

危険度ランクに対応した土石流の発生・非発生ルールと
防災事業優先箇所特定への適用に関する研究

Rules for the occurrence and non-occurrence of debris flow corresponding
to the risk level and its application to the designation of priority areas for
disaster prevention measures

2015年9月

小林 央宜

Kobayashi Hiroki

危険度ランクに対応した土石流の発生・非発生ルールと 防災事業優先箇所特定への適用に関する研究

学位論文の要旨

我が国における持続的な社会基盤の確立のためには、様々なリスクへの対応が不可欠である。特に、毎年のように発生し甚大な被害をもたらしている土砂災害への対応は急務である。しかしながら、防災事業の懸命な取り組みはあるものの、依然として被害が後を絶たないのが現状である。この状況の要因は、災害の危険性を有する箇所が莫大な数に上ることにあると考えられる（国内の土砂災害危険箇所数は約 52 万箇所）。これらの箇所に対策工事を行うためには多大な費用と時間が必要になることは想像に難くない。このため、対象となる箇所の危険度や被害規模等を勘案した優先度評価を行い、効率的な事業推進を行うことが必須といえる。

このような自然災害リスクへの対応については、個々の箇所の現況を把握するという作業が行われている段階である。ここから得られる基礎情報は、先に述べたような優先度評価を行うためには非常に有効なものと考えられる。既往の研究でも、同様な基礎情報を分析することでリスク評価を行う試みがなされてきた。例えば、パターン分類手法のひとつであるサポートベクターマシン（以下、SVM）を利用し、土砂災害の危険度を評価する手法が研究され、大きな成果が得られている。しかしながら、既往研究は限られたデータに対する検討のみが行われており、広い地域への適用性検証が必須の課題として指摘されていた。

一方、災害の発生するメカニズムは対象となる場所の環境や地形・地質等の素因的要素の組み合わせで全く異なったものとなることが推測される。これを特定することが出来れば、防災対策の必要となる箇所の絞り込み（あるいは対策実施を後回しにする箇所の特定）が可能となり、効率的な事業推進を実現できるものと考えられる。このテーマについては、地域の基礎情報をラフ集合により分析することで災害発生および非発生のルールを明示化する研究が実施されており、こちらもその有効性が確認されている。しかしながら、現状の成果はあくまで災害発生・非発生ルールのみを求めたものであることが課題として考えられる。災害の恐れがあると考えられる箇所の中でも、その危険度については差があることは前述の危険度評価研究でも明らかにされているが、そこで得られる危険度のランクに対応したルールを特定することが出来れば、「高い危険性を有する箇所」から「安全性の高い箇所」までの条件を段階的に示すことが可能となり、防災事業推進のための有効性を更に高めることが出来るものと考えられる。

本研究では、これら 2 つの大きな課題について検討を行い、それぞれに有効な手法を開発することができた。また実際の事業推進における具体的な活用方法についても検討し、その運用フローの提案も行っている。以下に本論文の研究内容を概説する。

まず、第一段階では SVM による土砂災害の危険度評価とラフ集合による災害発生・非発生ルール of 抽出手法について、これまで対象としていなかった地域（長崎市）のデータへの適用性検証を実施した。一連の検証作業の結果、他地域のデータに対しても既往論文と同等の活用が可能であることが確認できている。しかしながら、ここでの検討で抽出されたルールは災害発生・非発生のルールのみである。これに対し、よりきめ細かな評価を実現するための試みとして、危険度ランク毎のルールの抽出を検討した。

第二段階では同じ長崎市のデータについて SVM を利用することで災害発生の危険度に対応した 4 つのランクに分類し、それぞれのルールを検討することを試みた。ここでの検討結果では、各危険度ランクについて精度の高いルールが得られたものの、ルールで説明できないデータが多量に存在してしまうという課題が残ってしまった。そこで、第三段階の検討では最初に対象データから災害の危険・安全についての典型的な例となる「代表データ」を抽出し、それをを用いて分離超平面を求めることを行った。代表データから得られる分離超平面はその地域の典型的な災害発生・非発生の境界線と言え、そこから得られる危険度ランクも地域の典型的な特性を反映したものになると推定できる。ここで分類されたデータ群から各ランクのルールを求めた結果、より少ないルールで全てのデータを説明しつつ、ルールとしての精度は約 80% の高さを維持することに成功した。また、得られたルールのうち最も危険度が高いものには 10 箇所が該当しており、これこそが対策を最優先すべき箇所として認定することが出来た。逆に、危険度が非常に低いと考えられるルールに該当するものが 160 箇所もあることが判明し、これも事業の効率化のためには非常に有効な情報といえる。

本研究の成果は、社会資本整備や自然災害などあらゆるリスクの評価について適用できる可能性を有しており、将来において広範囲な分野で汎用的に活用されることが期待されるものである。

Rules for the occurrence and non-occurrence of debris flow corresponding to the risk level and its application to the designation of priority areas for disaster prevention measures

Abstract

Responding to various risks is necessary in order to establish a sustainable society and infrastructure in our country. There is an urgent need especially in responding to landslide disasters, which cause tremendous damages almost every year. Although much effort is being put forth into disaster prevention projects, the reality is that damages are still never ceasing. One cause is thought to be the vast number of hazardous places (approximately 520,000 areas with landslide risk in Japan). Countermeasure construction for these sites will certainly require an immense budget and time. For that reason, it can be said that evaluation of priority with consideration for such things as risk and damage scale of applicable sites as well as efficient promotion of projects are essential.

With regards to such responses for natural disaster risks, the current task being conducted involves grasping the present state of each site. Basic information obtained in this way is considered highly useful in conducting priority evaluation as mentioned above. Past studies likewise attempted to conduct risk evaluation by analyzing similar basic information. For instance, the support vector machine (herein referred to as SVM), one of the pattern classification methods, has been utilized in researching methods for evaluating landslide disaster risks, with substantial results. However, past studies have only examined limited data, and the need for testing applicability in broader regions has been identified.

The mechanisms by which disasters occur are presumed to vary greatly depending on the combination of causative factors such as environment, topography and geological features of the site. The ability to identify these would allow to narrow down the sites that require disaster prevention measures (or to identify the sites where implementation of measures can be postponed), helping to realize efficient promotion of projects. Pertaining to this subject matter, studies for clarifying rules for disaster occurrence and non-occurrence have been carried out through analysis of regional basic information by means of rough sets, and their effectiveness has been verified. The problem with the current results, however, is that they are results of only seeking rules for disaster occurrence and non-occurrence. Even in areas with possible disaster risks, the existence of variation in risk levels have been revealed in the risk evaluation studies mentioned above, meaning that if identifying the rules corresponding to such risk evaluation becomes possible, then the graded indication of conditions from “high risk area” to “high safety area” can become possible. It is believed that this would lead to further improvement of effectiveness in promoting disaster prevention projects.

This study examined these two important challenges, and successfully developed effective methods for both challenges. Specific utilization methods in promoting actual projects were also

considered, and its operation flow was suggested as well. The details of this study are summarized below.

For the first stage, with regards to the extraction method of disaster occurrence and non-occurrence rules by means of landslide risk evaluation and rough sets using SVM, applicability to the regional data (Nagasaki city) previously not considered to be applied was verified. A series of verification tasks resulted in confirmation that data from other regions can be utilized to the same degree as those in past papers. However, the rules extracted in this examination apply only to disaster occurrence and non-occurrence. By contrast, an examination was made for extracting rules for each risk ranking in attempting to achieve a more detailed evaluation.

In the second stage, the same data of the Nagasaki city are classified into four ranks corresponding to the disaster occurrence risks by utilizing SVM in an attempt to examine the rules for each rank. These examination results helped to obtain highly precise rules pertaining to the risk rankings, but the problem of excessive data unexplainable by rules remained. Accordingly, the examination in the third stage involved first extracting “representative data,” which is a typical example of disaster risk and safety, from the target data, and then using them to obtain a separating hyperplane. The separating hyperplane obtained from the representative data can be considered to be the boundary between the disaster occurrence and non-occurrence, so the risk rankings obtained from those data can likewise be presumed to reflect the typical properties of that region. As a result of obtaining the rules for each ranking from the classified data sets, we successfully maintained roughly 80% precision for the rules, while explaining all of the data using fewer rules. Among the obtained rules, 10 sites came under the highest risks and could therefore be recognized as the sites for highest priority of countermeasures. It was also revealed that as many as 160 sites corresponded to the rules considered as the lowest risk, and this information could become very useful in optimizing the project.

The results of this study have the potential for application in evaluating various risks including development of social infrastructure and natural disasters, leading to anticipation for their wide utilization in various fields in the future.

目 次

第1章 序論.....	1
1.1 研究の背景と目的.....	1
1.2 危険度評価手法に関する既往研究.....	2
1.3 災害の発生・非発生に関する素因誘因の特定に関する既往研究.....	3
1.4 本論文の構成.....	4
第2章 SVMとラフ集合を用いた災害ルールの評価手法の妥当性検証.....	8
2.1 緒言.....	8
2.2 手法の概要.....	9
2.2.1 SVMの概要.....	9
2.2.2 ラフ集合の概要.....	11
2.3 使用データの概要.....	12
2.4 土石流の発生・非発生ルールの抽出.....	18
2.4.1 SVMによる危険度評価.....	18
2.4.2 代表データの抽出.....	20
2.4.3 ラフ集合によるルール抽出.....	21
2.4.4 検討結果および考察.....	28
2.5 まとめ.....	34
第3章 災害危険度ランクに対応した発生・非発生ルールの明示化.....	36
3.1 緒言.....	36
3.2 手法および使用データの概要.....	37
3.3 災害発生危険度に対応した評価ルールの設定.....	37
3.3.1 SVMによる危険度ランクの設定.....	37
3.3.2 危険度ランクに対応した評価ルールの設定.....	39
3.4 まとめ.....	42
第4章 災害危険度ランクに対応した発生・非発生ルールの汎用性向上の検証.....	43
4.1 緒言.....	43
4.2 災害発生危険度に対応した評価ルールの設定方法.....	44
4.2.1 手法および使用データの概要.....	44
4.2.2 災害発生危険度に対応した評価ルールの設定手順.....	44
4.3 分析結果.....	49
4.3.1 危険度ランクの区分.....	49

4.3.2	危険度ランクに対応した評価ルールの設定.....	53
4.3.3	ルールの抽出結果.....	55
4.4	まとめ.....	61
第5章	今後の防災事業計画への利用に関する一提案.....	63
5.1	緒言.....	63
5.2	ルールに該当する箇所の危険度の確認.....	64
5.3	危険度評価システムの提案.....	66
5.4	まとめ.....	69
第6章	結論.....	70
6.1	本研究のまとめ.....	70
6.2	今後の課題と将来への展望.....	73

謝 辞

第1章 序論

1.1 研究の背景と目的

我が国における持続的な社会基盤の確立のためには、様々なリスクへの対応が不可欠である。特に、自然災害リスクについては、近年の気候状況の変化（集中豪雨の頻発など）から、年々その危険性が高まりを見せている。平成24年7月には九州北部豪雨による災害¹⁾が、平成26年8月には広島で豪雨による災害²⁾が発生し、それぞれ30名、74名と尊い命が失われ甚大な被害が発生したことは記憶に新しい。このような災害から生命・財産を守るための防災事業は、長年にわたり国家の重大課題として取り組まれているものの、厳しい気候・風土の上に立地するわが国では依然として悲惨な被害が後を絶たないのが現状である。この状況の要因の一つは、災害発生危険性の有する箇所数が莫大なものであるという点にあると考えられる。国内の土砂災害危険箇所の総数は約52万箇所と報告³⁾されている。これら全ての箇所を対象に詳細調査および有効な工事を施していくためには多大な費用と時間が必要になることは想像に難くない。そのため、防災対策を必要とする箇所毎の危険度や被害規模等を勘案した上での優先度評価を行い、効率的な事業推進を行うことが必須といえる。

このような中、自然災害リスクへの対応については、まず個々の対象の現況を把握するという作業が行われている段階である。土砂災害に関して言えば、平成13年に施行された土砂災害警戒区域等における土砂災害防止対策の推進に関する法律（以下、土砂災害防止法）に基づく基礎調査が全国で進められているところである。これらの基礎情報は、先に述べたような優先度評価を行うためには非常に有用であり、事業推進に大きく寄与するものと考えられる。これまでの研究でも、同様な基礎情報を分析することでリスク評価を行う試みがされてきた。例えば、パターン分類手法のひとつであるサポートベクターマシン⁴⁾（以下、SVM）を利用し、複雑な条件下で発生する土砂災害の危険度を評価する手法^{5),6)}についての検討がなされており、既に大きな成果が得られている。膨大な箇所を対象とした防災事業の推進に寄与するためには出来るだけ汎用性の高い評価手法を確立することが重要となるが、これについても同手法に関する既往研究^{5),6)}の中で実現の可能性が示されていた。しかしながら、既往研究^{5),6)}では対象としたデータの取得地域が山口県・広島県に限られており、同手法の汎用性を証明するためには他地域で取得されたデータへの適用性検証が必須の課題として指摘されていた。この適用性検証を実施し、SVMを用いた土砂災害の危険度評価手法の妥当性および汎用性の確認ができれば、今後さらに他地域への展開や他の災害形態への応用等、活用範囲の広がりが期待できる。

一方、災害の発生するメカニズムは画一的なものではなく、対象となる場所の環境や地形・地質等の素因的要素の組み合わせで全く異なったものとなることが推測される。これについては、箇所毎の基礎データに対応したリスク発生・非発生実績のデータを整

備し、それらをラフ集合により分析することで分析対象とした地域においてリスクが発現するルール（あるいは発現しないルール）を明示化する研究^{6), 7), 8), 9)}が実施されており、その適用性も検証されている。しかしながら、既往研究^{6), 7), 8), 9)}で、抽出されたルールは、過去の災害の履歴に基づいた「災害の発生するルール」、「災害の発生しないルール」のみであることが課題であった。この課題に対しては、危険度の「ランク」に対応したルールの抽出を行い素因・誘因の条件を危険度ランク毎に具体化することができれば、「高い危険性を有する箇所」から「安全性の高い箇所」までの条件を段階的に示すことができ、防災対策を推進する上でより優先的に進めるべき箇所を特定することが可能となると考えられる。

そこで本研究では、SVMを用いて危険度の評価を実施し、その危険度に対応した土石流災害の評価ルール（素因誘因の条件）を設定することを試みた。その結果として、優先的に事業を実施すべき箇所（事業を後回しにする箇所）を特定する事を目指し、防災事業の効率的な推進に寄与する事を目的として研究に取り組んだ。

1.2 危険度評価手法に関する既往研究

既往の研究事例にみられる土砂災害等の発生・非発生に関する危険度評価手法は、①採点方式、②力学的解析手法、③統計処理方式に大別することができる。

①の採点方式とは、技術者による現地調査を伴うもので、その調査結果に基づき危険度を評価するものである。これについては、砂防^{10),11),12)}や道路¹³⁾等のさまざまな分野で古くから取り入れられ実用化されているものである。道路分野では、安定度調査表¹³⁾と言われるチェックシートに現地調査での評価を記載し、その評価に基づく点数と技術者の判断により総合的に評価を実施している。これらの現地調査については、各箇所の現況を把握する上での重要な基礎資料となるものであるが、最終的な危険度評価は技術者個人の経験や力量などに左右されるところが大きく、その客観性に課題を残している。

②の力学的解析手法とは、地すべり等の円弧すべり計算によるものが代表的である。これは、計算に用いる斜面の物性等の諸条件を専門家が設定し、安定性を評価するものである。計算に当たっては、各基準書^{14), 15)}に物性の一般値等が示されているため参考にすることも可能であるが、詳細な対策等を実施するための検討に用いることが多いため、現地での土質試験結果等が必要である。そのため、個別の箇所に対する検討には非常に有効であるが、対象の危険箇所が膨大で面的な広がり分布している場合の評価には向かないと言える。

③の統計処理方式としては、数量化理論を用いた危険度評価^{16), 17)}や地形解析により要因を分析した上でAHP（階層型分析法）分析を実施する危険度評価¹⁸⁾、SVMを用いた危険度評価^{5), 19)}などが研究されている。

北園ら¹⁶⁾は、斜面崩壊に寄与するとされる素因と崩壊の有無の関係について、数量化Ⅱ類を用いた判別分析を実施し、どの素因が災害発生に影響を与える度合いが大きいか

を特定している。また、岡田ら¹⁷⁾は鉄道盛土を対象とし、数量化 I 類による多変量解析を実施することで、危険度評価基準の作成を試みている。しかしながら、これらの数理化理論を用いた評価では、複数の要因が複雑に絡み合い生じる自然災害に対して、その要因の組合せに関して明確にすることができない点に課題がある。

日本地すべり学会¹⁸⁾では、地震による斜面変動（地すべり、急傾斜地の崩壊）を対象とし、リモートセンシング技術を利用した数値標高モデル（以下、DEM）による地形解析を実施し要因を分析した上で、階層型分析法（以下、AHP 法）による点数付けを実施している。しかしながら、急傾斜地の崩壊では、現象のスケールが小さいことから、地形解析による要因抽出に課題を残している。

大石ら⁵⁾の研究では、高い分離能力を有する SVM を活用することにより、斜面災害における危険度を精度よく、かつ、汎用的に設定できることが示されている。しかしながら、災害発生条件となる要因の状況を説明することができないことが課題として残されていた。

1.3 災害の発生・非発生に関する素因誘因の特定に関する既往研究

頻発する土砂災害に対して、災害の発生した箇所の素因・誘因に関する具体的な特徴を特定できれば、防災事業計画を策定する上で非常に有効であると考えられる。

このような問題に対する既往の取組みの一つとして、山下²⁰⁾は花崗岩地域（広島）の土石流発生特性を地形地質要因（溪床勾配、流域面積、地被条件）に基づき統計的に評価している。しかしながら、本対象地域では、流域面積と被覆条件については、発生・非発生に差異が見られることが確認できたが、要因の組合せに関しては言及できていない。また、小橋ら²¹⁾は土石流の発生条件について、勾配と集水面積、谷次数区分と災害の発生率の関係を整理している。しかしながら、比較的大きな流域を対象としており、小流域に当てはめると別の要因が寄与する部分が大きくなり結果にばらつきが出るという課題が残った。またラフ集合を用いた発生・非発生ルールの抽出による評価^{8), 9)}では、地形地質データや降雨データと災害の発生・非発生の関係を用いて、具体的な規則性（以下、ルール）を抽出することに取り組んでおり、高い精度を確保した汎用性の高いルールの抽出に成功している。しかしながら、ラフ集合では災害の発生や非発生に寄与する要因の組合せを明確にすることはできるが、その危険度に関しては明確にできていないという課題があった。

このような研究状況を受けて取り組まれたのが SVM による危険度評価⁵⁾とラフ集合による災害発生、非発生ルールの抽出法^{8), 9), 19)}の組合せによるルール作成方法に関する研究⁶⁾である。この研究⁶⁾により、SVM の高精度な危険度評価結果で、「災害発生の可能性が高い」と評価された箇所の素因的条件（ルール）を明示することが可能となった。しかしながら、抽出されたルールは「発生」、「非発生」に関する 2 区分のみであり、対象箇所の持つ細分化された危険度に関しては、考慮されていないことが課題であった。

1.4 本論文の構成

本論文では、SVMを用いた土砂災害の発生危険度ランクの設定とそれに基づく評価ルールを設定を研究の柱として検討を行った。以下に本論文の構成と各章の概要を示す、

第1章 序論

第1章では本研究の背景と目的を述べた上で、既往の危険度評価と災害の発生・非発生に関する素因誘因の特定手法の現状を整理し、それらの手法に関する既往の研究成果と問題点を述べると共に、本論文の構成とその概要を述べた。

第2章 SVMとラフ集合を用いた災害ルールの評価手法の妥当性検証

対象地域における典型的な災害発生・非発生事例を代表するデータ（以下、代表データ）をSVMの利用により抽出し、この代表データを用いてラフ集合によるルールを抽出するという方法については、篠崎ら⁹⁾により提案されている。この研究においては、平成11年に発生した広島災害を対象とした検討を実施しており、高い精度と汎用性を有した実用的なルールの抽出に成功している。しかしながら、他の地域の災害事例を対象とした検討は行われておらず、手法としての汎用性は確認されていない。そこで、本章では、竹本ら⁹⁾が研究の対象とした昭和57年の長崎豪雨における土石流災害についてのデータに対して、篠崎ら⁹⁾の提案する手法を適用し、その汎用性についての検証を行った。

その結果、昭和57年の長崎豪雨災害における土石流発生・非発生事例の84.1%をわずか7つのルールで説明できる非常に汎用性の高いルールを得ることができ²²⁾、この手法の汎用性が確認できた。

第3章 災害危険度ランクに対応した発生・非発生ルールの明示化

第2章における篠崎ら⁹⁾の提案する手法に対する適用性検証の結果、新たな対象地域のデータについても要因数やルールを削減しながら高いサポートを有した汎用性の高いルールを得ることができ、SVMとラフ集合の組合せによる手法の優位性が明確となった。

しかしながら、抽出されたルールは「災害の起こるルール」か「災害の起きないルール」かのみを説明するものであったという点に課題を残している。

本来、土石流の災害発生や非発生に関する特徴は、対象箇所の持つ危険度に応じて差異があることが予想される。そのため、危険度に対応した類型的なルールを特定することができれば、優先的に防災事業を実施すべき箇所を容易に選定することが可能となる。そこで、本章ではSVMの基本機能を用いて、対象となるデータ群を①災害発生の可能性の高い危険領域のデータ群、②災害発生の可能性が比較的高いと考えられるデータ群、

③災害発生の可能性が比較的低いと考えられるデータ群, ④災害発生の可能性の低い安全領域のデータ群の4つのグループに区分し, それぞれに対応したルールを検討することを試みた.

