で高低差30mとなる等高線の間隔の最小値を計測し、逆算により、勾配(角度)を求める。
降雨集中度評価	15	流域の形状と谷の発達状況から降雨の集中度を5段階で評価したもの

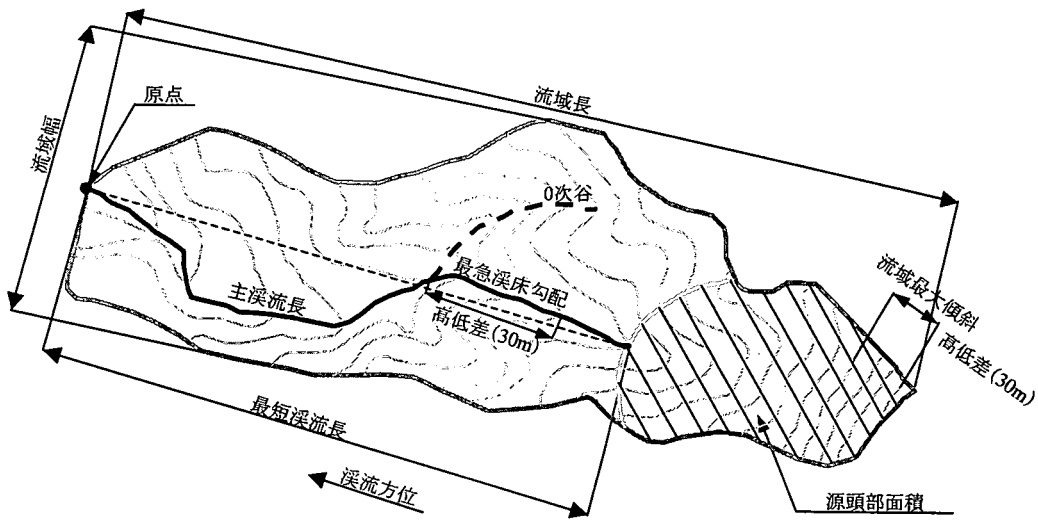


図-4 地形要因の概要図

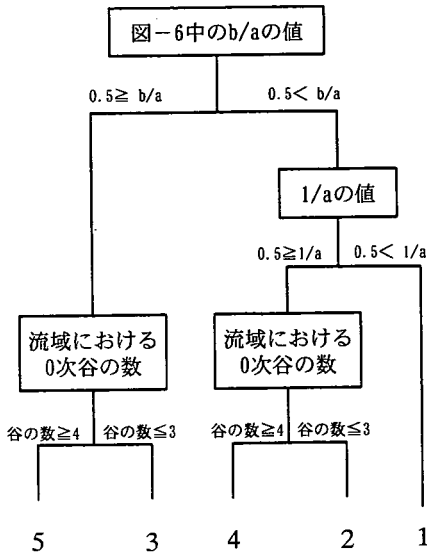


図-5 降雨集中度評価判定図

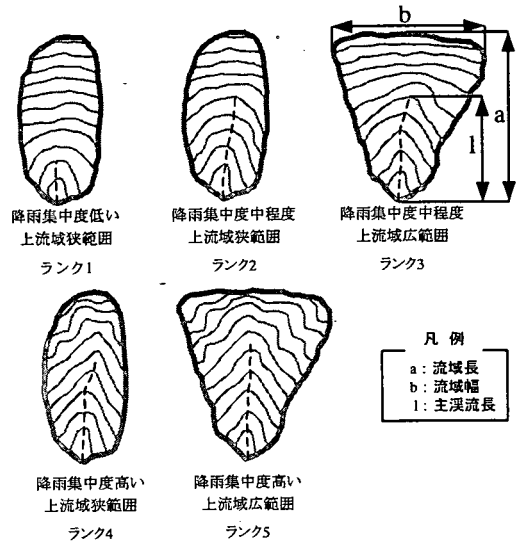


図-6 降雨集中度評価の概要図

険渓流について、土石流発生に重要と考えられる15要因を計測した。表-5に15個の地形要因を、また図-4に地形要因の概要について示す。

抽出された15項目の内、降雨集中度評価については地形図上では計測できない。この項目は、溪流の形状と谷の発達状況から降雨の集中度（集水のしやすさ）を5段階で評価したもので、図-5に従って分類する。図-5の下にある1~5の数値が図-6のランク番号に該当している。また図-5のa, b, lは図-6の形状例の中に示す。

4. 土石流発生に関わる重要要因の組合せの抽出

第4章以下の論文の構成を図-7に示す。本論文では、土石流ルールの設定を目的としていることから、まず最適な重要要因の組合せを求め、土石流ルールを定義し、その適用性などについて検討する。

(1) データのカテゴリ化

ラフ集合による重要要因の抽出では離散化されたデータを使用するため、地形要因の連続値データは表-6に示すような範囲で各要因のカテゴリ分類を行い離散値データとして扱った。

カテゴリ数が著しく異なると、一般的にカテゴリ数が多い要因が重要要因として抽出される。これは、多数のカテゴリによって細分化されること

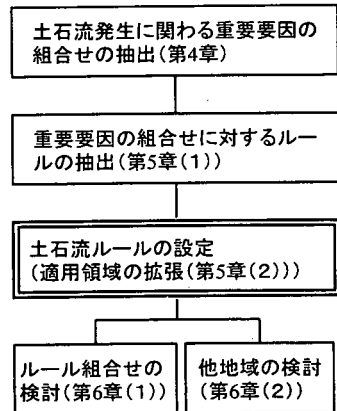


図-7 本論文の構成

で、各々の部分集合に属する溪流数が少なくなり、整合データとして評価される確率が増加するためである。したがって、ラフ集合を用いる場合は出来るだけカテゴリ数をそろえる必要がある。

さらに、地形要因中の溪流方位と降雨集中度評価の2要因が元のデータからカテゴリ分類されており、既にカテゴリ数が決定していることから、本研究では、他の地形要因に対しても同様のカテゴリ数（4個から5個）に分類して、分析に用いた。また、各要因のカテゴリ配分は、データの最大値、最小値及び分布状況を考慮し、等間隔に設定した。

表-6 カテゴリー分類

地形要因	カテゴリー分類				
	1	2	3	4	5
水系模様	~0.9	0.9~0.95	0.95~1	1~	
流域平均勾配 (°)	~15	15~20	20~25	25~30	30~
最急溪床勾配 (°)	~10	10~15	15~20	20~25	25~
主溪流長 (km)	~0.1	0.1~0.2	0.2~0.3	0.3~0.4	0.4~
流域面積 (km ²)	~0.05	0.05~0.1	0.1~0.15	0.15~0.2	0.2~
流域長 (km)	~0.15	0.15~0.3	0.3~0.45	0.45~0.6	0.6~
流域幅 (km)	~0.1	0.1~0.2	0.2~0.3	0.3~0.4	0.4~
流域形状比	~0.3	0.3~0.5	0.5~0.7	0.7~0.9	0.9~
谷深比	~0.2	0.2~0.4	0.4~0.6	0.6~0.8	0.8~
溪流方位	北	東	南	西	
0次谷の数	0個	1個	2個	3個	4個以上
流域最大傾斜 (°)	~35	35~40	40~45	45~50	50~
源頭部面積 (km ²)	~0.005	0.005~0.01	0.01~0.015	0.015~0.02	0.02~
源頭部最急谷勾配 (°)	~20	20~30	30~40	40~50	50~
降雨集中度評価	1	2	3	4	5

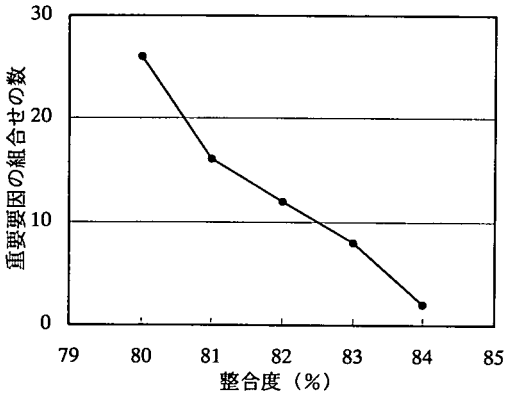


図-8 整合度と重要要因の組合せ数の推移

(2) 土石流発生・非発生の判定に重要な地形要因

どのような地形が土石流発生に関与しているのかを明確にするため、全ての地形要因を用いて分析を行った。なお整合度の要求水準は80%以上、エラーレベルは0で実施した。エラーレベルは榊原ら¹²⁾に詳細な記述がある。本研究ではルール抽出を行う際に確信度が1となるルールの抽出を行うために、重要要因抽出時から整合データ中に矛盾データを含まないようにエラーレベルを0として分析を行った。

表-7 土石流発生、非発生を最も効率良く判別する重要要因の組合せの抽出(整合度上位5位まで)

地形要因								整合度
3	6	7	8	9	10	12	14	
○	○		○		○	○		84.97%
○		○	○		○	○		84.48%
	○		○	○	○		○	83.99%
○	○		○	○	○			83.82%
○	○		○	○		○		83.50%

要求水準を満たす要因の組合せは全部で26通り存在した。図-8には80%から1%ごとにその組合せの累積された組合せの数を示す。例えば要求水準81%を満たす組合せは16通りであることを示している。図から整合度が高くなるにつれて急激に組合せ数が減少しており、84%を達成している組合せは2通りであった。分析結果を表-7に示す。80%以上の組合せは26組と非常に多いことから、ここでは5要因の組合せのうち整合度が上位5位まで表示している。このうち、最も整合度の高い組合せは、最急溪床勾配(3)、流域長(6)、流域形状比(8)、溪流方位(10)、流域最大傾斜(12)の組合せであり、84.97%となった。以下の検討には最も整合度の高