その結果, 既往の研究⁹⁾と比べると要因数やルール数は若干増えるものの, 抽出されたルールの精度やサポートについては, 同程度の結果となったことから, 危険度ランクごとのルールを設定することが可能となった.

第4章 災害危険度ランクに対応した発生・非発生ルールの汎用性向上の検証

第3章では, 危険度ランクに対応したルールの抽出方法の検討を実施した. しかしながら, 抽出したルールは, ルールの精度こそ高いが, そのルールで説明できる範囲が非常に狭く汎用性を確保するためには更なる検討が必要となった. これは, 危険度評価の段階で, SVMの優れた判別能力を利用するために, データを災害発生, 非発生の2群に完全分離する分離超平面(危険度評価の基準となる面)を採用したことが原因の一つであると考えられる. 母集団のデータを完全分離させた場合の分離超平面は, 一般の傾向とは異なる特殊な事例(以下, ノイズデータ)を含んだまま判別してしまう事になる. このようなデータを含んだ形でルールの抽出を行うと, 局所的な条件の組合せのルールができることにより, 確信度は向上するが, そのルールの汎用性は低下してしまうことが予測される.

一方, 防災事業を効率的に進めていくためには, 特異事例に引きずられることなく, 対象地域の典型的なルールを抽出し, その後の詳細調査や対策箇所の絞込み, および調査不要箇所の特定を行うことが重要である.

そこで本章では, 代表データを用いることで, 上記のノイズデータに引きずられない分離超平面の構築を行い, その分離超平面を基準とし, 危険度ランクを設定した上でランクごとの汎用的なルールを抽出することに取り組んだ.

その結果, 危険度ランクに対して分離性の良い要因を選別し, それらの要因を多く含む組合せを用いることで, 精度や汎用性の高いルールの抽出が可能であることが分かった. なお, 抽出された17個のルールで対象地域全ての箇所を約8割の精度で説明できることが確認された²³⁾.

第5章 今後の防災事業計画への利用に関する一提案

第4章で抽出された各ルールに該当する箇所の危険度($f(x)$ 値)に注目し, ルール毎の危険度の差異について分析した. その結果, 「危険」, 「比較的危険」に該当するルールの中でも特に危険なルールの特定が可能となった. これにより, そのルールに該当する10箇所が, 対象地域の中で最優先箇所として詳細調査や対策検討を行うべき箇所として抽出できた. 一方, 安全側のルールに関しても同様に, 特に危険性の低いルールを特定できた. 予算や時間に制約のある公共事業においては, 全ての箇所を同列で事業を

進めていくことは不可能であることから、このようなルールに該当する箇所は、調査実施の優先度を下げるなどの対応を取る必要があることを明らかにした。

また、本研究の結果を基に一連の分析手法を活用した防災対策箇所の特定フローを検討し、本研究成果の実用化に向けた提案を行っている。

第6章 結論

本研究の総括であり、各章の研究成果を要約するとともに、今後の課題や将来展望について述べた。

参考文献

- 1) 福岡管区気象台：災害時気象速報 平成24年7月九州北部豪雨，2012.7.31.
- 2) 内閣府（防災担当）：平成26年に発生した主な土砂災害，2014.12.4.
- 3) 国土交通省水管理・国土保全局砂防課：都道府県別土砂災害危険箇所，<http://www.mlit.go.jp/river/sabo/link20.htm>，2015.4.21.
- 4) Cortes, C. and Vapnik, V.: Support vector networks, *Machine Learning*, Vol.20, pp.273-297, 1995.
- 5) 大石博之，小林央宜，尹禮分，田中浩一，中山弘隆，古川浩平：サポートベクターマシンによる対策工効果を考慮した斜面災害危険度の設定，土木学会論文集，Vol.63, No.1, pp.107-118, 2007.
- 6) 篠崎嗣浩，森田真人，大石博之，古川浩平：SVMとラフ集合を用いた土石流の発生・非発生ルールの作成方法に関する研究，土木学会論文集 F, Vol. 65, No.4, pp.448-460, 2009.
- 7) 岡本正男，荒木義則，中山弘隆，古川浩平，水山高久：ラフ集合を用いたデータマイニングによる土砂移動現象の重要要因及びルール抽出に関する研究，砂防学会誌，Vol.54, No.6, pp4-15, 2002.
- 8) 佐藤丈晴，河野泰典，荒木義則，中山弘隆，水山高久，古川浩平：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，土木学会論文集，No.721, VI-57, pp.13-25, 2002.12.
- 9) 竹本大昭，杉原成満，佐藤丈晴，荒木義則，中山弘隆，水山高久，古川浩平：ラフ集合を用いた土石流発生・非発生 Rule の抽出と危険度の設定手法に関する研究，砂防学会誌，Vol.57, No.2, pp.4-15, 2004.
- 10) 建設省河川局砂防部：土石流危険渓流および土石流危険渓流調査要領（案），1999,4.
- 11) 建設省河川局砂防部急傾斜地保全課：急傾斜地崩壊危険箇所点検要領，1999，11.
- 12) 建設省河川局砂防部急傾斜地保全課：地すべり危険箇所調査要領，1996，10.
- 13) 財団法人 道路保全技術センター：道路防災点検の手引き（豪雨・豪雪等），2007，9.

- 14) 国土交通省河川局治水課：貯水池周辺の地すべり調査と対策に関する技術指針（案）・同解説，2009，7.
- 15) 公益社団法人 日本道路協会：道路土工 - 切土工・斜面安定工指針，2009，7.
- 16) 北園芳人，井上敦裕：斜面崩壊の素因に着目した斜面の危険度評価，第4回土砂災害に関するシンポジウム論文集，2008，8.
- 17) 岡田勝也，杉山友康，村石尚，野口達雄：統計的手法による鉄道盛土の降雨災害危険度の評価手法，土木学会論文集，No.448，Ⅲ-19，pp.25-34，1992，6.
- 18) 公益社団法人 日本地すべり学会：類型化に基づく地震による斜面辺土発生危険箇所評価手法の開発 報告書（概要版），2013，2.
- 19) 幸和範，小林央宜，大石博之，杉本博之，飯田毅，古川浩平：SVMを用いた土木設備の補修必要度評価，土木学会論文集 F4，Vol.68，No.2，pp.52-61，2012.
- 20) 山下祐一：花崗岩地域の土石流発生特性，日本応用地質学会 中国四国支部研究発表会，1995.
- 21) 小橋澄治，武居有恒，仲野公章：土石流発生の場合とその条件(2)，砂防学会誌，Vol.30，No.3，pp.1-6，1977，12.
- 22) 杉原成満，小林央宜，川崎雄大，篠崎嗣浩，大石博之，古川浩平：昭和57年7月長崎豪雨における土石流の発生・非発生ルールの抽出，土木学会論文集 F，Vol. 66，No. 4，pp.593-598，2010.
- 23) 小林央宜，杉原成満，大石博之，荒木義則，古川浩平：既存データの活用による災害発生危険に応じた土石流危険溪流の評価ルールの設定，土木学会論文集 F6，Vol.68，No.1，pp.1-12，2012.

第2章 SVMとラフ集合を用いた災害ルールの評価手法の妥当性検証

2.1 緒言

土砂災害からの被害を防ぐためには、ハード対策の実施が効果的であるが、ハード対策はそれを必要とする箇所が多いことに加え、近年の厳しい経済状況により施設整備のための財源確保が難しいために、思うように事業の進展が図れていないという実状がある。そのため、地域に広く散在する土砂災害危険箇所の中から特に危険な箇所や比較的安全な箇所の分布を把握する等のソフト対策による支援を行い、効率的に土砂災害の防止強化を行うことが重要であると考えられる。

このような分野に関する既往の研究としては、その一つとしてラフ集合¹⁾を用いて過去の災害事例から土砂移動現象の発生・非発生に関する規則性（以下、ルール）等の基礎知識を得る方法²⁾が提案されている。ラフ集合を用いたルールの抽出は、箇所毎の地形・地質等に関する情報（以下、要因）と災害発生・非発生の実績の組合せからなるデータベースを用いて行われる。ラフ集合を用いることで災害の発生・非発生という現象を説明することができる要因の組合せを「ルール」として導くことができる。このような特徴は、複雑な条件下で発生するものと考えられる土砂災害をルールとして再現する上では非常に適したものといえる。しかしながら、ラフ集合は膨大なデータベースを用いて簡約化を行う分析手法であるため、局所的かつ特例的な条件下の災害発生・非発生事例を含んだデータに適用した場合、過度に多様なルールが抽出されてしまうという傾向がある。このような特徴は、各斜面固有の災害メカニズムを導き出すためには適しているものの、防災事業計画への活用を考えた時には、対象地域における最も典型的な災害発生条件を把握することの方が肝要であるため、かえって問題となってしまう。よって、現象に対する説明精度を確保しながら高い汎用性を有した実用的なルールを抽出することが重要な課題となる。

この課題を克服するための方法としては、膨大なデータベースの中から、対象地域における典型的な災害発生・非発生事例を代表するデータ（以下、代表データ）をサポートベクターマシン³⁾（以下、SVM）を用いて抽出し、当該代表データのみを用いてラフ集合によりルールを抽出するという方法が篠崎ら⁴⁾により提案されている。その研究では、平成11年6月29日に発生した広島災害を対象とした検討において、ルール数を抑制させるとともに高い精度と汎用性を有した実用的なルールを抽出することに成功している。しかしながら、他の災害事例を対象とした検討は行われておらず、手法としての汎用性は確認されていない。また、手法の有用性に関する検討としては基本的なラフ集合のみを用いた既往研究²⁾との比較は行われているが、より実用的なルールの抽出方法であるルール領域の拡張⁵⁾、⁶⁾およびルールの組合せ⁵⁾の概念を導入した竹本ら⁵⁾の提案する手法との比較は行われていない。そのため、手法の有用性に関して十分な検討が行わ

れているとは言い難い。

そこで、本研究では竹本ら⁵⁾が研究の対象とした昭和57年7月23日の長崎豪雨における土石流災害（以下、S57.7長崎豪雨災害）の事例を用いて、篠崎ら⁴⁾が提案するルール抽出手法の汎用性についての確認を行った。その上で、ルールの精度および汎用性に関して既往研究⁵⁾との比較を行い、手法としての有用性の検証を試みた。

2.2 手法の概要

本研究では、SVMとラフ集合を用いた土石流の発生・非発生ルールの抽出を試みた。その抽出の流れについては、図-2.1に示すフローに基づく。

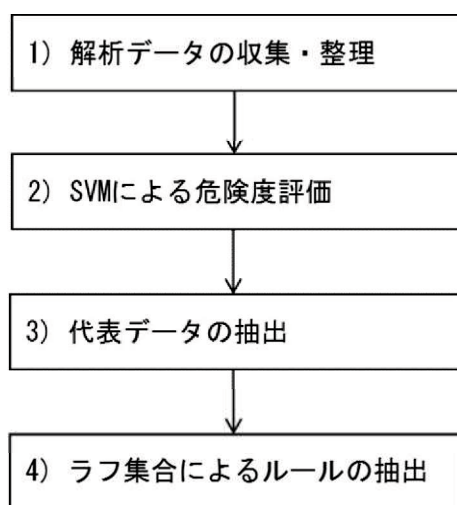


図-2.1 土石流の発生・非発生ルールの抽出フロー

2.2.1 SVMの概要

SVMは1995年にVapnikらによって提案されたパターン分類手法³⁾の一つである（図-2.2）。

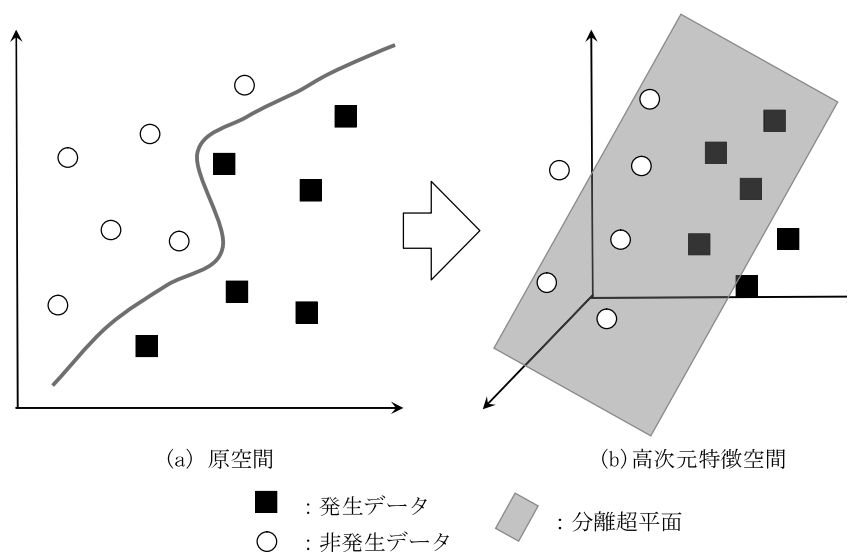


図-2.2 SVMの概念図

SVMは、線形分離不可能なデータ群を非線形関数を用いることにより高次元特徴空間にマッピングし、線形分離可能な状態とすることで明確に分離する機能を持つ。この時、高次元特徴空間でデータを分離する面を分離超平面と呼ぶ。

SVMでは、完全に分離できる場合をハードマージン、一部分離できない場合をソフトマージンという。いずれも線形条件のある2次関数の最大化問題に変換されるが、ソフトマージンはハードマージンを含む概念と考えられるため、前者についてのみ簡単に説明する。ソフトマージン法は、以下の式(1)で示される。

$$\begin{aligned} \underset{w, b, \xi}{\text{minimise}} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\ell} \xi_i \\ & x_i \in R^N, \quad b_i \in R, \xi_i \in R^{\ell} \\ \text{subject to} \quad & y_i (w \cdot x_i + b) \geq 1 - \xi_i \\ & i = 1, \dots, \ell, \quad \text{及び } \xi_i \geq 0, i = 1, \dots, \ell \end{aligned} \quad (1)$$

ここで、 ξ はスラック変数といい、分離面からどれだけ離れているかを表す指標である。 C はスラック変数 ξ に対する重みパラメータで、 C を大きくすれば誤判別の度合いをより最小化することができるため、完全分離に近い判別結果となる。また、 x は発生データ、非発生データであり、 y は教師値（発生の場合：-1、非発生の場合：1）を意味する。 w は分離面の法線ベクトルと呼ばれる変数で、 b はバイアス項と呼ばれる変数である。

ここで、平面による識別が一般に適切であるとは限らないことから、より複雑な識別のために曲面による分離を考える。まず、入力データ x を高次元特徴空間に写像する。ここで、カーネル関数を導入すると、式(1)は結局、以下の双対問題となり、Lagrange乗数 α_i^* に関する最適化問題となる。

$$\begin{aligned} \underset{\alpha}{\text{minimise}} \quad & w(\alpha) = \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ & \alpha \in R^{\ell} \\ \text{subject to} \quad & \sum_{i=1}^{\ell} \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C \quad i = 1, \dots, \ell \end{aligned} \quad (2)$$

また識別関数は式(3)で与えられる。

$$f(x) = \sum_{j \in SV} \alpha_j^* y_j K(x_j, x) + b^* \quad (3)$$

K はカーネル関数である。本研究ではガウシアンカーネルを用いる。 b^* はバイアスと呼ばれる値である。式(2)の最適解 α_i^* は大部分が0になるが、 $\alpha_i^* > 0$ の教師データがサポートベクターと呼ばれる。式(3)の SV はサポートベクターの集合を表す。

なお、本研究ではガウス関数として式(4)を用いる。

$$K(x_i, x_j) = \exp\left[-\frac{\|x_i - x_j\|^2}{2r^2}\right] \quad (4)$$

(i, j = 1, \dots, \ell)

rはデータの影響範囲を示す半径であり、式(1)のCとともに事前に入力が必要なパラメータとなる。

これらの機能を用いて、災害の発生と非発生を判別する分離超平面と各データとの距離f(x)を算出することで危険度を評価することもできる⁷⁾。

本研究ではこの危険度評価手法を用いて、対象箇所における危険度ランクの設定を行う。

2.2.2 ラフ集合の概要

ラフ集合は類似と近似を基本概念としたデータマイニング手法¹⁾の一つである。

図-2.3 にラフ集合の概念図を示す。全体集合を2つの要因を用いて9個の領域に区分しても部分集合は完全に分離されない。このように完全に分離されない集合をラフ集合という。これらの要因によって集合が区分された時、同じ領域内のデータのすべてが同種である領域をルールとして考えることができる(図-2.3の着色部)。この時、ルールの設定に用いる要因の組み合わせの評価においては、式(5)に示す整合度が用いられ、抽出したルールの精度と汎用性の評価においては、式(6)に示す確信度と式(7)に示すサポートが用いられる。

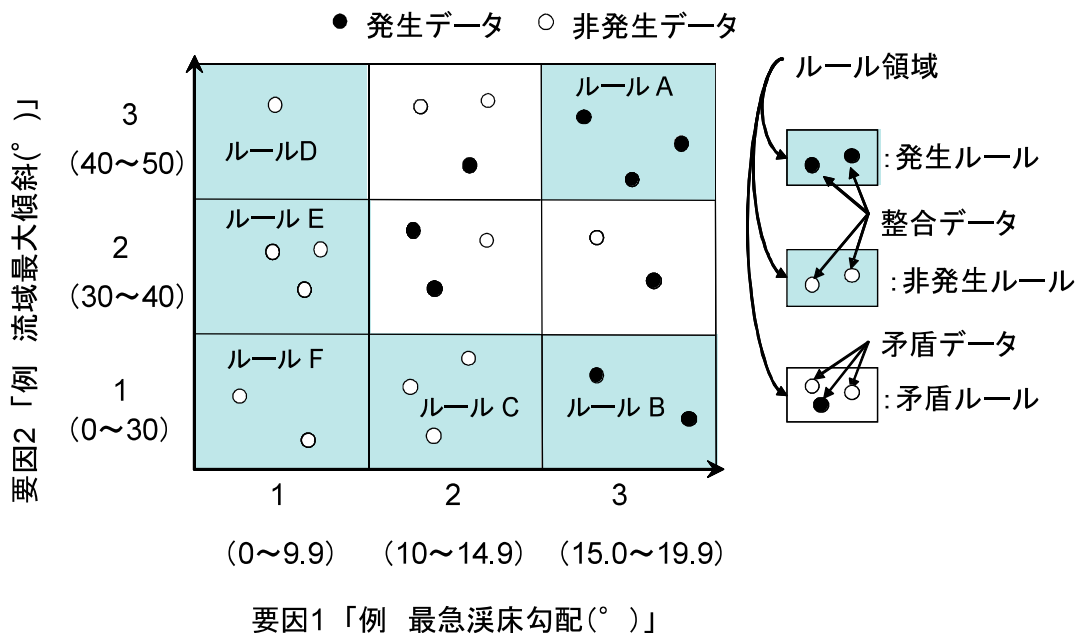


図-2.3 ラフ集合の概念図

$$\text{整合度} = \frac{\text{整合データ数}}{\text{全データ数}} \quad (5)$$

$$\text{確信度} = \frac{\text{ルール領域内の整合データ数}}{\text{ルール領域内の全データ数}} \quad (6)$$

$$\text{サポート} = \frac{\text{ルール領域内の全データ数}}{\text{全データ数}} \quad (7)$$

例えば図-2.3において、要因 1, 2 がそれぞれカテゴリ 3 である場合、これは発生のルールとなり、この時の確信度は 100.0% (3/3)、サポートは 13.6% (3/22) となる。なお、土砂災害などの自然現象に対してラフ集合を用いた場合、抽出した個々のルールで説明できる範囲は少なく、有用なルールとは言い難い場合がしばしば生じる。そのような場合、ルール領域の拡張⁵⁾、⁶⁾およびルールの組合せ⁵⁾を行うことが効果的であるとされている。そのため、ルールの設定にあたっては本研究でも同様の方法を用いる。ルール領域の拡張の概念としては、例えば図-2.3 に示すルール D が非発生ルールである場合、ルール D の領域上端を上限値として、より安全と考えられる部分集合を含む領域を非発生ルールとすることで（カテゴリ区分に「以下」の概念を与える）、ルールに含まれる非発生データ数が増加し、汎用性のあるルールとなる。この場合は「IF 流域最大傾斜=* AND 最急溪床勾配=1 以下 THEN 非発生」という拡張した非発生ルールを抽出することができる。なお、ここで流域最大傾斜について「*」としているのは、ルール D,E,F ともに非発生のルールとなっていることから、最急溪床勾配が 1 以下であれば、流域最大傾斜はどのような値でもよいことを意味している。