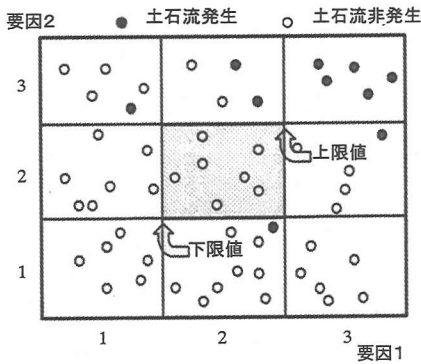


図-9 ルールの一例

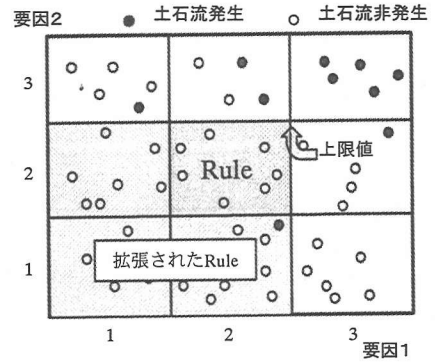


図-10 非発生ルールの拡張

いこの組合せを用いて検討を行う。

このように、どの地形要因が土石流発生、非発生に関与しているかを判断する手法としてニューラルネットワーク（以下 NN と記す）による感度分析が報告されている。荒木ら⁸⁾は、土石流の発生予測について降雨の感度分析を行った。このときの NN モデルは地形と降雨の要因を組合せた構造となっており、地形の要因としては代表的な溪流の地形要因を用いている。このように、この研究では、NN モデルの構造が降雨要因を含めて検討を行っており、地形要因のみでの検討は行われていない。また、地形要因に関しては代表的な溪流のデータを用いるか、平均値、最頻値などの地形データを用いるかについても議論しなければならないが、それらの検討は行われていない。

菊池ら¹³⁾では切土のり面に対して NN のシナプスウェイトの差を用いて各斜面要因の特徴分析を行っている。この手法では、最もよく効いている地形要因の抽出は可能であるが、地形要因の組合せの精度に関しては検討できない。さらに、重要と思われる地形要因の組合せを削除して分析を行い、その精度低下の大きさを削除した要因の組合せの重要度として与える手法も行っているが、全ての組合せに対して検討を行う必要があり多大な労力と時間が必要となる。したがって、重要要因の組合せの抽出に関しては、ラフ集合を用いることは非常に簡潔であり有効な手法であると言える。

5. 土石流ルールの設定

(1) 土石流ルールの設定の手順

土石流の発生・非発生の事象の境界を明確にする重要要因の組合せから、土石流ルール抽出を行う。

ラフ集合における整合データは、カテゴリ別に分割された格子内のデータがすべて同種（土石流発生あるいは非発生）のデータである。これらの整合データを含む格子は全体集合の中で分散しており、一つ一つが有意なルールであるとは言い難い。その一例を図-9 に示す。要因1、要因2はカテゴリが大きくなるにつれて危険度が大きくなる要因とする。すべて非発生データである灰色の格子が土石流ルールとして抽出されたとする。この土石流ルールよりカテゴリのより小さい部分集合（要因1のカテゴリが1で要因2のカテゴリが1である集合等）においては、灰色の部分集合よりも安定な集合であると推測される。ところが、より安定と推測されるカテゴリ値の小さい部分集合が安定であるか否かはこの土石流ルールから判断できない。これは自然災害を対象としている土石流ルールとしては問題があると思われる。

そこで、本研究では土石流発生・非発生に有用な土石流ルールを設定するために、ラフ集合において抽出されるルールから土石流ルールを新しく定義することとした。上記したように、自然現象に矛盾のない土石流ルールを設定するには、例えば土石流非発生のルールであれば、より安全と思われる部分集合をすべて含めることで矛盾をなくすることができる。その結果、1つの土石流ルールに含まれる溪流数が増加することから、サポートが増加し、より実用的な土石流ルールを設定することができる。1つの土石流ルールに含まれる領域を土石流ルールの適応領域と呼び、この適応領域を拡張する新しい土石流ルールを定義する。以下に土石流ルールの設定手順を示す。

1. 最も整合度の高い重要要因の組合せに対して土石流ルールを設定する。また、整合度が同じである場合には、整合度の高い組合せで多く用い

表-8 長崎市の地形要因の組合せから抽出される土石流非発生ルール (確信度 1, サポート 0.02 以上まで)

ルール	地形要因番号					判定	確信度	サポート
	3	6	8	10	12			
RuleA	*	2	2	*	1	1	1	0.049
RuleB	*	2	*	1	2	1	1	0.041
RuleC	*	2	*	3	2	1	1	0.034
RuleD	2	2	*	*	1	1	1	0.029
RuleE	*	2	2	3	*	1	1	0.026
RuleF	2	2	*	*	2	1	1	0.026
RuleG	3	1	*	*	*	1	1	0.024
RuleH	*	2	*	3	3	1	1	0.024
RuleI	4	*	2	1	*	1	1	0.024
RuleJ	2	2	*	4	*	1	1	0.024
RuleK	4	2	*	3	*	1	1	0.024
RuleL	4	2	2	*	*	1	1	0.023
RuleM	*	1	5	*	*	1	1	0.021
RuleN	2	*	3	*	2	1	1	0.021
RuleO	5	*	*	*	1	1	1	0.020
RuleP	*	1	*	1	*	1	1	0.020
RuleQ	3	2	*	2	*	1	1	0.020
RuleR	2	2	*	1	*	1	1	0.020

られている要因を多く含む組合せから土石流ルール抽出を行う。抽出するルールは確信度が 1 となる土石流ルールとする。

- 抽出された土石流ルールにおいて、発生ルールの場合はその格子が土石流ルールを構成する各カテゴリーの下限に、非発生ルールの場合には上限となるように、それ以下のカテゴリーをすべて含有する範囲をとる。非発生の例を図-10 に示す。図の例では非発生ルールであるからその上限値以下のカテゴリーをすべて含むとする。ただし、要因が離散データである場合（例えば 渓流方位等）は、拡張されない。
 - その範囲で確信度とサポートを計算する。
- 以上の設定手法より、次節で土石流ルールを抽出し、その検討を行う。

(2) 土石流ルールの抽出

長崎市で同様の豪雨が起きた場合、土石流が発生する渓流あるいは非発生である渓流を明確にするために、全渓流 (612 渓流) を使用して土石流ルール抽出を行った。地形要因を使用して土石流ルールを抽出する重要要因の組合せについては表-7 に示した整合度が 84.97% と最も高い、最急渓床勾配(3)、

流域長(6)、流域形状比(8)、渓流方位(10)、流域最大傾斜(12)の 5 要因の組合せで行った。上記組合せから得られた土石流ルール数は計 213 個認められた。表-8 には 213 個の土石流ルールの内、確信度が 1、サポートが 0.02 以上の土石流ルールのみを示す。判定については、1 が安定、0 が土石流発生時の土石流ルールを示している。表中に示された土石流ルールは全て非発生 (判定 1) の土石流ルールとなり、発生時の土石流ルールは 1 つも得られなかった。発生時の土石流ルールは確信度 1 の土石流ルールは抽出されたものの、サポートに関しては、最大の土石流ルールで 0.008 となり、表-8 に示される 0.02 には届かなかったため、表示されなかった。発生時の土石流ルールが得られない原因として、以下の理由が挙げられる。

- 1) 発生時の渓流数が非発生時の渓流数と比較して少ない (発生: 非発生=150:462) ことから、どのカテゴリーにも少なくとも非発生渓流が入る確率が高い。1 つの非発生渓流が入ってしまうと、そのカテゴリーは発生時の土石流ルールにはなれない。したがって、渓流数の差以上に発生時のルールの抽出される数が非常に少ない。

表-9 新しく定義したルール抽出法に基づいて抽出された土石流非発生ルール (確信度 0.90 以上, サポート 0.05 以上まで)

ルール	地形要因番号					判定	ルール満足数	整合データ数	矛盾データ数	確信度	サポート
	3	6	8	10	12						
Rule1	*	2	*	3	3	1	51	51	0	1	0.083
Rule2	2	2	*	*	2	1	42	42	0	1	0.069
Rule3	*	2	*	3	2	1	35	35	0	1	0.057
Rule4	*	2	2	*	1	1	31	31	0	1	0.051
Rule5	*	2	2	*	3	1	58	57	1	0.983	0.095
Rule6	*	2	*	1	2	1	55	54	1	0.982	0.090
Rule7	3	2	*	*	3	1	121	118	3	0.975	0.198
Rule8	3	3	2	*	1	1	34	33	1	0.971	0.056
Rule9	2	*	3	*	2	1	31	30	1	0.968	0.051
Rule10	4	2	*	3	*	1	54	51	3	0.944	0.088
Rule11	4	2	*	*	3	1	160	151	9	0.944	0.261
Rule12	3	3	3	*	3	1	70	66	4	0.943	0.114
Rule13	4	2	2	*	*	1	56	52	4	0.929	0.092
Rule14	*	5	*	1	2	1	95	88	7	0.926	0.155
Rule15	*	2	4	*	3	1	40	37	3	0.925	0.065
Rule16	*	4	*	1	2	1	88	81	7	0.921	0.144
Rule17	4	*	*	1	1	1	37	34	3	0.919	0.061
Rule18	3	3	2	*	3	1	59	54	5	0.915	0.096
Rule19	4	3	2	*	3	1	78	71	7	0.910	0.128
Rule20	3	5	3	*	3	1	95	86	9	0.905	0.155
Rule21	*	5	2	*	2	1	91	82	9	0.901	0.149
Rule22	5	*	2	*	2	1	91	82	9	0.901	0.149
Rule23	*	2	4	*	5	1	50	45	5	0.900	0.082