2.3 使用データの概要

本研究では、竹本ら⁵⁾が研究対象とした長崎市内の土石流危険溪流 567 溪流（土石流発生 150 溪流、非発生 417 溪流）を検討対象データとした。ここで検討対象とした土石流は、「山腹が崩壊して生じた土石等又は溪流の土石等が水と一体となって流下する自然現象」と定義されている⁸⁾。そのため、土石流の発生においては、①流域内の土砂、②水、③流下しやすさの条件が重要であると考え、以下の要因を選定した。

- ①流域内の土砂：土砂の供給の場と成り得る「斜面要因」と「地質要因」を採用。具体的には、斜面崩壊の起こりやすさ（傾斜）や土砂生産の場（面積）の要因。
- ②水：短期・長期の降雨指標として、「1 時間雨量」と「半減期 72 時間の実効雨量」を採用。
- ③流下しやすさ：溪床勾配などの「溪流要因」と集水のしやすさといった「流域評価」を採用。

ここで、ラフ集合は連続値データを扱うことができないため、連続値データをカテゴリ値に離散化する必要がある。本研究では、既往研究⁵⁾を参考にデータの最大値、最小値等を考慮し、等間隔の5区分を基本としてカテゴリ区分を決定した(表-2.1)。

なお、各要因の概要を以下に示す。

1) 溪流要因

① 水系模様

流水の流下形態を表す。数値が1に近いほど直線的に流下してくるため、危険性が高い。最短溪流長/溪流長で求められる。

② 流域平均勾配

土石の流下のしやすさを表す。勾配が大きいほど流下しやすい。高低差/(主溪流長+上流域長)で求められる。

③ 最急溪床勾配

溪床の移動のしやすさを表す。最急溪床勾配が大きいほど土石流発生時には移動速度が上昇するため危険となる。

④ 主溪流長

流水の流下距離を表す。流域の氾濫開始点から最長流路の1次谷の上流端までの距離を測定する。

⑤ 流域面積

集水の大小、ならびに流域の規模を表す。氾濫開始点より上流の面積を計測する。

⑥ 流域長

溪流の流下方向を長手方向としたときの長さであり、流域の形状や規模を表す指標の一つである。

⑦ 流域幅

溪流の流下方向を長手方向としたときの幅であり、流域の形状や規模を表す指標の一つである。

⑧ 流域形状比

流域幅/流域長より求め、流域形状の比を表す。

⑨ 谷深比

谷の発達程度を表す。主溪流長/流域長で求められる。1に近いほど流域の長さと同程度の谷地形が発達していることから、危険となる。

⑩ 溪流方位

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。

2) 斜面要因

⑪ 0次谷の数

谷地形ではないが，崩壊の危険性が高く，土石流発生の場合と考えられる。流域内の0次谷の数を計測する。

⑫ 流域最大傾斜

斜面崩壊の発生のしやすさを表す。傾斜が急なほど崩壊が起こりやすく危険となる。

⑬ 源頭部面積

土石流発生の場合の大小を表す。1次谷の上流端を基準点としたときの流域面積を計測する。

⑭ 源頭部最急谷勾配

土石流発生源の傾斜を表す。源頭部内の谷部で高低差30mとなる等高線の間隔の最小値を計測し，逆算により勾配を求める。

3) 流域評価

⑮ 降雨集中度評価

流域形状の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あり，5段階評価は，**図-2.4**を参考にして以下のように行う。なお，土石流危険溪流が0次谷であれば降雨集中度評価は1とする。ここで， a ：流域長， b ：流域幅， l ：主溪流長として，以下のフローで評価を実施する。

- ① b/a の値が 0.5 以上・・・②へ
0.5 未満・・・③へ
- ② 流域における0次谷の数が
4 以上 → 降雨集中度 5
3 以下 → 降雨集中度 3
- ③ l/a の値が 0.5 以上・・・④へ
0.5 未満 → 降雨集中度 1
- ④ 流域における0次谷の数が
4 以上 → 降雨集中度 4
3 以下 → 降雨集中度 2

4) 地質要因

⑩ 地質小分類

長崎市の地質は基盤が結晶片岩で構成され、その上を覆うように長崎火山岩類が分布する典型的な火山地形を呈している⁹⁾。地質構造の差は、各溪流の崩壊に対する危険度に影響する要因と考え、表-2.1 に示す代表地質に分類した。

5) 降雨要因

⑪ 時間雨量

短期降雨の指標として、時間雨量を採用した。気象庁、県、市町村等の各雨量観測所の時間雨量から等雨量線を設定し、その等雨量線から国土地理院3次メッシュ(1km×1km)の各格子点における時間雨量を推定し、その平均をとることで算出した¹⁰⁾。

⑫ 実効雨量(半減期72時間)

長期降雨の指標として、過去に降った雨が地中にどの程度残存しているかを示す指標の一つとなる実効雨量(半減期72時間)を採用した。

表-2.1 検討に用いる要因と各要因のカテゴリ区分

要 因	カテゴリ区分								
	1	2	3	4	5	6	7	8	
地形要因	1 水系模様	0.95~0.99	0.90~0.94	~0.90	-	-	-	-	
	2 流域平均勾配(°)	0~14.9	15.0~19.9	20.0~24.9	25.0~29.9	30.0~	-	-	
	3 最急溪床勾配(°)	0~9.9	10.0~14.9	15.0~19.9	20.0~24.9	25.0~	-	-	
	4 主溪流長(km)	0~0.09	0.10~0.19	0.20~0.29	0.30~0.39	0.40~	-	-	
	5 流域面積(km ²)	0~0.049	0.050~0.099	0.100~0.149	0.150~0.199	0.200~	-	-	
	6 流域長(km)	0~0.149	0.150~0.299	0.300~0.449	0.450~0.599	0.600~	-	-	
	7 流域幅(km)	0~0.09	0.10~0.19	0.20~0.29	0.30~0.39	0.40~	-	-	
	8 流域形状比	0~0.29	0.30~0.49	0.50~0.69	0.70~0.89	0.90~	-	-	
	9 谷深比	0~0.19	0.20~0.39	0.40~0.59	0.60~0.79	0.80~	-	-	
	10 溪流方位 (注)	北	南	西	東	-	-	-	
	11 O次谷の数	0個	1個	2個	3個	4個以上	-	-	
	12 流域最大傾斜(°)	0~34.9	35.0~39.9	40.0~44.9	45.0~49.9	50.0~	-	-	
	13 源頭部面積(km ²)	0~0.0049	0.0050~0.0099	0.0100~0.0149	0.0150~0.0199	0.0200~	-	-	
	14 源頭部最急谷勾配(°)	0~19.9	20.0~29.9	30.0~39.9	40.0~49.9	50.0~	-	-	
	15 降雨集中度評価	1	2	3	4	5	-	-	
	16 地質小分類 (注)	結晶片岩類	安山岩質凝灰岩	輝石安山岩	第三紀層	下平閃緑岩類	変朽安山岩	角閃石安山岩	その他
	17 降雨短期	~100	100~115	115~130	130~145	145~	-	-	
	18 降雨長期	~240	240~270	270~300	300~330	330~	-	-	

(注)既往研究⁴⁾において発生率で並べ替えを行った後のカテゴリ区分

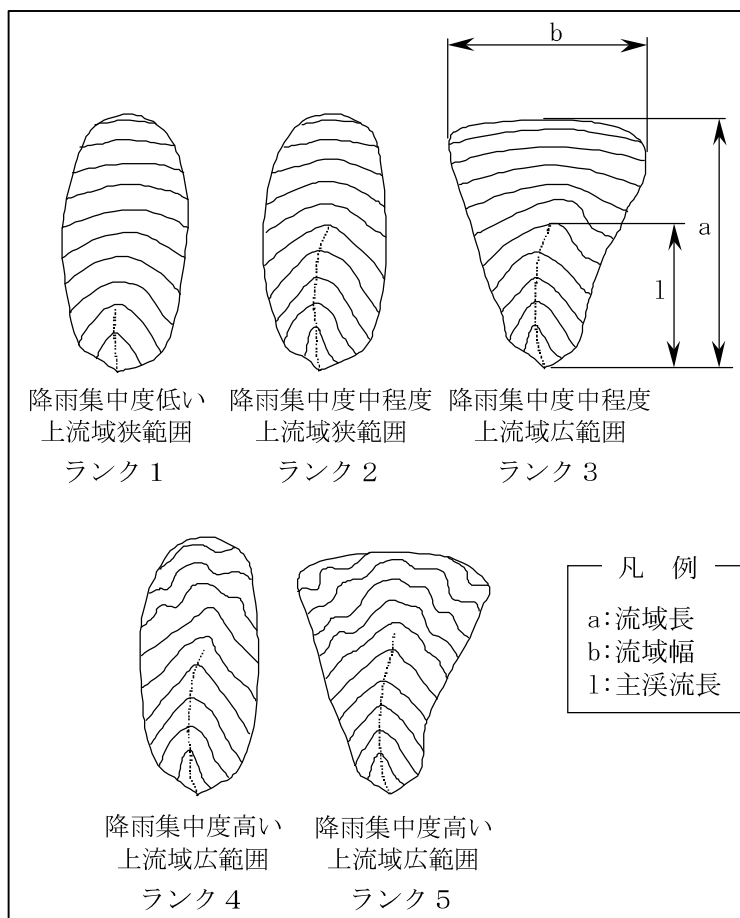


図-2.4 降雨集中度評価

2.4 土石流の発生・非発生ルールの抽出

前述のデータを用い、図-2.1のフロー図に示す手順に従って、S57.7長崎豪雨災害に関する新たな土石流発生・非発生ルールの抽出を行った。また、抽出したルールの精度および汎用性について既往研究⁵⁾との比較を行った。

2.4.1 SVMによる危険度評価

SVMの分析を行う際には、最適な解析を行うためのパラメータスタディが必要である。SVMでは、誤判別の度合いを調整する C とガウシアンカーネルの半径（データの影響度）である r の2つのパラメータを設定する必要がある。本研究では、 $C=5, 10, 50, 100, 200, 300, 400, 500$ の8ケース、 $r=0.1, 0.5, 1, 2, 3, 4, 5$ の7ケースの計56ケースでパラメータスタディを行った。パラメータスタディでは、次式に定義する的中率により各ケースにおいて構築される分離超平面の精度（いかに正確にデータ分類が行われるか）を検証した。

$$\text{的中率} = \frac{\text{的中データ数}}{\text{全データ数}} \quad (8)$$

ここで、SVMによる危険度設定に用いるパラメータは、既往研究⁴⁾におけるパラメータの選定方法を参考に、的中率100.0%のパラメータセットのうち、サポートベクターとなるデータ（ $|f(x)| \leq 1$ となるデータ）数が最も少なくなる $C=500, r=4$ を採用した（表-2.2）。また、当該パラメータを用いて各溪流の危険度評価を行った。

表-2.2 的中率 100.0%となるパラメータにおける $f(x)$ 値カテゴリ毎のデータ分布

パラメーター		$f(x)$ 値カテゴリ毎のデータ個数			
C	r	$f(x) < -1$	$-1 \leq f(x) \leq 0$	$0 < f(x) \leq 1$	$1 < f(x)$
5	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
10	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
50	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
	2	4	146	231	186
100	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
	2	4	146	231	186
200	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
	2	4	146	231	186
	3	25	125	183	234
300	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
	2	4	146	231	186
	3	25	125	183	234
400	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
	2	4	146	231	186
	3	25	125	183	234
500	0.1	0	150	417	0
	0.5	0	150	417	0
	1	0	150	400	17
	2	4	146	231	186
	3	25	125	183	234
	4	32	118	157	260

2.4.2 代表データの抽出

代表データは、分析対象地域における典型的な災害発生・非発生事例を代表するデータであることから、災害発生データとしてはより危険度の高いもの ($f(x)$ 値が-1よりも小さいもの)、非発生データとしてはより安全度の高いもの ($f(x)$ 値が1よりも大きいもの) を抽出する⁴⁾ (図-2.5)。ここで、抽出する代表データの割合は元のデータベースにおける発生・非発生データの割合を乱さない方法が最も有効である⁴⁾。そのため、本研究でも同様の方法を採用した。

上記パラメータに基づく SVM 解析により算出した $f(x)$ 値に着目すると、 $f(x)$ 値が-1 よりも小さくなる発生の代表データ数は32個であった (表-2.2)。そのため、非発生の代表データは、元のデータベースにおける発生・非発生データの割合 (150 : 417) が保持できるように $f(x)$ 値が大きい87個を抽出した。結果として代表データは119個となった。

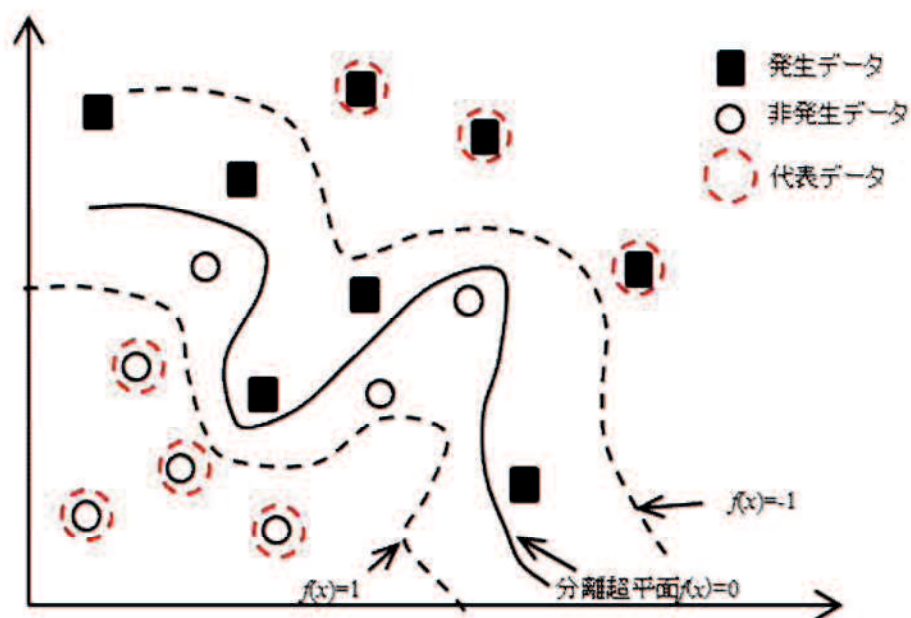


図-2.5 代表データのイメージ図

2.4.3 ラフ集合によるルール抽出

(1) ラフ集合における整合度の設定と最小要因

代表データからルールを抽出するには、ラフ集合を用いる。ここで重要となるのがラフ集合を用いた解析に必要な最小要因数と整合度との関係である。一般的に整合度を高くすると要因数が増えルールが煩雑になり、要因数を減らすとルールの簡素化は可能となる一方で整合度が低下する傾向を示すため、解析時には適切な整合度を設定することが必要となる。

本研究では土砂災害の発生と非発生を典型的に分離する代表データのみを用いてルールの抽出を行うため、整合度 100.0%を満足する要因の組合せのうち、最小要因数で構成されるものを採用した (図-2.6)。

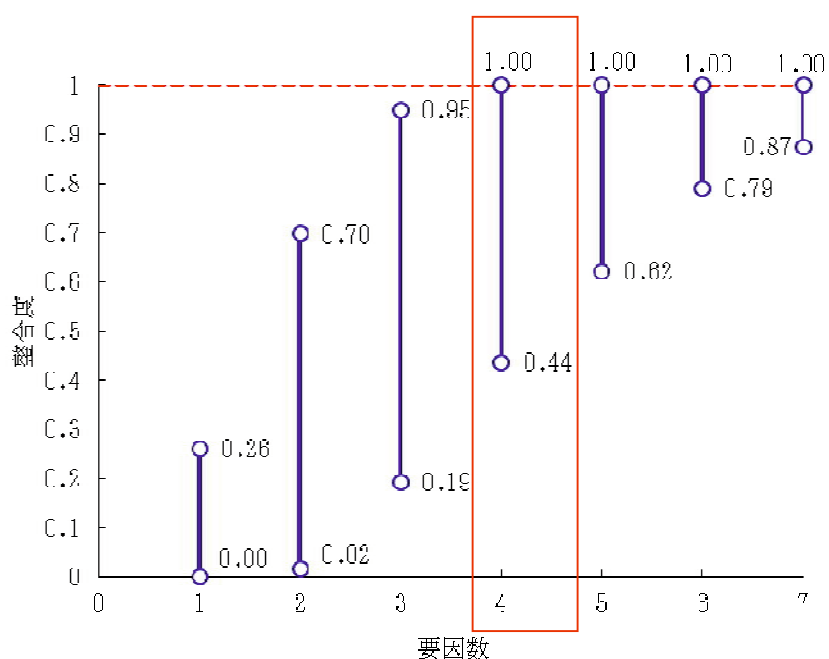


図-2.6 要因数と整合度の関係

なお、整合度 100.0%を満足する要因の組合せが多数存在する場合は、土石流の発生・非発生に対して分離性が高い要因⁵⁾ (以下、高分離性要因) (主溪流長, 流域長, 溪流方位, 流域幅, 流域最大傾斜, 源頭部面積, 地質小分類, 実効雨量 (半減期 72hr)) を多数含む組合せを採用した。なお、この高分離性要因とは、図-2.7 の各カテゴリのグラフで発生率が右肩上がりになっている要因の事を示しているものであり、これらの要因が多いルールほど拡張に伴う精度の低下を抑制できると考えられる。