- 2) 本研究ではエラーレベルを考慮していない (=0) ことから, カテゴリーに1つでも矛盾データを含むと土石流ルールとして抽出されない。
- 3) 土石流ルールとして抽出されても, 適応領域を拡張することで, 多数ある非発生溪流を含み, 確信度が急激に下がることから精度の良い発生土石流ルールは抽出されない。
- 4) 最も危険と思われるカテゴリーでも非発生溪流が含まれる場合が多々ある。

1) ~4) より, 土石流発生ルールの抽出は非常に難解であると思われる。

表-8 には得られた 213 個の土石流ルールの中でサポートの高い土石流ルールのみを示しているにもかかわらず, 最もサポートの高い土石流ルール (ruleA) でさえ, 溪流全体の 4.9%しか適応できない。ruleA は流域長(6)のカテゴリーが 2, 流域形状比(8)のカテゴリーが 2, 流域最大傾斜(12)のカテゴリーが 1 であれば, その溪流が長崎災害時に経験

した同じ雨量を再び経験しても土石流は発生しないということを示している。しかし, この土石流ルールは重要要因の 1 つである流域長(6)を見るとカテゴリー2 に含まれる溪流は安全であるが, その溪流より短い溪流では安全であるかどうか不明である。したがって, 前節で検討した手順により土石流ルールの拡張を行い, 新しい土石流ルールの設定を行う。

前節に示された土石流ルール抽出法にしたがって, 重要要因の組合せから有用な土石流ルールを抽出する。表-9 には抽出された全 213 個の土石流ルールのうち, 確信度 0.90 以上でサポートが 0.05 以上の土石流ルールを示し, それ以下の土石流ルールは省略した。ここで, 確信度が 0.90 以上でサポートが 0.05 以上の土石流ルールのみを示しているが, この基準で土石流ルールを定義しているわけではない。ルールを採用する場合には, 最も高い確信度とサポートを得ることが重要であり, どちらをどの程度重要視するかは適用する場合によって, 変化すると思われる。表-9 より領域を拡張したため, 確信度が 1 であった土石流ルールが若干確信度を下げる結果と

なった。これは安全側に拡張したとはいえ、矛盾するデータも若干含まれていることが伺える。しかしながら、確信度と比較して、サポートは数倍もの上昇率を示し、サポートが0.1を越える土石流ルールも珍しくない。このように土石流ルールの適応領域を拡張することで、ラフ集合で定義されるルールと比較してサポートを格段に増加させることができ、適応領域が、あるカテゴリだけでなく、あるカテゴリ以下のカテゴリすべてを含むことで、土石流という自然現象に対して、矛盾のない土石流ルールが設定されたといえる。

例えば Rule1 の場合、流域長(6)はカテゴリ2 (0.3km) 以下、溪流方位(10)はカテゴリ3 (南向き)、流域最大傾斜(12)はカテゴリ3 (45°) 以下の溪流が 8.3%あり、それらは長崎災害時と同程度の降雨に対しても極めて土石流が発生しにくいことが示されている。Rule7 の場合、最急溪床勾配(3)はカテゴリ3 (20°) 以下、流域長(6)はカテゴリ2 (0.3km) 以下、流域最大傾斜(12)はカテゴリ3 (45°) 以下の溪流は 19.8%あり、それらは長崎災害時と同程度の降雨を受けても、97.5%の確率で土石流が発生しないことを意味している。

また、発生した土石流ルールについても同様に適応領域の拡張を行ったが表-9 に示される条件 (確信度 0.90 以上、サポート 0.05 以上) を満たすルールは存在しなかった。発生した土石流ルールでは確信度 0.9 以上のルールで、サポートの最も大きいもので 0.013 であった。

6. 土石流ルールの適用

(1) ルールの組合せによる非発生溪流の抽出

前章までに抽出した土石流ルールによりサポートが 0.1 を越える土石流ルールが多く抽出できたが、それでも1つの土石流ルールでは10%ほどしかサポートできていないことから単独で使用するには、まだサポートが小さいといえる。したがって、本節では、サポートは小さいものの土石流ルール数は非常に多いことから、これらの土石流ルールを組合せて、より多くの溪流に対してサポートできる手法を検討する。

土石流ルールを組合せることによってサポートは上昇するが、その達成すべき目的によって、それぞれの土石流ルールの組合せかたが当然変化する。例えば、確信度が1で、より多くの非発生溪流を抽出することが目的であれば、表-9 にある rule1~rule4 を組合せること (rule3 は rule1 に含まれるこ