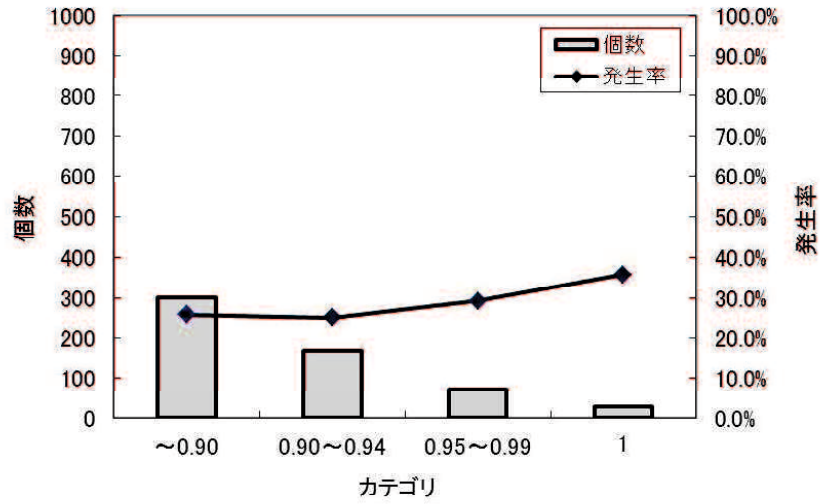


図-2.7(1) 水系模様のカテゴリ区分

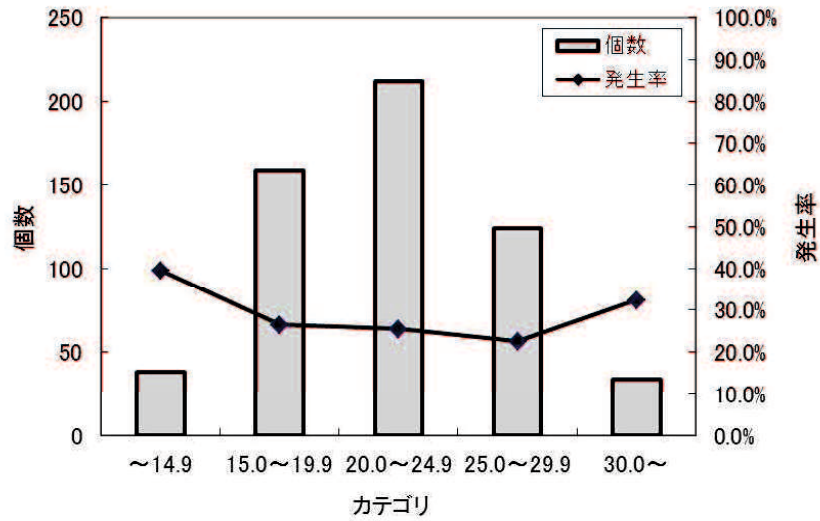


図-2.7(2) 流域平均勾配のカテゴリ区分

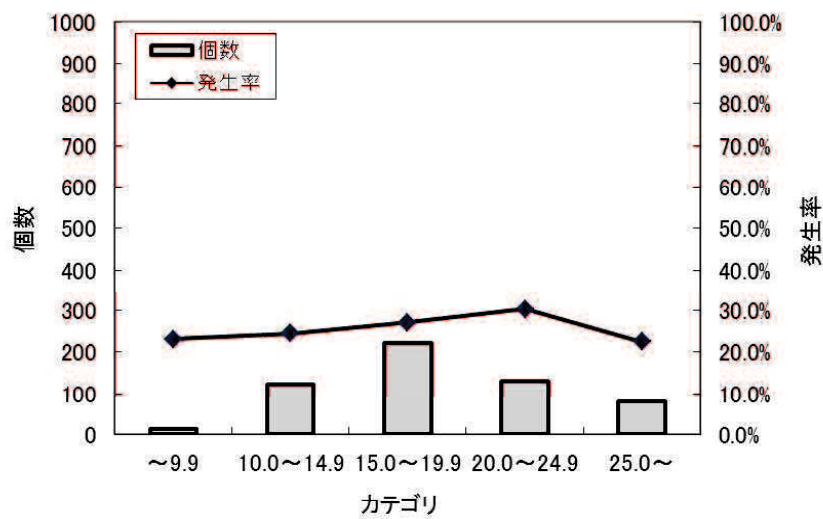


図-2.7(3) 最急渓床勾配のカテゴリ区分

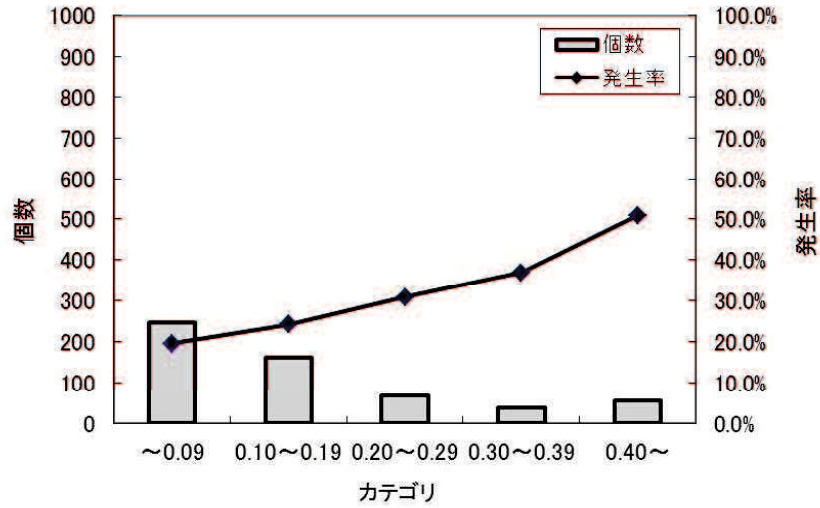


図-2.7(4) 主溪流長のカテゴリ区分（高分離性要因）

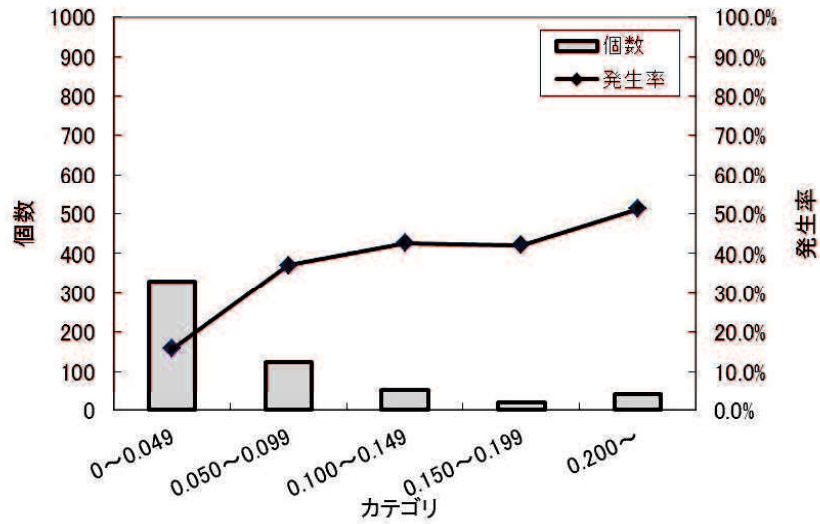


図-2.7(5) 流域面積のカテゴリ区分

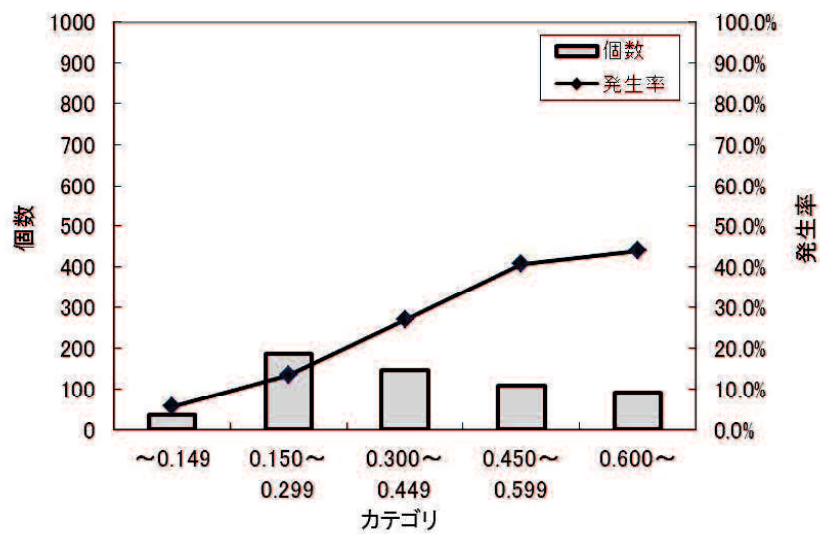
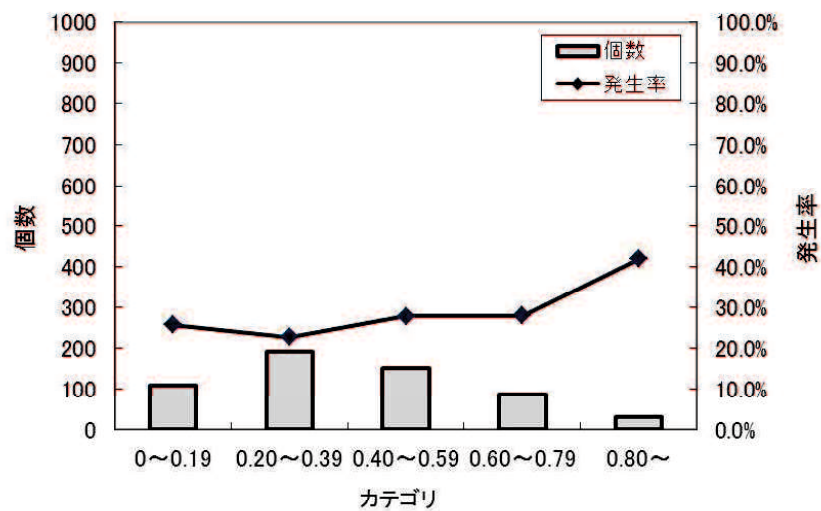
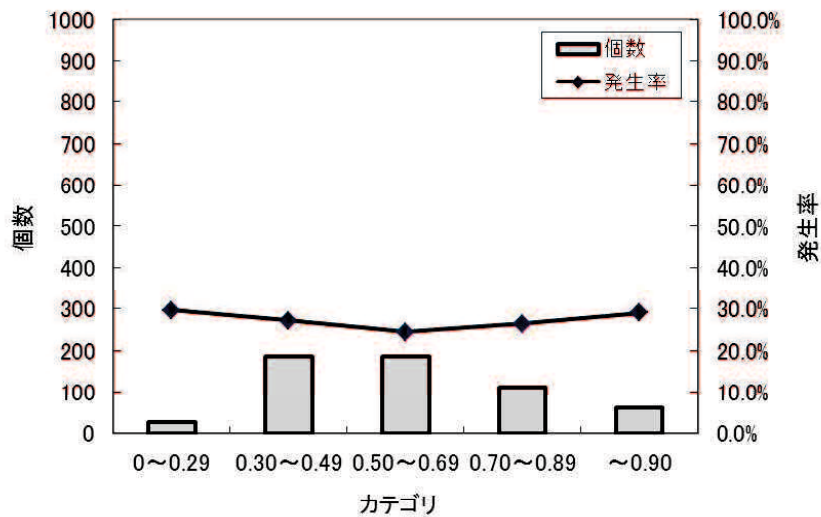
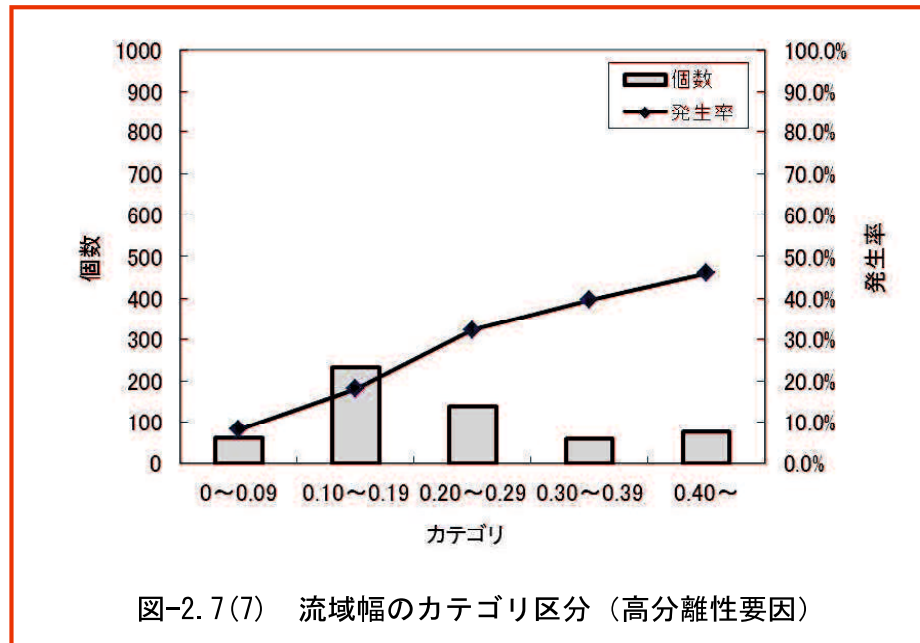
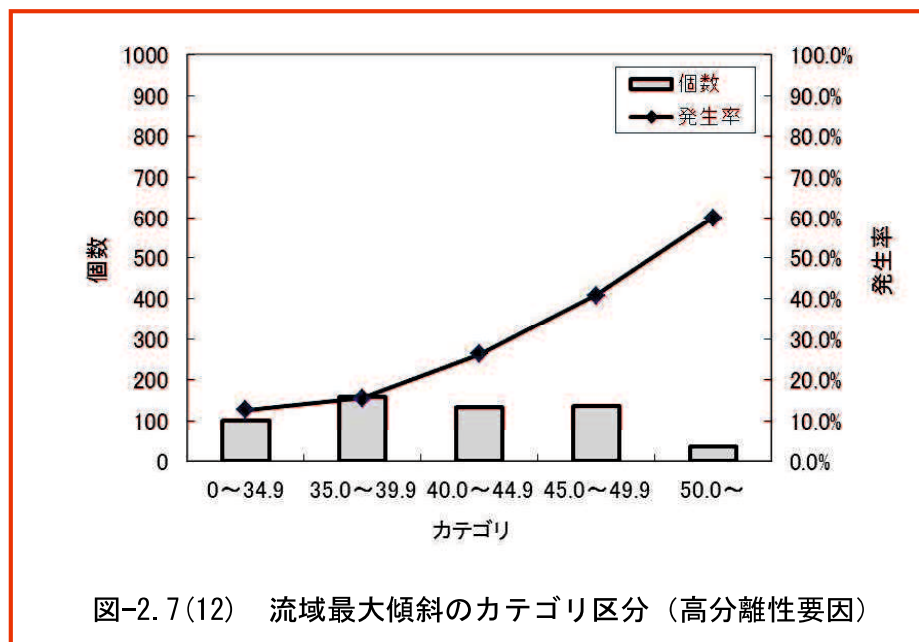
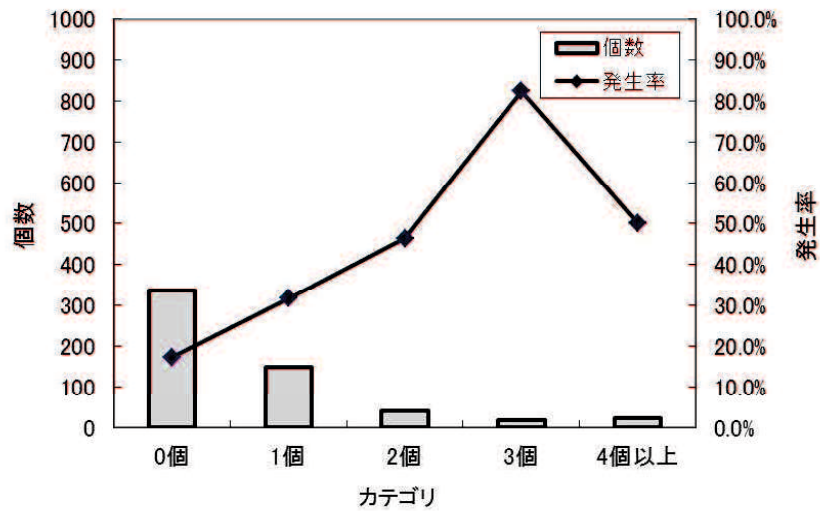
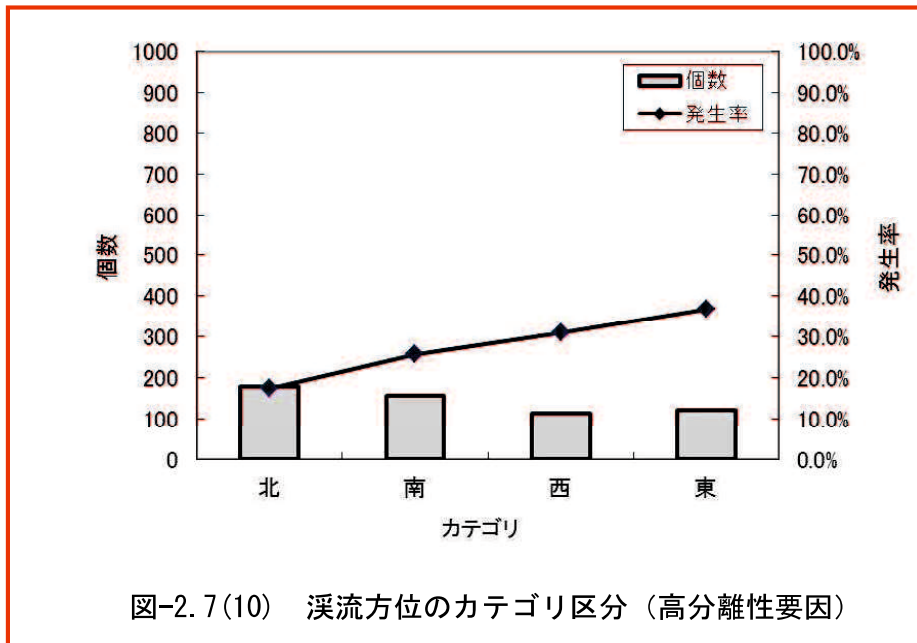


図-2.7(6) 流域長のカテゴリ区分（高分離性要因）





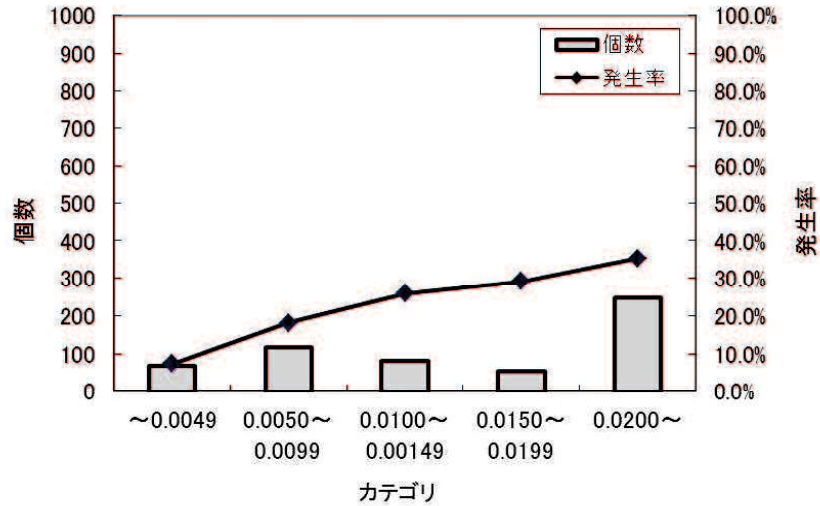


図-2.7(13) 源頭部面積のカテゴリ区分 (高分離性要因)

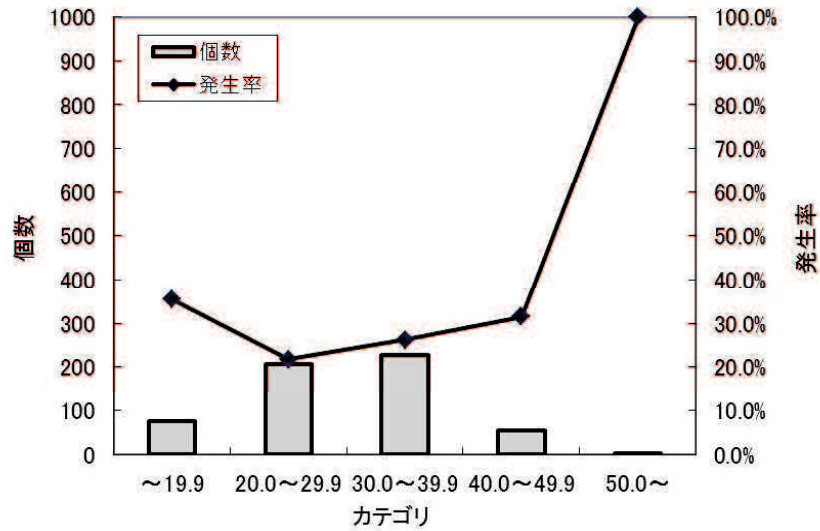


図-2.7(14) 源頭部最急谷勾配のカテゴリ区分

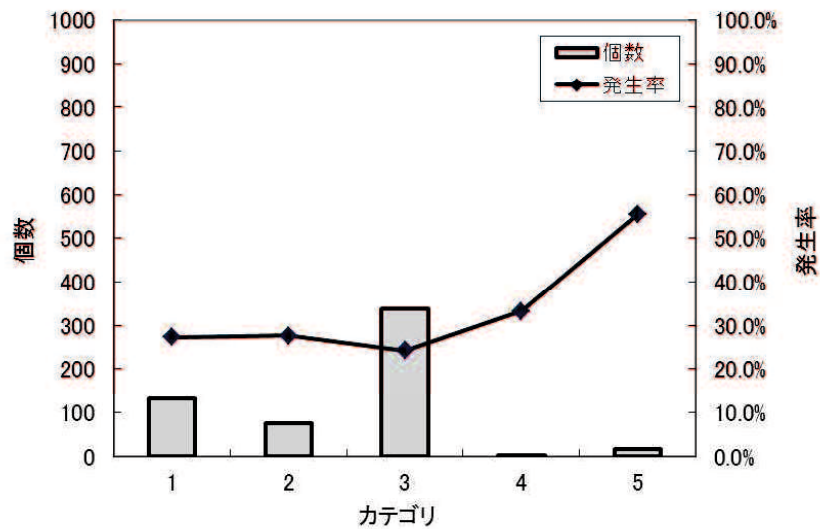


図-2.7(15) 降雨集中度評価のカテゴリ区分

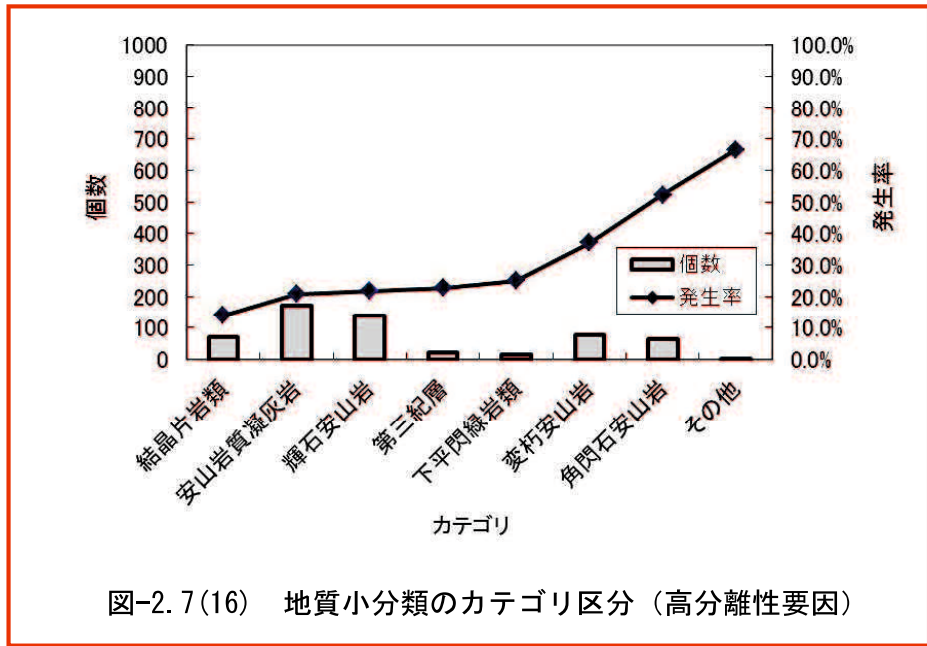


図-2.7(16) 地質小分類のカテゴリ区分 (高分離性要因)

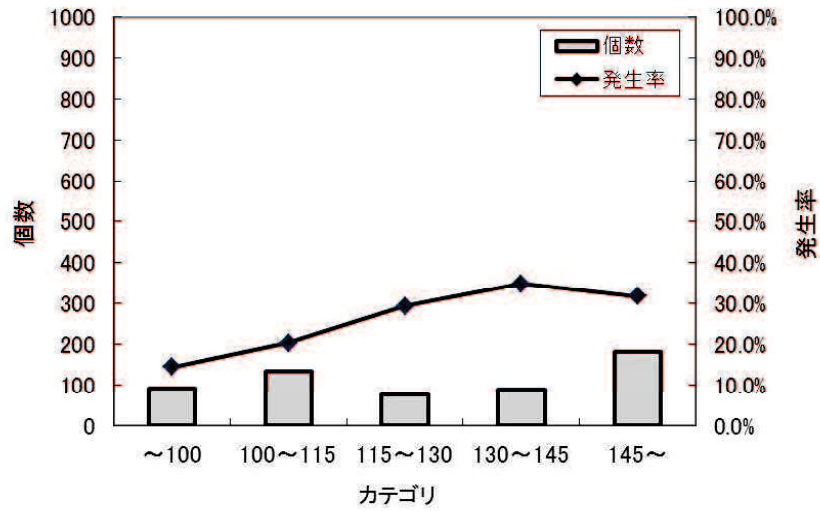


図-2.7(17) 時間雨量のカテゴリ区分

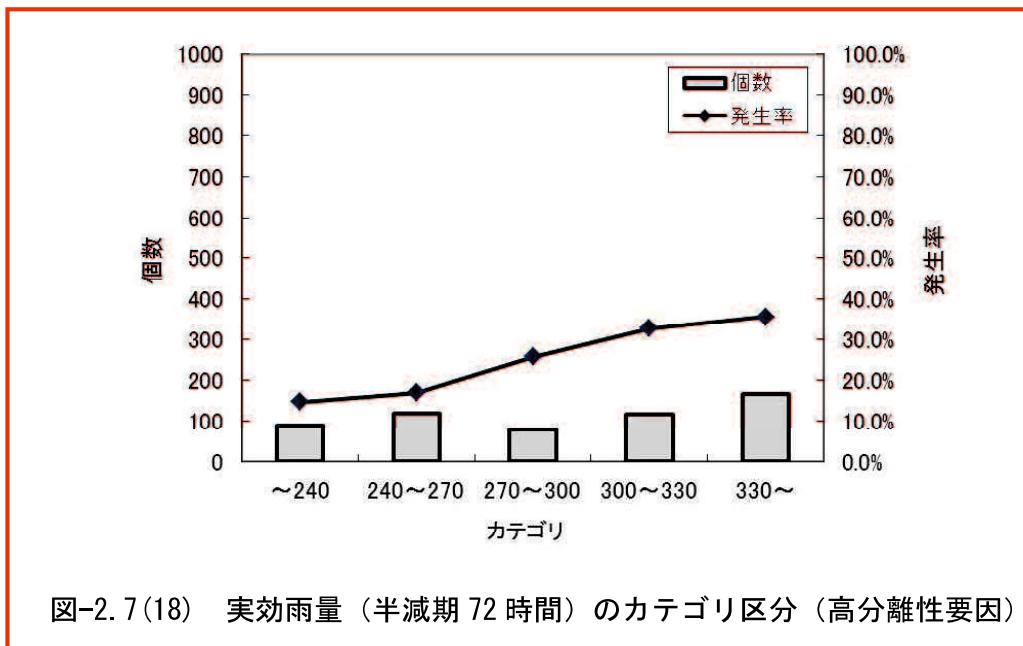


図-2.7(18) 実効雨量 (半減期 72 時間) のカテゴリ区分 (高分離性要因)

(2) 組合せ拡張ルールの抽出

汎用性が高いルールを抽出するため、ラフ集合を用いて抽出されたルールを基に、ルール領域の拡張⁵⁾、⁶⁾およびルールの組合せ⁵⁾を行った（以下、組合せ拡張ルール）。ここで、組合せ拡張ルールに用いるルールは、ある程度の誤判別を許容できる割合として確信度が90.0%以上⁴⁾のもののみとした。

2.4.4 検討結果および考察

代表データから抽出した組合せ拡張ルールによる精度（以下、適用精度）とそのルールを元のデータベースに適用した場合の適用精度を表-2.3に示す。

表-2.3(1) 組合せ拡張ルールの適用精度（代表データを対象）

ルール					適合結果						
ルール番号	要因				判定	含まれるデータ数 D1	整合データ数 D2	矛盾データ数 D3	残り箇所数 D4	累計	
	最急 溪床勾配 (°)	主溪流長 (km)	流域幅 (km)	地質小分類						確信度 D5 =	サポート D6 =
									119	D2/D1	D1/567
ルール16	*	0.09 ↓	0.29 ↓	*	非発生	50	46	4	69	92.0%	42.0%
ルール8	*	0.29 ↓	*	安山岩質凝灰岩 結晶片岩類	非発生	73	69	4	46	94.5%	61.3%
ルール7	19.9 ↓	*	0.19 ↓	*	非発生	79	75	4	40	94.9%	66.4%
ルール1	*	*	*	結晶片岩類	非発生	85	81	4	34	95.3%	71.4%
ルール5	*	*	0.19 ↓	輝石安山岩 安山岩質凝灰岩 結晶片岩類	非発生	86	82	4	33	95.3%	72.3%
ルール22	*	*	0.20 ↑	変朽安山岩 角閃石安山岩 その他	発生	102	98	4	17	96.1%	85.7%
ルール19	*	0.40 ↑	0.40 ↑	*	発生	107	103	4	12	96.3%	89.9%

表-2.3(2) 組合せ拡張ルールの適用精度（全データを対象）

ルール					適合結果						
ルール 番号	要因				判定	含まれる データ数 D1	整合 データ数 D2	矛盾 データ数 D3	残り 箇所数 D4 567	累計	
	最急 渓床勾配 (°)	主溪流長 (km)	流域幅 (km)	地質小分類						確信度	サポート
										D5= D2/D1	D6= D1/567
ルール16	*	0.09 ↓	0.29 ↓	*	非発生	222	184	38	345	82.9%	39.2%
ルール8	*	0.29 ↓	*	安山岩質凝灰岩 結晶片岩類	非発生	326	269	57	241	82.5%	57.5%
ルール7	19.9 ↓	*	0.19 ↓	*	非発生	361	299	62	206	82.8%	63.7%
ルール1	*	*	*	結晶片岩類	非発生	387	319	68	180	82.4%	68.3%
ルール5	*	*	0.19 ↓	輝石安山岩 安山岩質凝灰岩 結晶片岩類	非発生	394	324	70	173	82.2%	69.5%
ルール22	*	*	0.20 ↑	変朽安山岩 角閃石安山岩 その他	発生	460	364	96	107	79.1%	81.1%
ルール19	*	0.40 ↑	0.40 ↑	*	発生	477	376	101	90	78.8%	84.1%

(1) 組合せ拡張ルールの評価

表-2.3 より、ルールを構成する要因に着目すると、最急渓床勾配と地質小分類は竹本ら⁵⁾の研究において抽出された重要要因と一致しているとともに、主溪流長、流域幅、地質小分類は高分離性要因⁵⁾に該当している。このことから、本研究で抽出したルールを構成する要因は、土石流の発生・非発生に起因する重要要因として妥当な要因であると考えられる。これらの4要因を用いた組合せ拡張ルールは、7個のルール（発生2個、非発生5個）で構成されており、当該ルールの適用精度は、確信度 78.8%、サポート 84.1%と、ともに高い水準を示しており、現象に対する精度を確保しながら高い汎用性を有したルールが抽出されているものと考えられる。

抽出したルールから推測される S57.7 長崎豪雨災害の実態としては、「渓床勾配が緩くて流域の規模が小さく、かつ結晶片岩類もしくは安山岩質凝灰岩、結晶安山岩が分布する地域の溪流」では土石流は発生しておらず、「渓床勾配が急で流域の規模が大きく、かつ変朽安山岩もしくは角閃石安山岩が分布する地域の溪流」で土石流が発生したという特徴が挙げられる。これは、カテゴリ区分の中でも発生率が高い区分で発生ルールが構成され、発生率が低い区分で非発生ルールが構成されていることから S57.7 長崎豪雨災害の実態と良く整合した妥当な結果であると考えられる（図-2.8～図-2.11）。

これらのことから、篠崎ら⁴⁾の提案する手法は、S57.7 長崎豪雨災害の事例に対しても現象に対する精度を確保しながらも高い汎用性を有した実用的なルールを抽出することが可能な手法であると考えられる。

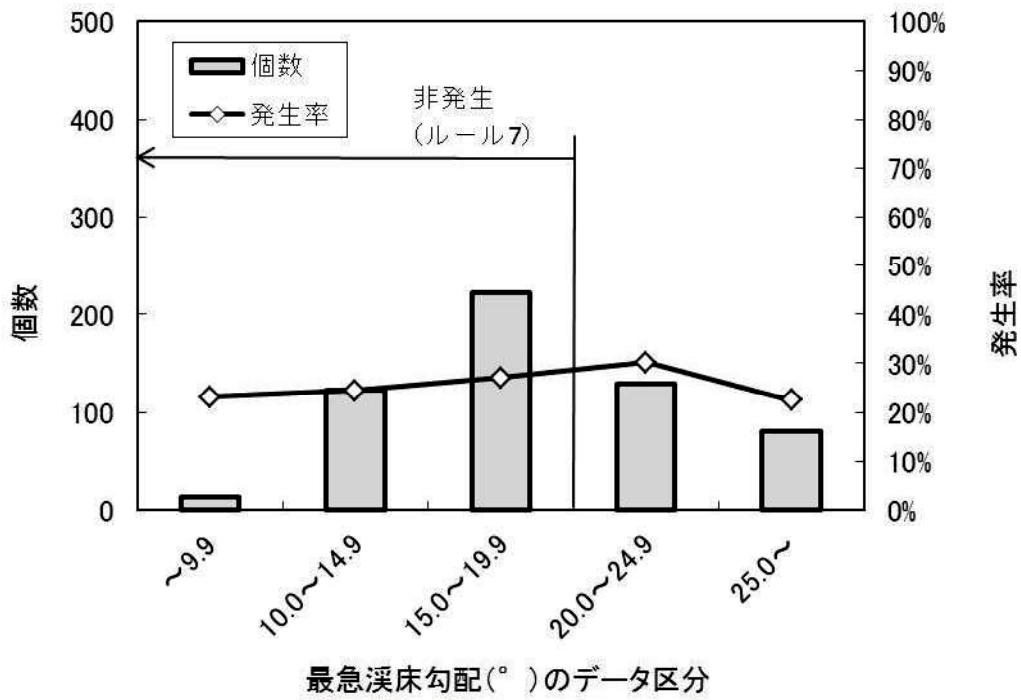


図-2.8 最急溪床勾配と発生率との関係

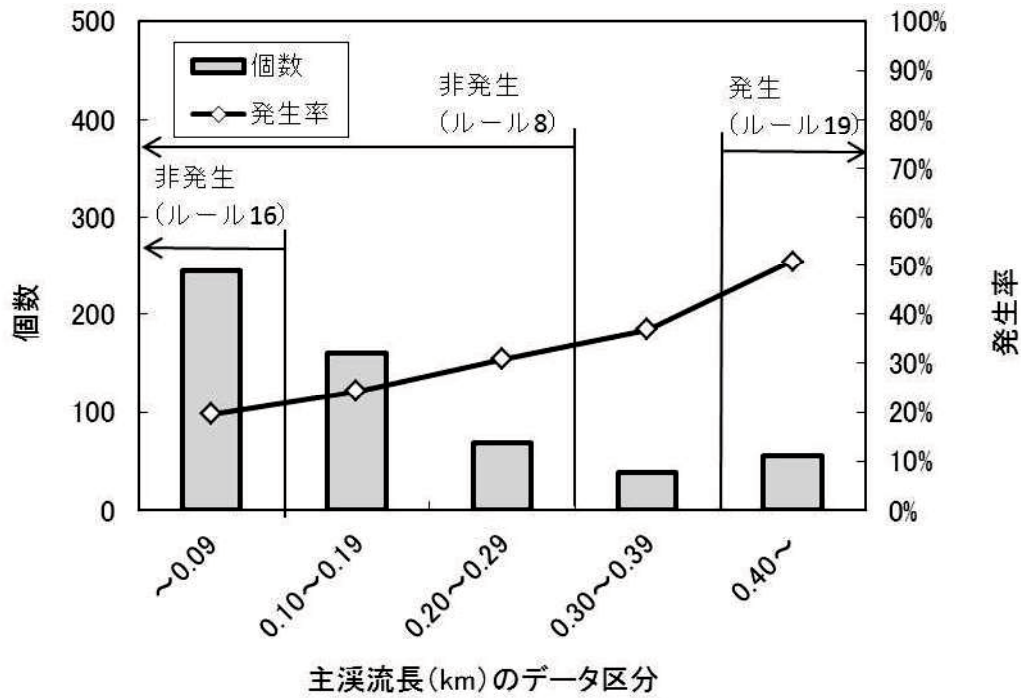


図-2.9 主溪流長と発生率との関係

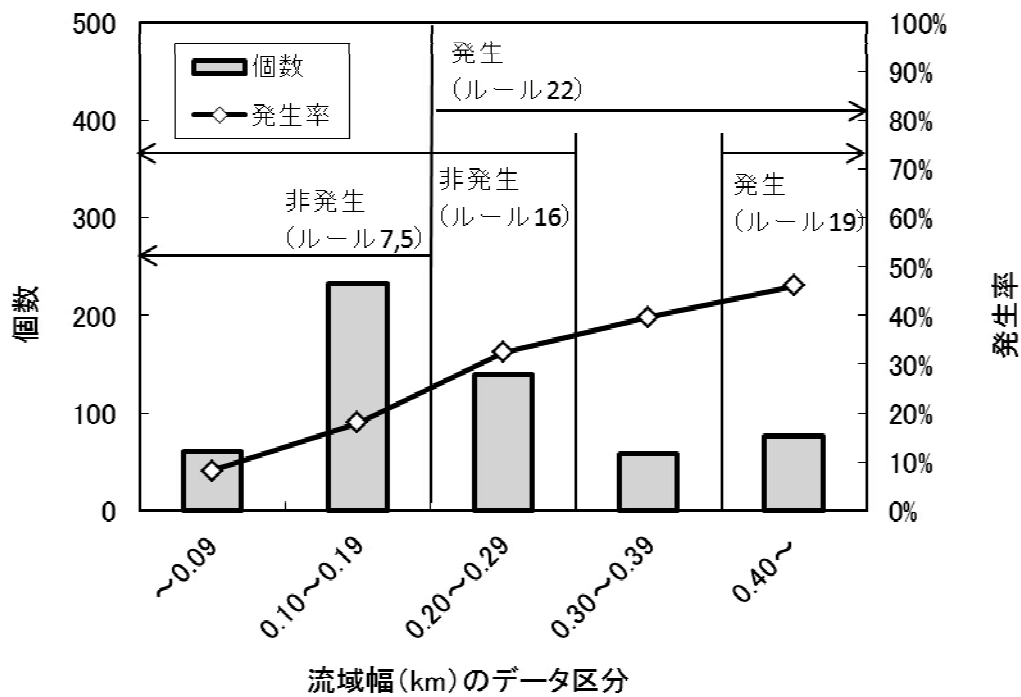


図-2.10 流域幅と発生率との関係

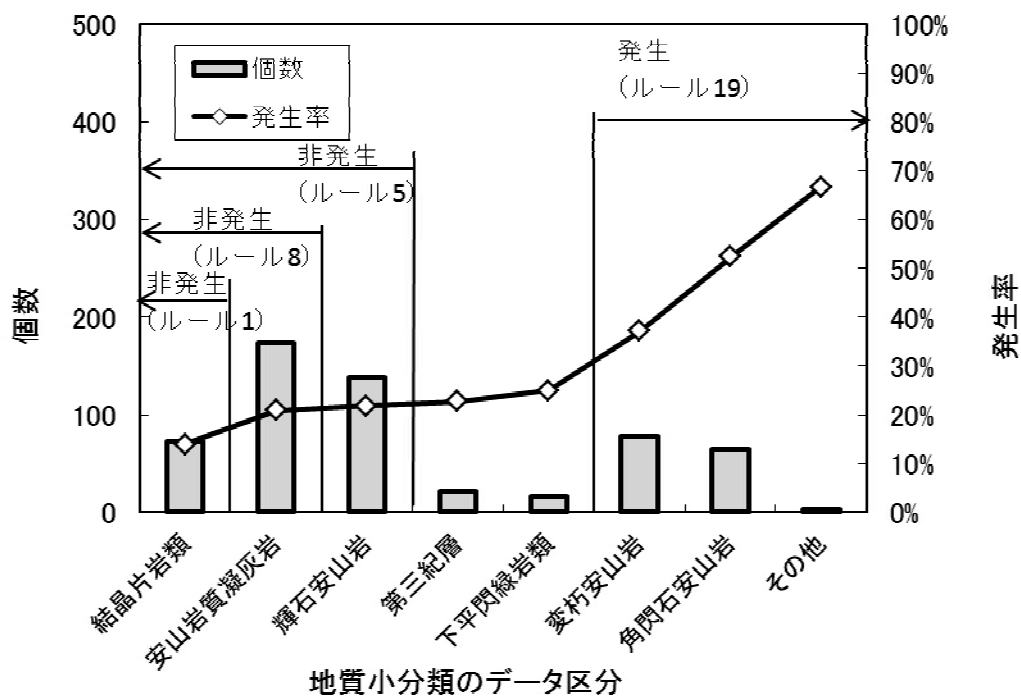


図-2.11 地質小分類と発生率との関係

(2) 既往研究との比較

篠崎ら⁴⁾の提案する手法の妥当性を確認するために、竹本らの研究⁵⁾との比較を行った¹¹⁾。表-2.4より、ラフ集合にルールの特長概念^{5), 6)}を導入した竹本ら⁵⁾の研究では、6つの要因を用いた22個のルール(発生10個, 非発生12個)が抽出されている。また、確信度が87.4%と非常に高く、サポートも61.6%と高い水準を示している。

一方、篠崎ら⁴⁾の提案するルール抽出方法を用いた本研究では、4つの要因を用いた7個のルール(発生2個, 非発生5個)となっており、要因やルール数が大きく削減されている。また、当該ルールの確信度は78.8%と竹本ら⁵⁾のルールに比べて劣る結果となったものの、その低下量は約9%に留まっていることから、現象に対する精度は大きく損なわれていない。一方、サポートは84.1%と大きく上回る結果となっている。これらのことを総括すると、本研究で抽出したルールは、要因やルール数の減少と汎用性の面において優れたルールであると評価できる。

要因やルール数を削減することは、今後のデータ作成において地形図からデータを抽出する時間が大幅に削減できるほか、抽出したルールをもとに現地を確認する場合においても、注視すべき箇所を少なくすることが可能となり、調査時間の短縮にも寄与するものと考えられる。また、ルールの汎用性を向上させることにより、抽出したルールを用いた危険箇所の特定がより容易になることから、詳細な現地調査を行う際の優先順位の検討や危険箇所の把握等、防災事業計画の策定において有効に活用できるものと考えられる。

これらのことから、篠崎ら⁴⁾の提案するルール抽出方法は、要因やルール数を大きく削減しながらも、現象に対する高い精度と汎用性を有したルールを抽出することが可能な有用性が高い手法であると考えられる。

表-2.4 既往研究との比較結果

		竹本らの研究 ⁵⁾ [ラフ集合のみ]	本研究 ¹⁾ [SVM+ラフ集合]	
データ使用	データ数	[長崎市] 567溪流 (発生150 非発生417)		
	要因	地形・地質16要因, 降雨 2要因		
SVMによる危険度設定	パラメータ	—	$C=500, r=4$	
	代表データの抽出方法	—	全発生データに占める-1より小さいデータ21%(32個)を抽出する ↓ 全非発生データの21%(87個)を抽出	
ラフ集合によるルール抽出	整合度の要求水準	95.0%以上	100.0%	
	確信度の要求水準	発生:67.5%以上 非発生:92.4%以上	発生:90.0%以上 非発生:90.0%以上	
	データ数	発生:150 非発生:417	発生:32 非発生:87	
	抽出された重要要因		最急渓床勾配	最急渓床勾配
				主溪流長
			流域長	
				流域幅
			溪流方位	
			流域最大傾斜	
		地質小分類	地質小分類	
		実効雨量(半減期72hr)		
	ルール数	発生:10 非発生:12	発生:2 非発生:5	
適用精度	累計確信度	87.4% (305/349)	78.8% (376/477)	
	累計サポート	61.6% (349/567)	84.1% (477/567)	

2.5 まとめ

本章では篠崎ら⁴⁾の提案するルール抽出手法を用いて、竹本ら⁵⁾が研究対象とした S57.7 長崎豪雨災害における土石流の発生・非発生ルールを改めて抽出し、ルールの適用精度に関する既往研究⁵⁾との比較から手法としての有用性の検証を試みた。本研究における主要な結論を以下に示す。

- ①本研究で抽出したルールでの分析の結果、S57.7 長崎豪雨災害における土石流発生・非発生事例の 84.1%を、わずか7つのルールで説明できる非常に汎用性の高いルールを得ることができた。
- ②ルールの確信度は 78.8%と既往研究⁵⁾に比べやや劣る結果となったものの、全データの 84.1%を説明できる汎用性の高いルールであることを勘案すると、精度としては十分な水準を有しているものと考えられる。
- ③篠崎ら⁴⁾の提案するルール抽出方法を活用することで、要因やルール数を大きく削減しながらも、現象に対する高い精度と汎用性を有したルールを抽出でき、この方法に汎用性があることが確認できた。
- ④本研究での取り組みにより、元々優れた利点があったラフ集合が SVM を組み合わせることでルール数や要因数が多くなるという課題を解消できることとなり、災害の発生・非発生ルールを検討するための手法としてさらに適用性が高いものであることが明らかとなった。

参考文献

- 1) 河野浩之：データベースからの知識発見の現状と動向，人工知能学会誌，Vol.12，No.4，pp.497-504，1997.
- 2) 岡本正男，荒木義則，中山弘隆，古川浩平，水山高久：ラフ集合を用いたデータマイニングによる土砂移動現象の重要要因及びルール抽出に関する研究，砂防学会誌，Vol.54，No.6，pp.4-15，2002.
- 3) Cortes C. and Vapnik, V.: Support vector networks, *Machine Learning*, Vol.20, pp.273-297, 1995.
- 4) 篠崎嗣浩，森田真人，大石博之，古川浩平：SVM とラフ集合を用いた土石流の発生・非発生ルールの作成方法に関する研究，土木学会論文集 F, Vol. 65, No.4, pp.448-460, 2009.
- 5) 竹本大昭，杉原成満，佐藤丈晴，荒木義則，中山隆弘，水山高久，古川浩平：ラフ集合を用いた土石流発生・非発生ルールの抽出と危険度の設定手法に関する研究，砂防学会誌，Vol.57，No.2，pp.4-15，2004.