とから、3 個の土石流ルールでよい) で最も多くの溪流を抽出することが出来る。

したがって、確信度の高い土石流ルールから順に組合せることが当然考えられるが、サポートの小さい土石流ルールを多数組合せても、それ程の効果は得られない。サポートの大きな土石流ルールを用い、できる限り組合せる土石流ルール数を減少させることが望ましい。つまり、要求水準を満たし、さらに最も少ない土石流ルール数を用いることが、効率的な抽出方法であると言える。

本節では、表-9 で単独の土石流ルールで最もサポートの高い値を示した Rule11 (確信度 0.944, サポート 0.261) を要求水準として用いる。そして単独の土石流ルールよりも、土石流ルールを組合せることで、さらに多くの溪流を抽出できることを示すことを目的として、土石流ルールの組合せかたの手法を以下に示す。

- 1) 確信度 1 でサポートが最大の土石流ルールを抽出し、該当する非発生溪流を削除する (表-9, Rule1)。確信度が 1 の土石流ルールは他に 3 個あるがサポートが小さいことから、抽出できる溪流数が少なく効率的でないため、ここではサポートが最大の土石流ルールのみを用いた。
- 2) 1) で抽出された溪流を削除した残りの溪流で、土石流ルールを適用させたときに、要求水準である確信度 0.944 を越える土石流ルールで最もサポートの高い土石流ルールを適用する (表-9, Rule7)。
- 3) この作業を繰り返して、必要数の溪流を抽出する。または残存させる (Rule14 以下を繰り返す)。

この手法は最初に抽出した土石流ルールを用いて、段階的に安定な溪流を抽出していることから、2 回目以降からラフ集合で扱う母集団が変化するため、確信度とサポートが表-9 の値から変化する。したがって、段階毎に確信度、サポートを分析し、最適な土石流ルールを用いる必要がある。

1)~3) より、抽出される溪流数を表-10 に示す。3 段階の土石流ルールで非発生溪流として抽出された溪流の精度は 95.5%で、抽出される溪流数は全溪流数の 33.0%となり、Rule11 のみで抽出できる非発生溪流の精度及び溪流数を越える結果となった。

また、溪流数についての要求水準も当然考えられる。全 612 溪流の内、調査の必要な溪流を 400 溪流以下に抑え、なおかつ確信度 90%を保つように抽出したいという場合も考えられる。その場合組合せを上記 1)~3) より考慮すると表-10 の Rule20 までで、

表-10 数個の土石流ルールを用いた極めて土石流が発生する可能性の低い溪流の抽出

	地形要因番号					判定	削除される溪流			残りの溪流	削除した非発生溪流数 削除した溪流数	削除した溪流数 全溪流数 (612)
	3	6	8	10	12		全体	非発生 溪流数	発生 溪流数			
溪流の分布数										612		
Rule1	*	2	*	3	3	1	51	51	0	561	1	0.083
Rule7	3	2	*	*	3	1	144	141	3	468	0.979	0.235
Rule14	*	5	*	1	2	1	202	193	9	410	0.955	0.330
Rule20	3	5	3	*	3	1	247	230	17	365	0.931	0.404

表-11 過去の災害事例

災害箇所	災害時期	主な岩種
長崎県 長崎市	1982.7	安山岩
香川県 小豆島	1974.7	安山岩・花崗岩
広島県 広島市	1999.6	花崗岩
	加計町	1988.7

表-12 他地域で検討した土石流ルール

	地形要因番号					判定	確信度	サポート
	3	6	8	10	12			
Case1	3	2	*	*	3	1	0.975	0.198
Case2	5	4	*	*	3	1	0.849	0.618

全溪流の40.4% (247 溪流) が削除され、残りの溪流が365 溪流となることから、表-10 に示される4つのルールを用いればよい。

このように土石流ルールを単独で用いることも有効ではあるが、上記のように複数の土石流ルールを用いることによって、より高度な要求精度を満足することができる。この作業がすべて室内作業であり、現場での調査が不要であることから、溪流数の絞込みに土石流ルールを用いることは非常に有用であるといえる。

(2) 他地域への適用

本研究において長崎災害をもとに土石流ルールの設定を行った結果、有用な土石流ルールが抽出された。しかしながら、この土石流ルールは長崎市内で設定されたものであるため、他地域にもこの土石流ルールが有効に活用できるかは不明である。他地域においてこの土石流ルールが適用できるようであるならば、地域ごとに設定することなく、この結果をそのまま引用することができる。

本節では長崎市以外の地域で発生した土砂災害に対して、第5章で得られた土石流ルールが活用できるか検討を行った。第5章で得られた適用領域を拡張したルールが適用できると前節の土石流ルールの組合せも可能となることから、ここでは元の土石流ルールに対して検討を行った。