- 6) Greco, S.: An Algorithm for Induction of Decision Rule Consistent with the Dominance Principle, Lecture Notes in Artificial Intelligence, Vol.2005, Springer-Verlag, Berlin, pp.304-313, 2001.
- 7) 大石博之, 小林央宜, 尹禮分, 田中浩一, 中山弘隆, 古川浩平: サポートベクターマシンによる対策工効果を考慮した斜面災害危険度の設定, 土木学会論文集 F, Vol.63, No.1, pp.107-118, 2007.
- 8) 社団法人全国治水砂防協会: 土砂災害防止法令の解説—土砂災害警戒区域等における土砂災害防止対策の推進に関する法律—, 2003.
- 9) 建設省土木研究所砂防部砂防研究室: 昭和 57 年 7 月長崎豪雨による土砂災害調査報告書, 土木研究所資料, ISSN0386-5878, 土研資料第 2107 号, p.44-45, 1984.
- 10) 岡本正男, 仲野公章, 秦耕二, 荒木義則, 古川浩平, 水山高久: 局地的な集中豪雨に対する土石流警戒避難基準雨量の設定に関する研究～1999.6.29 広島災害における事例～, 砂防学会誌, Vol.55, No.1, p.15-25, 2002.
- 11) 杉原成満, 小林央宜, 川崎雄大, 篠崎嗣浩, 大石博之, 古川浩平: 昭和 57 年 7 月長崎豪雨における土石流の発生・非発生ルールの抽出, 土木学会論文集 F, Vol. 66, No. 4, pp.593-598, 2010.

第3章 災害危険度ランクに対応した発生・非発生ルールの明示化

3.1 緒言

第2章では、過去に検討実績のある土砂災害危険箇所を対象としたSVMによる危険度評価および災害発生・非発生ルールの抽出手法¹⁾について、これまで対象としていなかった地域（長崎市）から取得したデータを対象として適用性検証を実施した²⁾。一連の検証作業を実施した結果、長崎市のデータに対しても既往論文¹⁾と同等の活用が可能であることが確認できた²⁾。また、同じ長崎市のデータでラフ集合のみを使った検討を実施していた竹本ら³⁾の研究成果と比較した場合にも、要因数やルール数を削減しながら高いサポートを有した汎用性の高いルールを得ることができており、SVMとラフ集合の組合せによる手法の優位性が明確となった。

しかしながら、前章で抽出されたルールは、「災害の発生するルール」か「災害の発生しないルール」のみを説明するものであったという点に課題を残している。

例えば、土石流の恐れのある箇所として分類されたデータについても、本来は危険の程度に差があると考えられる。データを危険度に対応したグループに更に細分し、その危険度に対応した類型的なルールを特定することができれば、優先的に防災事業を実施すべき箇所を容易に選定することが可能となり、また有効な対策方法の検討にも資するものと考えられる。

そこで、本章ではSVMの基本機能を用いて、対象となるデータ群を①災害発生の可能性の高い危険領域のデータ群、②災害発生の可能性が比較的高いと考えられるデータ群、③災害発生の可能性が比較的低いと考えられるデータ群、④災害発生の可能性の低い安全領域のデータ群の4つのグループに判別し、それぞれに対応したルールを検討することを試みた。

3.2 手法および使用データの概要

本章の検討に用いる手法は、危険度ランクの設定にSVM⁴⁾、ルールの抽出にはラフ集合⁵⁾を用いることとした。なお、手法の詳細については、2章の2.2にそれぞれ記載している。

また、使用するデータについては、2章と同様に長崎市内の土石流危険溪流567溪流(土石流発生150溪流、非発生417溪流)を検討対象データとした。検討に用いる要因についても、地形地質要因：16要因、降雨要因：2要因の計18要因（2章の表-2.1参照）とし、2章での検討を踏襲した。

3.3 災害発生危険度に対応した評価ルールの設定

3.3.1 SVMによる危険度ランクの設定

SVMを用いた危険度設定では、地形地質要因等の土砂災害の発生に起因する要因と過去の災害履歴データを用いて災害の発生と非発生を分離する分離超平面の構築を行う。その際には、分析パラメータとして誤判別の度合いを調整する C とガウシアンカーネルの半径(データの影響度)である r の2つのパラメータを設定する必要がある。ここでは、前章と同様の考え方で、的中率100.0%のパラメータセットのうち、サポートベクターとなるデータ ($|f(x)| \leq 1$ となるデータ)数が最も少なくなる $C=500$ 、 $r=4$ を採用した。SVMの特徴としては、発生・非発生を判別(分離)する分離超平面 ($f(x)=0$)を境界として、その分離超平面を構成するのに寄与するサポートベクター ($|f(x)| \leq 1$)及びその外側のデータ ($|f(x)| \geq 1$)とに区分することができる(図-3.1)。そこでこの特徴を活かし、データ群を①災害発生の可能性の高い危険領域のデータ群、②災害発生の可能性が比較的高いと考えられるデータ群、③災害発生の可能性が比較的低いと考えられるデータ群、④災害発生の可能性の低い安全領域のデータ群の4つのグループに判別した(図-3.2)。

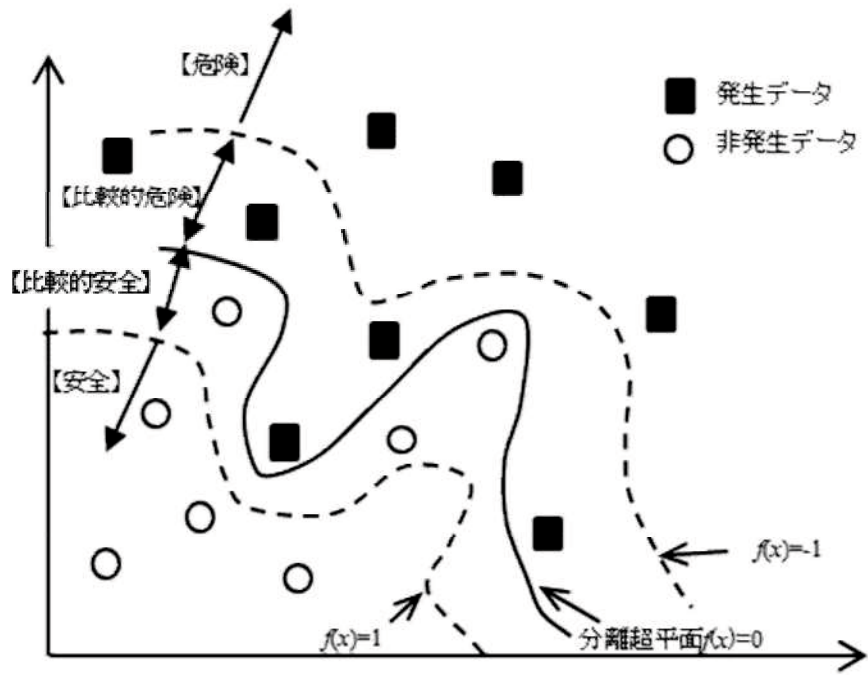


図-3.1 危険度ランクのイメージ図

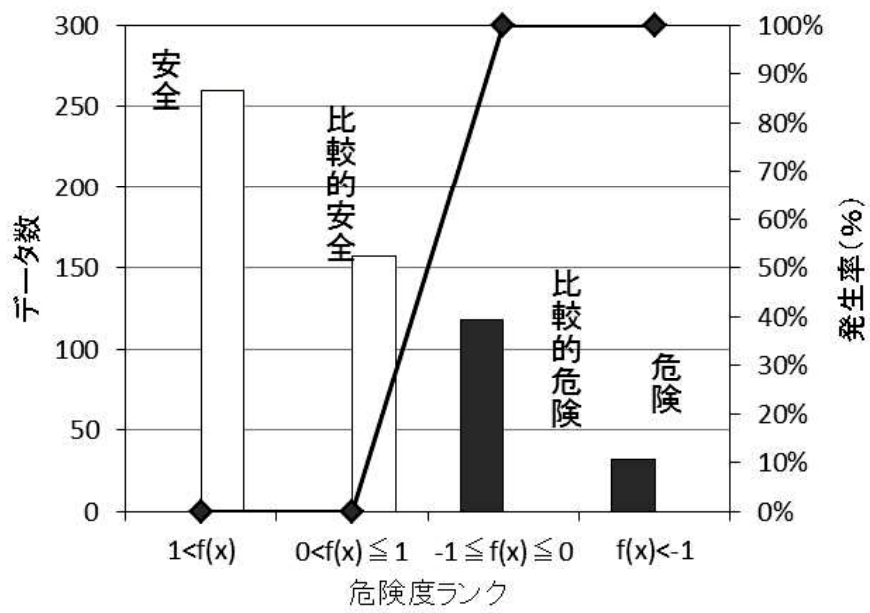


図-3.2 危険度ランク毎のデータ数

3.3.2 危険度ランクに対応した評価ルールの設定

(1) ラフ集合の分析に用いるデータ

本検討でラフ集合の分析に用いるデータは、前節で区分した危険度ランクの領域に属するデータをそれぞれ用いることとした。各ランクに属するデータ群はSVMにより災害発生、非発生を完全に分離したことにより形成されたものであることから、抽出されるルールの精度は高いものとなる事が予想される。

(2) 整合度の設定と重要要因の選定

ラフ集合で土石流の評価ルールを設定する場合、要因数を増やすと整合度は高くなるが、ルールが煩雑になってしまう。逆に要因数を減らすとルールの簡素化は可能となるが整合度が低下してしまうことから、適切に設定することが必要となる。前章のように発生と非発生を代表するデータを用いてルールを抽出する場合は、データの質の低下を防ぐために整合度100.0%を要求水準とすることが有効である。しかしながら、本検討では、全てのデータを用いてルール抽出を行うことから、同じく全てのデータを用いている既往研究³⁾に習って整合度の要求水準は95.0%とした。その上で要求水準を満足する最少要因数である7要因を採用する事とした。

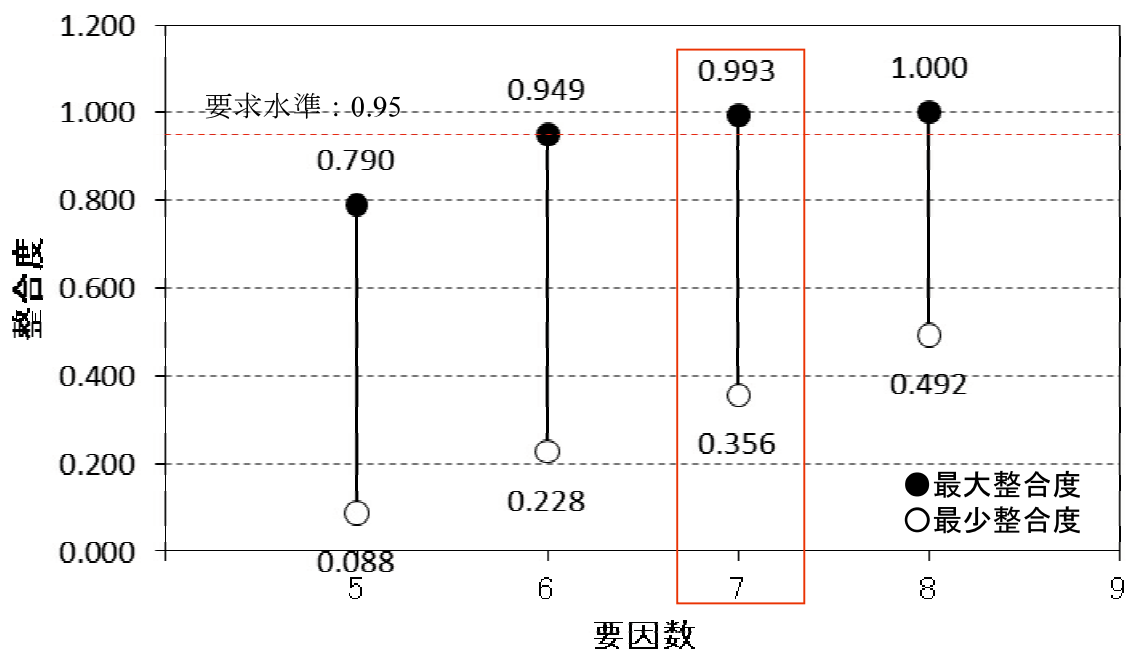


図-3.4 要因数と整合度の関係

なお、要求水準を満足する組合せが多数存在する場合には、土石流の発生・非発生に対して分離性の高い要因（以下、高分離性要因）（主溪流長、流域長、溪流方位、流域幅、流域最大傾斜、源頭部面積、地質小分類、実効雨量（半減期 72hr））を多数含む組合せの中で最も整合度の高いものを採用した。

(3) 組合せ拡張ルールの抽出

汎用性の高いルールを抽出するため、ラフ集合を用いて抽出されたルールを基に、ルール領域の拡張³⁾およびルールの組合せ³⁾を行った(以下、組合せ拡張ルール)。ここで、組合せ拡張ルールに用いる確信度は、既往研究³⁾と同じように発生側(危険、比較的危険)を67.5%以上、非発生側(安全、比較的安全)を92.4%以上とした。

(4) 検討結果および考察

危険度ランクごとに抽出した組合せ拡張ルールの適用精度を表-3.1に示す。

表-3.1より、設定した評価ルールは、最急溪床勾配、流域長、溪流方位、流域最大傾斜、源頭部面積、地質小分類、実効雨量(半減期72hr)の7要因から構成されている。抽出されたルールは、危険:3ルール、比較的危険:4ルール、安全:12ルール、比較的安全:5ルールの計24ルールであった。設定した安全側のルールに着目すると、「溪床勾配や流域内の傾斜が緩く、流域の規模が小さい(流域長が短く、源頭部面積が小さい)、かつ結晶片岩類もしくは安山岩質凝灰岩が分布する地域の北向き溪流」では、土石流の危険性が低い事が分かる。中でも源頭部面積は、「安全」のルールにのみ出現していることから、特に危険性が低いと判断できる。逆に危険側のルールに着目すると、「溪床勾配や流域内の傾斜が急で、流域の規模が大きい(流域長が長く、源頭部面積が大きい)、かつ角閃石安山岩が分布する地域の東向き溪流」では、土石流の危険性が高いと言える。特に最急溪床勾配と流域最大傾斜については、「危険」のルールの方が「比較的危険」のルールに比べて危険側の条件となっており、違いが明確になっている。当該ルールの実現象に対する精度(確信度)は90.0%を維持しており、非常に精度の高いルールが抽出されている。一方、ルールの汎用性を示すサポートについては、56.6%となっており、全体の6割弱ほどの特定に留まってしまった。

以上のことから、今回抽出したルールは、ルールの精度こそ高いものの、そのルールで説明できる範囲が非常に狭く汎用性が高いとは言えないものであると言える。

よって実用化のためには汎用性向上についての更なる検討が必要と考えられる。

表-3.1 危険度ランクごとに抽出した組合せ拡張ルールの適用精度

ルール番号	要因										判定	適合結果					
	最急 渓床勾配 (%)	流域長 (km)	流域方位	流域最大 傾斜 (%)	源頭部 面積 (km ²)	地質 小分類	実効雨量 (T=72hr) (mm)	含まれる データ数 D1	整合 データ数 D2	矛盾 データ数 D3		残り 箇所数 D4	累計				
													確信度 D5=	D6=	DI/567		
Rule77	*	0.299 ↓	*	*	*	結晶片岩類 安山岩質凝灰岩 輝石安山岩	*	154	144	10	413	93.5%	27.2%				
Rule51	*	*	*	34.9 ↓	0.0199 ↓	*	*	184	172	12	383	93.5%	32.5%				
Rule35	*	0.599 ↓	北, 南	*	*	結晶片岩類 結晶片岩類 安山岩質凝灰岩 輝石安山岩 第三紀層 下平閃線岩類 菱朽安山岩	*	197	185	12	370	93.9%	34.7%				
Rule96	*	0.299 ↓	北, 南	*	0.0099 ↓	*	*	207	195	12	360	94.2%	36.5%				
Rule98	14.9 ↓	*	北	39.9 ↓	*	*	*	215	203	12	352	94.4%	37.9%				
Rule43	24.9 ↓	0.299 ↓	*	*	0.0049 ↓	*	*	220	208	12	347	94.5%	38.8%				
Rule81	*	*	北	*	*	結晶片岩類 安山岩質凝灰岩	240 ↓	224	212	12	343	94.6%	39.5%				
Rule94	19.9 ↓	0.299 ↓	北, 南	39.9 ↓	*	*	*	221	215	12	340	94.7%	40.0%				
Rule13	*	*	*	39.9 ↓	0.0099 ↓	*	330 ↓	229	217	12	338	94.8%	40.4%				
Rule108	24.9 ↓	*	北, 南	*	0.0099 ↓	結晶片岩類 安山岩質凝灰岩	*	231	219	12	336	94.8%	40.7%				
Rule121	24.9 ↓	0.449 ↓	北	*	*	*	240 ↓	233	221	12	334	94.8%	41.1%				
Rule55	*	0.299 ↓	*	*	*	*	270 ↓	234	222	12	333	94.9%	41.3%				
Rule174	9.9 ↓	0.449 ↓	北, 南, 西	*	*	*	*	236	224	12	331	94.9%	41.6%				
Rule215	*	0.449 ↓	*	34.9 ↓	*	*	270 ↓	238	226	12	329	95.0%	42.0%				
Rule129	*	0.149 ↓	*	*	*	*	300 ↓	239	227	12	328	95.0%	42.2%				
Rule204	*	*	北, 南	34.9 ↓	*	結晶片岩類 安山岩質凝灰岩 輝石安山岩	270 ↓	240	228	12	327	95.0%	42.3%				
Rule230	19.9 ↓	0.299 ↓	北, 南	44.9 ↓	*	*	*	241	229	12	326	95.0%	42.5%				
Rule328	*	*	南, 西, 東	*	*	角閃石安山岩 その他	270 ↑	278	256	22	289	92.1%	49.0%				
Rule319	*	*	東	50.0 ↑	0.0050 ↑	*	*	286	264	22	281	92.3%	50.4%				
Rule324	25.0 ↑	*	*	50.0 ↑	0.0100 ↑	*	*	287	265	22	280	92.3%	50.6%				
Rule303	15.0 ↑	0.450 ↑	西, 東	*	*	*	330 ↑	309	280	29	258	90.6%	54.5%				
Rule317	*	0.600 ↑	*	45.0 ↑	0.0200 ↑	*	300 ↑	319	287	32	248	90.0%	56.3%				
Rule251	*	*	*	45.0 ↑	*	角閃石安山岩 その他	330 ↑	320	288	32	247	90.0%	56.4%				
Rule260	*	*	東	45.0 ↑	*	角閃石安山岩 その他	*	321	289	32	246	90.0%	56.6%				

3.4 まとめ

本章では、竹本ら³⁾が研究対象とした S57.7 長崎豪雨災害における土石流のルールを危険度ランクごとに抽出することを試みた。本研究における主要な結論を以下に示す。

- ①区分した 4 つの危険度ランクに対応したルールを抽出することを試みた結果、最急
渓床勾配、溪流長、溪流方位、流域最大傾斜、源頭部面積、地質小分類、実効雨量
(半減期 72hr) の 7 要因で構成される計 24 個のルール (危険 : 3, 比較的危険 : 4,
比較的安全 : 5, 安全 : 12) を抽出することができた。
- ②今回抽出したルールは、確信度こそ 90.0% と高いが、そのルールで説明できる範囲
が非常に狭く (サポート 56.6%) 汎用性に関しては高いとは言えないものであった。

よって、汎用性の向上のための更なる検討が必要と考えられるため、次章で検討を行った。

参考文献

- 1) 篠崎嗣浩, 森田真人, 大石博之, 古川浩平 : SVM とラフ集合を用いた土石流の発生・非発生ルールの作成方法に関する研究, 土木学会論文集 F, Vol. 65, No.4, pp.448-460, 2009.
- 2) 杉原成満, 小林央宜, 川崎雄大, 篠崎嗣浩, 大石博之, 古川浩平 : 昭和 57 年 7 月長崎豪雨における土石流の発生・非発生ルールの抽出, 土木学会論文集 F, Vol. 66, No. 4, pp.593-598, 2010.
- 3) 竹本大昭, 杉原成満, 佐藤丈晴, 荒木義則, 中山隆弘, 水山高久, 古川浩平 : ラフ集合を用いた土石流発生・非発生ルールの抽出と危険度の設定手法に関する研究, 砂防学会誌, Vol.57, No.2, pp.4-15, 2004.
- 4) Cortes C. and Vapnik, V.: Support vector networks, *Machine Learning*, Vol.20, pp.273-297, 1995.
- 5) 河野浩之 : データベースからの知識発見の現状と動向, 人工知能学会誌, Vol.12, No.4, pp.497-504, 1997.

第4章 災害危険度ランクに対応した発生・非発生ルール汎用性向上の検証

4.1 緒言

第3章では、危険度ランクに対応したルールの抽出方法の検討を実施した。しかしながら、抽出したルールは、精度こそ高いものの、そのルールで説明できる範囲が非常に狭く汎用性を確保するためには更なる検討が必要となった。これは、危険度評価の段階でSVMを利用する際に、データを災害発生・非発生の2群に完全分離する分離超平面（危険度評価の基準となる面）を採用したことが原因の一つであると考えられる。

母集団のデータを完全分離させた場合の分離超平面は、一般の傾向とは異なる特殊な事例（以下、ノイズデータ）を含んだまま判別してしまうことになる（図-4.1）。このようなデータを含んだデータ群を母集団として、確信度の高いルールを抽出すると、結果として説明できる範囲（図-4.1の着色部）が狭くなりサポートが低下してしまうと考えられる。

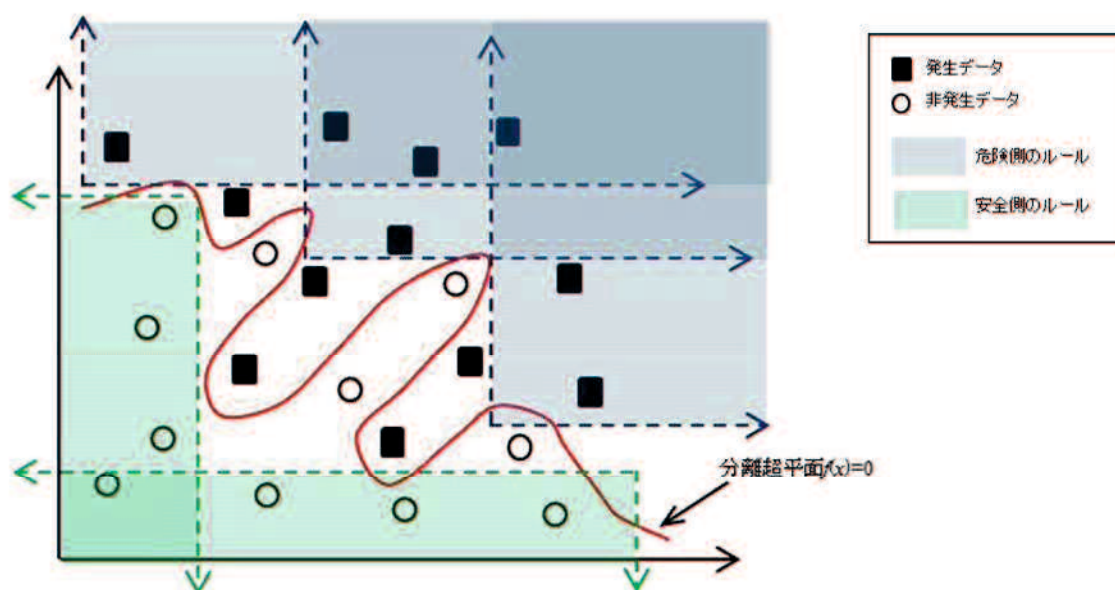


図-4.1 完全分離によるルール設定のイメージ図

防災事業を効率的に進めていくためには、特異事例に引きずられることなく、対象地域の典型的なルールを抽出し、詳細調査や対策工事の必要となる箇所を絞込み、および調査不要箇所を特定を行うことが重要である。

そこで本研究では、上記のノイズデータに引きずられない分離超平面を構築し、そこから危険度ランクを設定してルールを抽出する手法の検討に取り組んだ¹⁾。

4.2 災害発生危険度に対応した評価ルールの設定方法

4.2.1 手法および使用データの概要

本章の検討に用いる手法は、危険度ランクの設定にSVM²⁾、ルールの抽出にはラフ集合³⁾を用いることとした。なお、手法の詳細については、2章の2.2にそれぞれ記載している。また、使用するデータについては、2章と同様に長崎市内の土石流危険溪流567溪流(土石流発生150溪流, 非発生417溪流)を検討対象データとした⁴⁾。検討に用いる要因についても、地形地質要因:16要因, 降雨要因:2要因の計18要因(2章の表-2.1参照)とし、2章での検討を踏襲した。

4.2.2 災害発生危険度に対応した評価ルールの設定手順

SVM とラフ集合を用いた災害発生危険度に対応した評価ルールの設定手順を図-4.2に示す。また、その詳細を以下に示す。

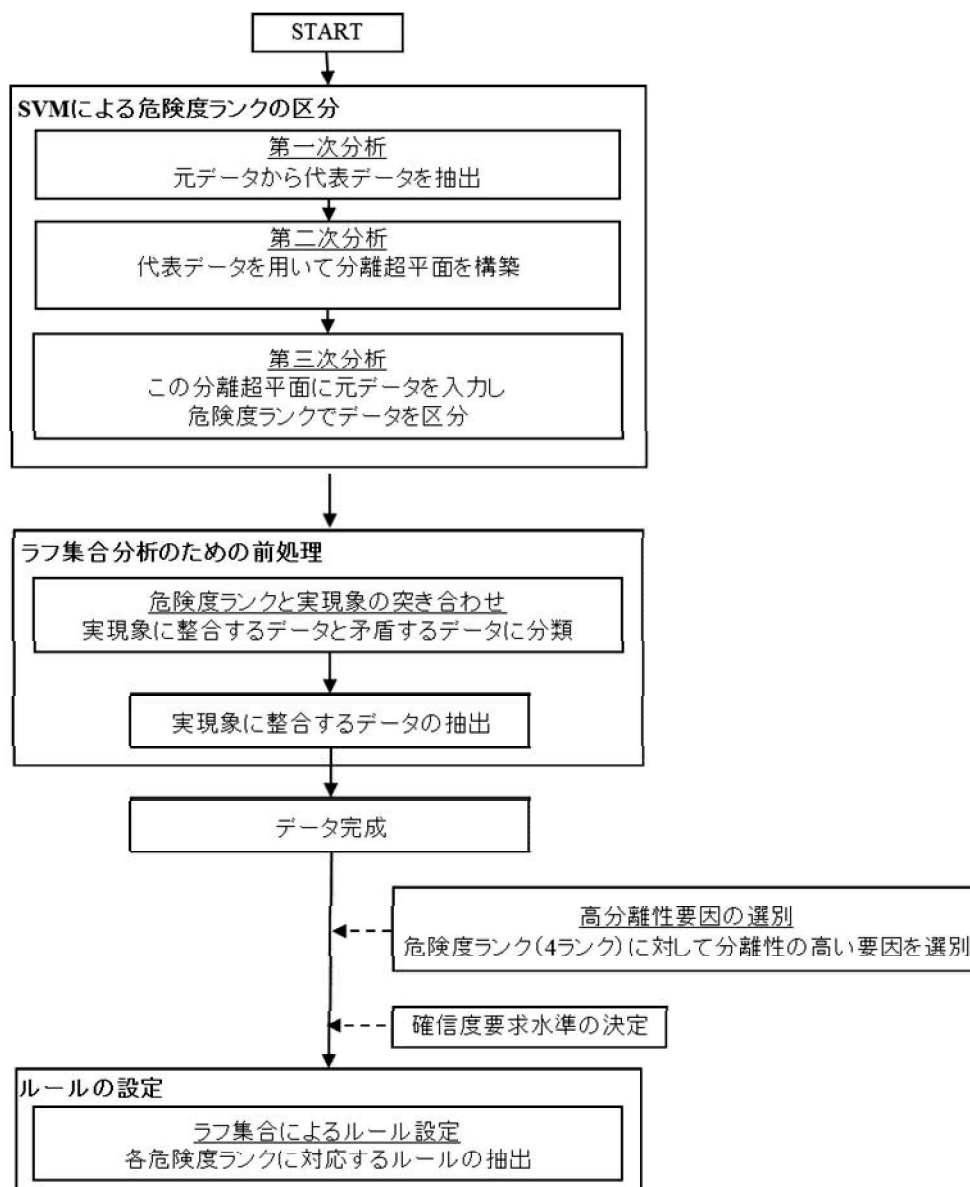


図-4.2 発生危険度に対応した評価ルールの設定手順

(1) SVMによる危険度ランクの区分

1) 代表データを用いた危険度ランク設定の流れ

ここでは、ノイズデータに引きずられない分離超平面を構築し、そこから危険度ランクを設定するまでの流れを整理する。

①一次分析（代表データの抽出）

代表データは、分析対象地域における典型的な災害発生・非発生事例を代表するデータであることから、災害発生データとしてはより危険度の高いもの ($f(x)$ 値が-1 よりも小さいもの)、非発生データとしてはより安全度の高いもの ($f(x)$ 値が1 よりも大きいもの) を抽出する⁵⁾ (図-4.3)。

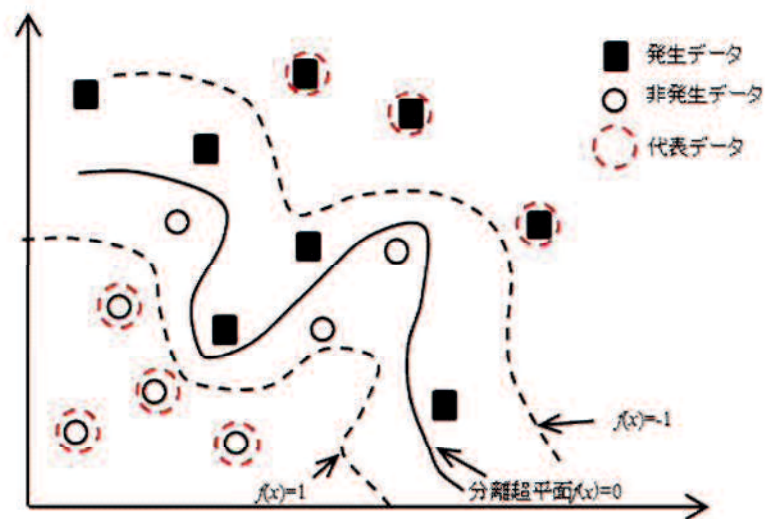


図-4.3 一次分析のイメージ図（代表データの抽出）

②二次分析（代表データによる分離超平面の構築）

二次分析では、一次分析で抽出された代表データのみを用いて、分離超平面を再構築する。ここで再構築された分離超平面は、ノイズデータに影響を受けない汎用的な分離超平面と言える。

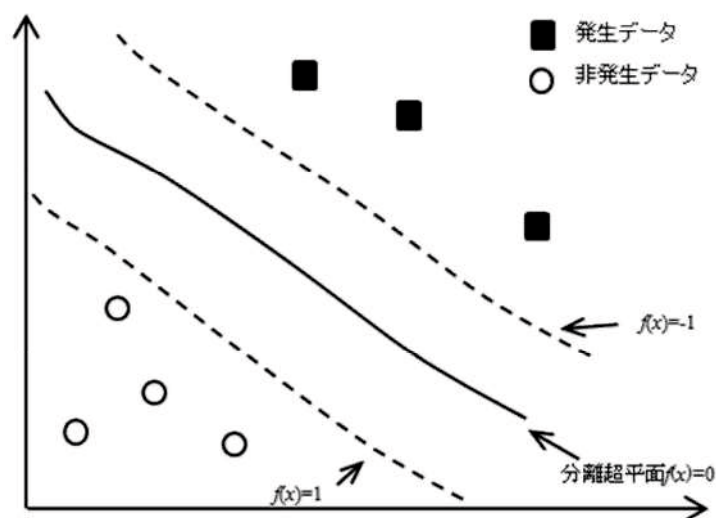


図-4.4 二次分析のイメージ図（代表データによる分離超平面の再構築）

③三次分析（危険度ランクの設定）

二次分析で再構築した分離超平面を基準として、対象地域の全データについて分離面からの距離 $f(x)$ 値を算出する。この $f(x)$ 値を各箇所の危険度の指標として用いる⁶⁾。これによりノイズデータに引きずられることなく（誤判別を許容）各箇所の危険度設定が可能となる。なお、危険度ランクの設定に関しては、3章と同様に①災害発生の可能性の高い危険領域のデータ群、②災害発生の可能性が比較的高いと考えられるデータ群、③災害発生の可能性が比較的低いと考えられるデータ群、④災害発生の可能性の低い安全領域のデータ群の4つのグループに区分した。

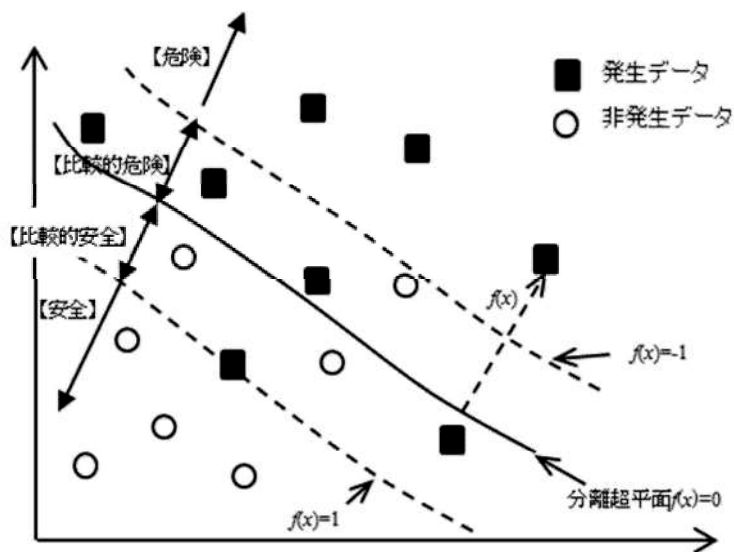


図-4.5 三次分析のイメージ図

(2) ラフ集合による危険度ランクに対応したルールの設定

1) ラフ集合の分析に用いるデータ

前節での検討に誤判別を許容することから、危険側に区分された非発生データ、安全側に区分された発生データといった実際の現象と区分した危険度ランクが矛盾するデータは少なからず確認される。これらのデータは「危険」な状況でありながらも局所的かつ特例的な条件により災害が起らなかった事例や「安全」な状況でありながらも局所的かつ特例的な条件により災害が起こった事例であると考えられる。

表-4.1(1)および表-4.1(2)に具体的な事例データの一例を示す。表-4.1(1)に示す箇所については、代表データを用いて求めた分離超平面からの距離 $f(x)$ 値により【比較的危険】なランクに区分したデータである。各要因のカテゴリ値を見ると、発生データであるID:239に比べて、非発生データであるID:251, 252の方が大きくなっている。これは、前述の「危険」な状況でありながらも局所的かつ特例的な条件により災害が起らなかった事例と言える。

一方、表-4.1(2)に示す箇所は、【比較的安全】なランクに区分したデータである。非発生データである ID:155 に比べて、発生データである ID:114, 148 の方がカテゴリ値が小さくなっている。これは、前述の「安全」な状況でありながらも局所的かつ特例的な条件により災害が起きた事例と言える。

このようなデータは、机上で抽出できる一般的な要因では捉えきれない要因が影響している可能性が高いと考えられる。

表-4.1(1) ノイズデータの一例（危険側）

箇所ID	カテゴリ値							教師値	f(x)	危険度区分
	最急溪床勾配	流域幅	流域形状比	溪流方位	流域最大傾斜	地質小分類	実効雨量(T=72hr)			
239	3	3	2	3	2	6	5	発生	-0.61658626	比較的危険
251	4	4	2	3	2	6	5	非発生	-0.08686971	比較的危険
252	4	4	3	3	2	6	5	非発生	-0.71544391	比較的危険

表-4.1(2) ノイズデータの一例（安全側）

箇所ID	カテゴリ値							教師値	f(x)	危険度区分
	最急溪床勾配	流域幅	流域形状比	溪流方位	流域最大傾斜	地質小分類	実効雨量(T=72hr)			
114	2	2	2	4	3	2	1	発生	0.224180803	比較的安全
148	3	2	1	4	3	2	1	発生	0.361423874	比較的安全
155	5	2	2	4	3	2	1	非発生	0.999999857	比較的安全

 : ノイズデータ

ラフ集合を用いた既往研究⁵⁾によると、ノイズデータを多く含むデータベースからルールを設定を行った場合、ノイズデータを説明するためにルールが増大し、結果として有効なルールを得ることができないことが報告されている。そのため、有効な評価ルールを設定にあたっては、ノイズデータを除外したデータを用いることが望ましい。そこで、ラフ集合の分析に用いるデータは、区分した危険度ランクと実現象の関係が一致するデータのみを対象とすることとした（図-4.6）。

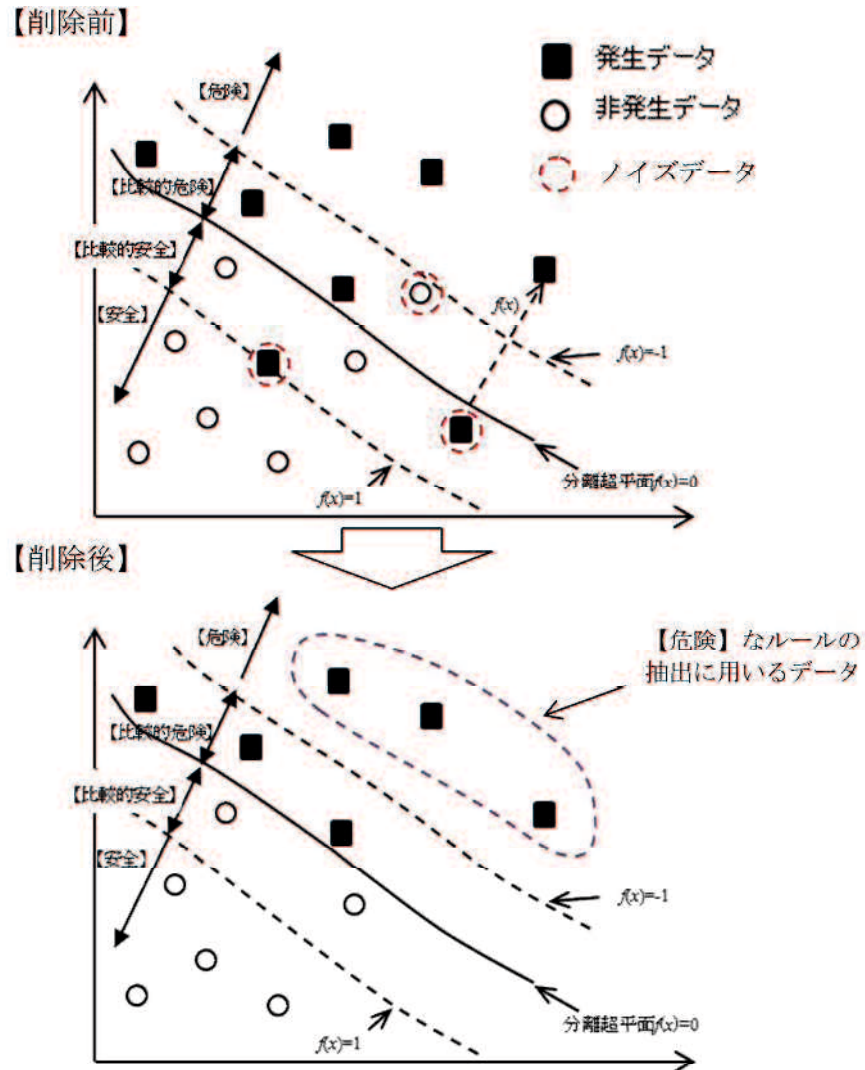


図-4.6 ラフ集合の分析に用いるデータ

2) 高分離性要因の選別

ラフ集合を用いた既往研究⁷⁾によると、汎用的かつ精度の高いルールを導き出すためには、対象とする現象に対して分離性が高い要因（以下、高分離性要因）を多数含む要因の組み合わせを用いてルールを設定することが有効であると報告されている。そのため、2章,3章と同様に高分離性要因の検討を行い、それらを多数含む要因の組み合わせからルールを設定を行うこととした。