ここで用いた過去の災害事例について表-11 に示す。本節で採用した土石流ルールを表-12 に示す。本節では、確信度が非常に高い(0.95 以上) 土石流ルールのうちで最もサポートの大きいRule7(Case1)

と確信度0.8 以上でサポートが最も高い土石流ルール(Case2) を用いて他地域において確信度とサポートを求めた。また、Case2 の土石流ルールについては確信度が0.9 以下であるので表-9 には記載されていない。

他地域に適用した結果を表-13、表-14 に示す。Case1 では扱った土石流ルールのサポートが比較的小さいため他地域ではほとんど使用できない。広島市のみで使用できる程度である。Case2 では扱った土石流ルールのサポートが大きいため、他地域でも使用は可能である。広島市や加計町では、サポートが10~20%下がっているものの、確信度は大きくなっており、この土石流ルールの適用は可能であると思われる。しかしながら、小豆島の花崗岩地域においては極端に確信度が低下している。また、小豆島の安山岩地域でも確信度は上昇しているがサポートがかなり下がっている。このように、長崎市、広島市、加計町と小豆島では土石流ルールの確信度およびサポートが全く異なる結果となった。これは各地域の地形が異なっているものと考えられる。

ここで、長崎地域と他地域の地形について検討を行う。図-11 に各地域の全ての溪流における流域最大傾斜の頻度分布を示す。この図から、小豆島の流域最大傾斜は長崎市と比較して、非常に緩い傾向を示すことが分かった。このように土石流ルールを全く異なる地形を有する地域に適用するとカテゴリ分布に大きな偏りが生じるため、ほとんど分類できない現象が起きることから、土石流ルールの適用は困難である。

また、長崎市の全溪流数が612 溪流であるのに対

表-13 Case1の結果

地域	岩種	全溪流数	ルール満足 溪流数	整合データ 数	矛盾データ 数	確信度	サポート	
長崎県	長崎市	安山岩類	612	121	118	3	0.975	0.198
香川県	小豆島	安山岩	40	0	0	0	-	0.000
		花崗岩	137	6	4	2	0.667	0.044
広島県	広島市	花崗岩	1235	149	144	5	0.966	0.121
	加計町	花崗岩	92	1	1	0	1.000	0.011

表-14 Case2の結果

地域	岩種	全溪流数	ルール満足 溪流数	整合データ 数	矛盾データ 数	確信度	サポート	
長崎県	長崎市	安山岩類	612	378	321	57	0.849	0.618
香川県	小豆島	安山岩	40	10	10	0	1.000	0.250
		花崗岩	137	94	52	42	0.553	0.686
広島県	広島市	花崗岩	1235	635	596	39	0.939	0.514
	加計町	花崗岩	92	40	39	1	0.975	0.435

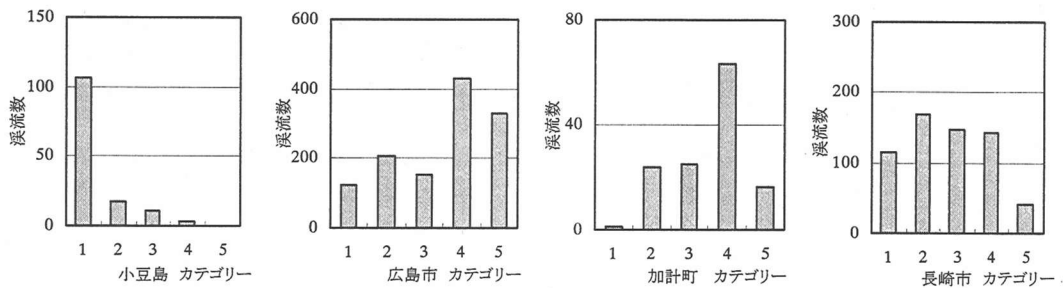


図-11 各地域の流域最大傾斜(12)の頻度分布

して、小豆島の全溪流数は177と溪流数が少ないため、このような確信度とサポートの結果となると考えることもできるが、加計町の場合を考慮すると、全溪流数が92溪流であるにもかかわらず、長崎市の土石流ルールが適応できている。したがって、単に溪流数が少ないことによる影響と考えることも困難である。加計町の場合もカテゴリー1にはほとんど溪流が含まれていないが、他の4要因には特に溪流数が少ないということもないことから、サポートが減少している程度の影響である。小豆島、加計町と比較して広島市では図中のカテゴリー分布がそれほど偏りなく、どのカテゴリーにも多数の溪流が存在している。つまり、広島市は長崎市と類似した溪流の地形形状を示していることがこの図から読み取れる。

このことより、他地域に土石流ルールを適用するには、各要因でカテゴリーごとの頻度分布が類似し

ている地域でないと、各カテゴリーに含まれる溪流数に偏りができ、土石流ルールのサポートが下がることから、土石流ルールの適用は難しいと考えられる。

7. おわりに

本研究では長崎市周辺の地形形状より、土石流発生・非発生に関わる重要要因の抽出および土石流ルールを定義することが出来た。土石流ルールに関しては、ラフ集合で定義されているルールを、土石流に適応できるようにルールの領域を拡張した。また、これらの土石流ルールから、極めて土石流が発生しにくい溪流の抽出を試み有用性を検討した。以下に結論を列挙する。

- 1) 長崎市周辺地域において土石流発生・非発生を

分類する重要な地形要因は、最急溪床勾配・流域長・流域形状比・谷深比・溪流方位・流域最大傾斜である。

- 2) ラフ集合で定義されたルールの領域を拡張することで、土石流に適合したルールを定義することが出来る。
- 3) 土石流ルールを数個用いることで土石流の発生しにくい溪流の抽出に利用できる。
- 4) 長崎市で抽出された土石流ルールを他地域に適用させたところ、同様の地形を保有している地域(広島市)では適用できるが、全く異なる地域(小豆島)では、カテゴリーに偏りが出来るため、適用は難しい。

なお、本研究の結果は、限られた地域に対して土石流ルールを設定したため、地形要因には、その地域の特徴が表れている。また、ラフ集合のカテゴリー分類についても、その地域で最適と思われる分類を行っている。従って他地域で適用するには、6.で検討されているように、長崎市と同様の地形であるか検討を行い、小豆島のように全く異なる地形を呈しているような地域であれば、再度該当地域において分析を行う必要がある。

参考文献

- 1) 建設省河川局砂防部：土砂災害に関する警報の発令と避難の指示のための降雨量設定指針(案)，1984。
- 2) 建設省土木研究所砂防部砂防研究室：土石流の警戒避難基準雨量の設定に関する研究(Ⅱ)，1986。
- 3) 建設省土木研究所砂防部砂防研究室：土石流の警戒避難基準雨量の設定に関する研究(Ⅲ)，1988。
- 4) 瀬尾克美，船先昌継：土砂害(主に土石流的被害)と

降雨量について，新砂防，88，pp.22-28，1974。

- 5) 瀬尾克美，横部幸裕：土砂害(主に土石流的被害)と降雨量について(第2報)，新砂防，108，pp.14-18，1978。
- 6) 瀬尾克美，五代均，原義文，塩島由道：土石流警戒・避難基準としての降雨指標について，新砂防，Vol.38，No.2，pp.16-21，1985。
- 7) 荒木義則，古川浩平，松永悟，小笠原貴人，石川芳治，水山高久：土石流危険溪流における地形特性を考慮した土砂崩壊の発生限界線の設定に関する研究，土木学会論文集，No.574/VI-36，pp.33-47，1997。
- 8) 荒木義則，古川浩平，松井範明，大木戸孝也，石川芳治，水山高久：ニューラルネットワークを用いた土石流危険溪流における土砂崩壊のリアルタイム予測に関する研究，土木学会論文集，No.581/VI-37，pp.107-121，1997。
- 9) Pawlek, Z. : Rough Sets, International Journal of Computer and Information Sciences, Vol.11, pp.341-356, 1982.
- 10) 長崎県土木部河川砂防課：昭和55年度土石流危険溪流および危険区域調査，1980。
- 11) 長崎県土木部砂防室：土石流危険溪流調査表，1990。
- 12) 榊原弘之，倉本和正，菊池英明，中山弘隆，鉄賀博己，古川浩平：ラフ集合を用いたデータマイニングによるがけ崩れ発生要因の抽出に関する研究，土木学会論文集，No.658/VI-48，pp.221-229，2000。
- 13) 菊池英明，古川浩平，大前伸司，奥園誠之，西岡勲：ニューラルネットワークによる切土のり面の豪雨時のリアルタイム崩壊予測と予測に用いる降雨要因に関する研究，土木学会論文集，No.616/VI-42，pp.63-76，1999。

(2001.6.27 受付)

DEBRIS FLOW RULE BASED ON ROUGH SET THEORY EXTRACTION FROM DISASTER CAUSED BY THE NAGASAKI HEAVY RAIN IN JULY, 1982

Takeharu SATO, Yasunori KAWANO, Yoshinori ARAKI,
Hirotaka NAKAYAMA, Takahisa MIZUYAMA and Kohei FURUKAWA

Record concentrated heavy rain hit the Nagasaki Prefecture in July, 1982. It caused a large number of debris flows. It is considered that rainfall makes a prediction of debris flow. Rainfall measures in various regions. However, there are cases where it does not cause debris flows. Therefore, we have to study the natural feature of the region to predict debris flow, since it is not clear to distinguish which torrent is dangerous. This study is clear that essential factor which extracts from natural feature data causes debris flow. We have tried to expand area of rule based on rough set theory. Finally, we have defined new rule and examined validity.