3) 確信度の要求水準

ラフ集合を用いた分析では、一般的に整合度を高くするとデータの質の低下は防げるものの、現象を説明するために必要となる要因数が増える。そのため、結果としてルールが煩雑になる。一方、整合度を低くすると現象を説明するために必要となる要因数を減らせるため、ルールの簡素化が期待できる。ただしこの場合、データの質が低下することが懸念される。そこで、本研究ではデータの質の低下を防ぐため、整合度の要求水

準を100.0%とし、当該要求水準を満足する要因の組み合わせのうち、最小要因数であり、かつ高分離性要因を多数含む組み合わせを用いてルールの設定を行った。

また、汎用性が高いルールを設定するため、ラフ集合の分析結果から得られたルールを基に、ルール領域の拡張^{7), 8)}およびルールの組み合わせ⁷⁾を行った。ここで、組み合わせに用いる拡張ルールの要求水準には、既往研究⁷⁾と同様に確信度を用いる。ただし、確信度はサポートと相反関係にある。そのため、確信度を高く設定すると、多くの事例を説明するためには多くのルールを組み合わせることが必要となることや一部の事例にしか説明できない汎用性に乏しいルールが得られることが懸念される。一方、確信度を低く設定すると、サポートが向上し組み合わせに用いるルール数を低減することが可能となるが、実現象に対する整合性の低下が懸念される。これらのことから、組み合わせに用いる拡張ルールの要求水準は、ラフ集合の分析に用いたデータの特性を踏まえ、適切な確信度を設定する必要がある。

本研究においては、ラフ集合の分析対象データは、SVMを用いて設定した分離超平面に基づいた危険度ランクであり、当該分離超平面の元データに対する判別精度は75.8%である。そのため、75.8%よりも高い要求水準は、分析対象データ自体の有する判別精度に対して過剰な水準であり、75.8%よりも低い要求水準は、分析対象データ自体の有する判別精度を劣化させることに繋がるものと考えられる。

これらのことから、組み合わせに用いる拡張ルールの要求水準は、危険度ランクの設定に用いた分離超平面の精度を踏まえ、確信度75.0%以上とした。

4.3 分析結果

前述のデータを用い、図-4.2 のフロー図に示す手順に従って、災害発生危険度に対応した評価ルールの設定を試みた。

4.3.1 危険度ランクの区分

(1) 代表データの抽出結果（第一次分析）

SVMの分析を行う際には、最適な解析を行うためのパラメータスタディが必要である。ソフトマージン法のSVMでは、誤判別の度合いを調整するCとガウシアンカーネルの半径（データの影響力）である r の2つのパラメータを設定する必要がある。本研究では、 $C=5, 10, 50, 100, 200, 300, 400, 500$ の8ケース、 $r=0.1, 0.5, 1, 2, 3, 4, 5$ の7ケースの計56ケースでパラメータスタディを行った。パラメータスタディでは、次に定義する的中率により各ケースにおいて構築される分離超平面の精度（いかに正確にデータ分類が行われるか）を検証する。

$$\text{的中率} = \frac{\text{教師値と } f(x) \text{ 値が整合するデータ数}}{\text{全データ数}} \times 100 \quad (1)$$

パラメータスタディの検討結果を表-4.2 および表-4.3 に示す。ここで、SVMによる危険度設定に用いるパラメータは、既往研究⁵⁾におけるパラメータの選定方法を参考に、的中率 100.0%のパラメータセットのうち、サポートベクターとなるデータ ($|f(x)| \leq 1$ となるデータ) 数が最も少なくなる $C=500$, $r=4$ の組み合わせとする (表-4.2, 表-4.3)。

表-4.2 SVM パラメータ検討結果一覧

		C							
		5	10	50	100	200	300	400	500
r	0.1	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	0.5	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	1	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	2	98.2%	99.3%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	3	90.5%	94.4%	99.1%	99.8%	100.0%	100.0%	100.0%	100.0%
	4	85.0%	88.7%	96.1%	96.8%	98.8%	99.5%	99.6%	100.0%
	5	81.8%	83.4%	91.0%	94.7%	96.3%	96.8%	98.2%	98.6%

代表データの抽出は、当該パラメータを用いて設定した $f(x)$ 値に基づいて行う。代表データは、分離超平面に近接するデータを削除したデータであることから、災害発生データとしてはより危険なもの ($f(x)$ 値が -1 よりも小さいもの)、非発生データとしてはより安全なもの ($f(x)$ 値が 1 よりも大きいもの) を抽出する。ここで、抽出する代表データの割合は、元のデータベースにおける発生・非発生データの割合を乱さない方法が最も有効である⁵⁾とされていることから、本研究でも同様の方法を採用することとした。

上記パラメータに基づく SVM 解析により算出した $f(x)$ 値に着目すると、 $f(x)$ 値が -1 未満となる発生データ数は 32 個であった (表-4.3)。それに対して $f(x)$ 値が 1 以上の非発生データは 260 個もあった。そのため、全データにおける発生・非発生データの比率 (150 : 417) の関係が保持できるよう、非発生データからは $f(x)$ 値が 1 より大きくなるデータのうち $f(x)$ が大きい方から 87 個を代表データとして抽出した。その結果代表データは 119 個となった。

表-4.3 完全分離可能なパラメータにおける $f(x)$ 値の分布

パラメータ		$f(x)$ 値カテゴリ毎のデータ個数				サポート ベクターの数
C	r	$f(x) < -1$	$-1 \leq f(x) \leq 0$	$0 < f(x) \leq 1$	$1 < f(x)$	
5	1	0	150	400	17	550
10	1	0	150	400	17	550
50	1	0	150	400	17	550
	2	4	146	231	186	377
100	1	0	150	400	17	550
	2	4	146	231	186	377
200	1	0	150	400	17	550
	2	4	146	231	186	377
	3	25	125	183	234	308
300	1	0	150	400	17	550
	2	4	146	231	186	377
	3	25	125	183	234	308
400	1	0	150	400	17	550
	2	4	146	231	186	377
	3	25	125	183	234	308
500	1	0	150	400	17	550
	2	4	146	231	186	377
	3	25	125	183	234	308
	4	32	118	157	260	275

(2) 代表データによる分離超平面の構築結果（第二次分析）

抽出した代表データを用いて分離超平面の構築を行った。なお、分離超平面の構築に用いるパラメータは、先の検討において最適パラメータとして選定した $C=500$, $r=4$ の組み合わせとした。

代表データより構築した分離超平面の精度を表-4.4 に示す。表より、先の検討において最適パラメータとして選定した $C=500$, $r=4$ の組み合わせを用いることにより、的中率は 75.8%と高い水準を示していることが確認された。この的中率は、SVM とラフ集合を用いてルールの設定を試みた既往研究^{4),5)}における累計確信度と同等であることから、代表データより構築した分離超平面を用いた場合であっても、元データの危険度を精度良く設定できるものと考えられる。

表-4.4 代表データより構築した分離超平面の精度

	データ数	教師値と $f(x)$ 値が 整合するデータ	的中率
代表データ	119	119	100.0%
元データ	567	430	75.8%

(3) 危険度ランクの区分 (第三次分析)

災害発生危険度に対応した評価ルールの設定に用いる教師値は、代表データより構築した分離超平面を用いて設定した危険度に基づき、「危険 ($f(x) < -1$)」、「比較的危険 ($-1 \leq f(x) < 0$)」、「比較的安全 ($0 \leq f(x) \leq +1$)」、「安全 ($1 < f(x)$)」の4つに区分した危険度ランクを用いる。

危険度ランクと土石流の発生率 (該当する全データのうち発生データが含まれる割合) との関係を図-4.7 に示す。図より、土石流の発生率は危険度ランクが上昇するにつれて増加していることが確認されたことから、区分した危険度ランクは実際の災害発生状況と一致する妥当なランクであると考えられる。また、これにより代表データを用いて構築した分離超平面より設定した危険度は元データにおける災害発生状況を精度よく表わすことが可能な有効な方法であることが明らかとなった。

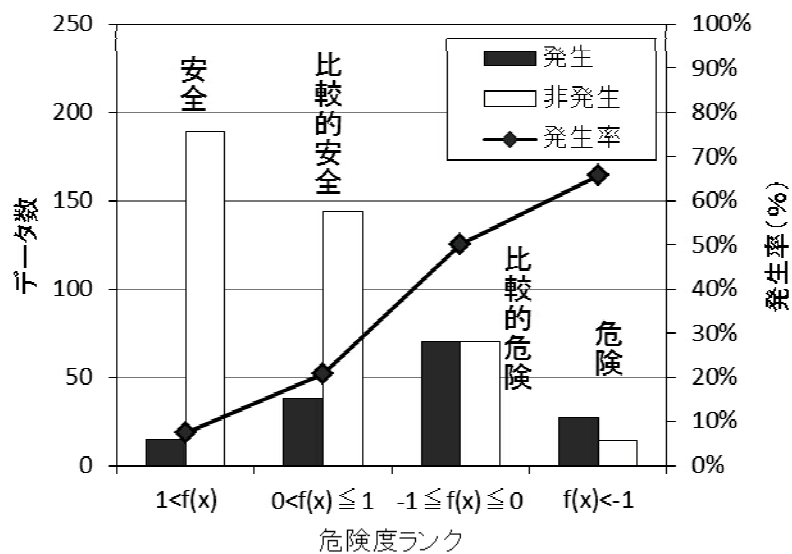


図-4.7 区分した危険度ランクと土砂災害の発生率の関係

4.3.2 危険度ランクに対応した評価ルールの設定

(1) 高分離性要因の検討

本研究では、土砂災害の発生危険度に対応した評価ルールを設定を行うことから、本研究における高分離性要因とは土石流の発生・非発生に対する分離性が高いことに加え、4つに区分した危険度ランクに対しても分離性が高い要因となる。

そこで、本研究では既往研究⁷⁾において土石流の発生・非発生に対して分離性が高い要因として示されている8要因（主溪流長、流域長、流域幅、溪流方位、流域最大傾斜、源頭部面積、地質区分、実効雨量（半減期72hr））を対象に、危険度ランクに対する分離性の検証を行った。ここで、危険度ランクに対する分離性が高い要因とは図-4.8に示す流域最大傾斜のようにカテゴリ値の増加にともない、「危険」が含まれる割合と「比較的危険」が含まれる割合の双方が単調増加する要因とする。流域最大傾斜と同様の傾向を示す要因を図-4.9および図-4.10に示す。図より、流域長、流域幅、流域最大傾斜、地質小分類の4要因が設定した危険度ランクに対して分離性が高い要因（以下、高分離性要因）であることが確認された。

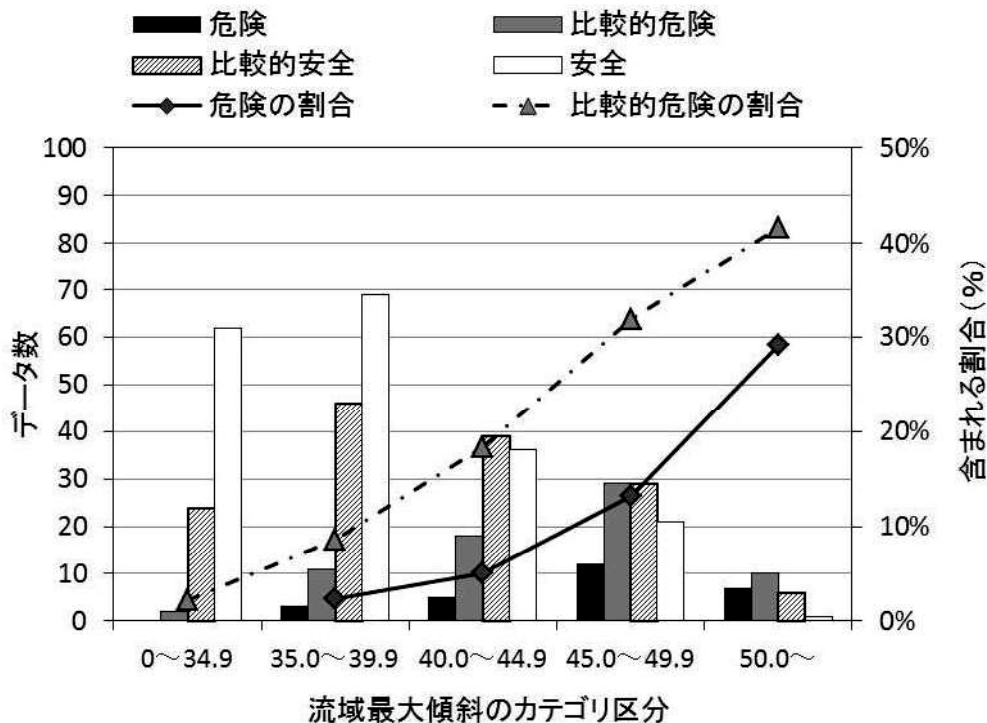


図-4.8 流域最大傾斜と危険度ランクの関係

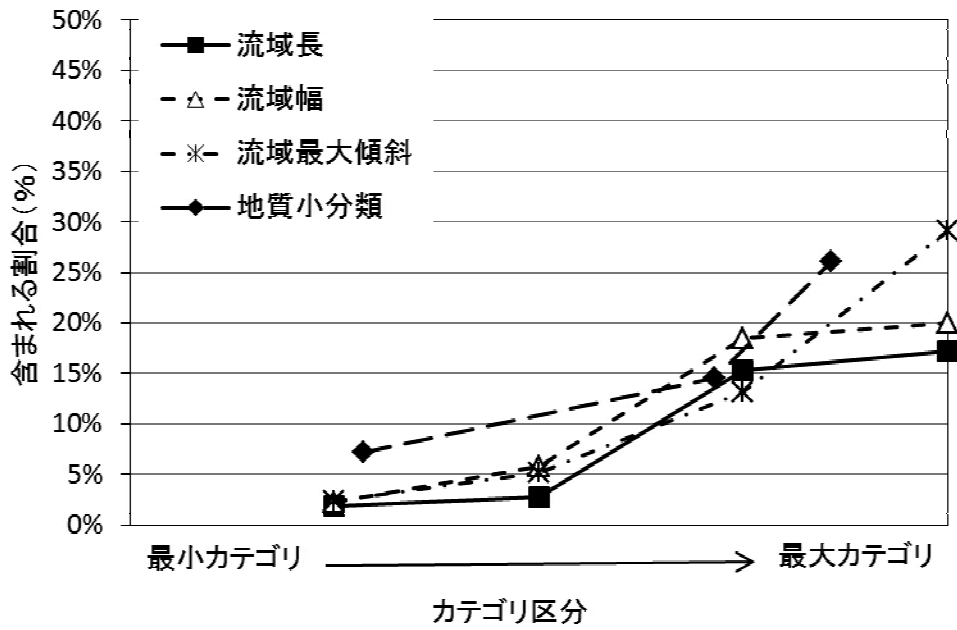


図-4.9 危険の割合が単調増加する要因

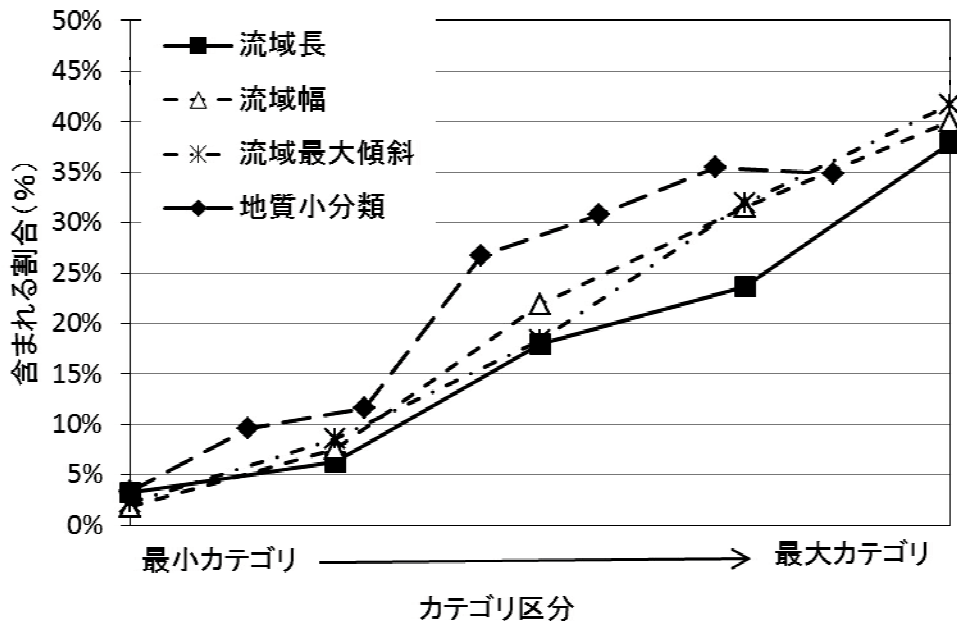


図-4.10 比較的危険の割合が単調増加する要因

4.3.3 ルールの抽出結果

危険度ランクに対応した評価ルールを元のデータベースに適用した場合の精度（以下、適用精度）を表-4.5に示す。ここで、「安全」もしくは「比較的安全」のルールに対しては、非発生データを整合データ、発生データを矛盾データと評価している。一方、「危険」もしくは「比較的危険」のルールに対しては、発生データを整合データ、非発生データを矛盾データと評価している。なお、表中においてルールを構成する要因のカテゴリ値は、各カテゴリに含まれる最大、最小値を記載しており、「↓」「↑」の表記は、ルール領域の拡張の概念よりそのカテゴリ値以下、以上を意味している。また、表中の「*」についてはどのようなカテゴリ値でも良いことを意味する。

表-4.5より、設定した評価ルールは、最急溪床勾配、流域幅、流域形状比、溪流方位、流域最大傾斜、地質小分類、実効雨量（半減期72hr）の7つの要因から構成されている。当該ルールの実現象に対する精度は78.7%を示しており、危険度の設定に用いた分離超平面の精度（75.8%）を若干上回る結果となっている。また、ルールの汎用性を示すサポートは100.0%となっていることから、当該ルールを用いることでS57.7長崎災害の事例の全データに対して約8割の精度で説明することが可能である。

設定したルール内容に着目すると、「安全」と評価できる溪流と「比較的安全」と評価できる溪流の特徴は比較的類似した結果となっており、「溪床勾配及び流域最大傾斜が緩くて流域幅が小さく、かつ安山岩質凝灰岩、輝石安山岩、第三紀層等が分布する北向きの溪流であり、かつ長期降雨が少ない地域」という特徴を有する溪流は危険性が低いと評価することができる。この中でも「北向きの溪流」という特徴は「安全」のルールにのみ出現していることから、このような特徴を持つ溪流については危険性は特に低いものと考えられる。また、流域最大傾斜と地質小分類に関しては「比較的安全」のルールよりも「安全」のルールの方がより安全側の条件となっており、「比較的安全」と「安全」の違いが明確に表れている。一方、危険性が高いと評価される溪流の特徴に着目すると、これらの溪流は「溪床勾配及び流域最大傾斜が急で流域幅の大きい溪流であり、かつ角閃石安山岩等が分布する地域の溪流」という特徴を有している。その中でも「角閃石安山岩が分布する地域」という特徴は、「危険」のルールにのみ出現していることから、このような特徴を持つ溪流については危険性が特に高いものと考えられる。また、流域最大傾斜と流域幅に関しては「比較的危険」のルールよりも「危険」のルールの方がより危険側の条件となっており、「比較的危険」と「危険」の違いが明確に表れている。

表-4.5 災害発生危険度に対応した評価ルール

ルール番号	ルール要因										適用結果					累計	
	最急 渓床勾配 (°)	流域幅 (km)	流域 形状比	溪流方位	流域 最大傾斜 (°)	地質小分類	実効雨量 (半減期 72hr) (mm)	判定	含まれる データ数 D1	整合 データ数 D2	矛盾 データ数 D3	残り 箇所数 D4	確信度 D5 = D2/D1	サポート D6 = D1/567	累計		
															確信度 D5 = D2/D1	サポート D6 = D1/567	
ルール31	*	0.19 ↓	*	*	*	結晶片岩類 安山岩質凝灰岩	300 ↓	安全	160	138	22	567	86.3%	28.2%			
ルール62	19.9 ↓	0.19 ↓	0.49 ↓	*	*	輝石安山岩	*	安全	176	152	24	391	86.4%	31.0%			
ルール47	*	0.29 ↓	*	*	*	結晶片岩類	270 ↓	安全	190	165	25	377	86.8%	33.5%			
ルール16	*	*	0.49 ↓	北	34.9 ↓	結晶片岩類	*	安全	196	170	26	371	86.7%	34.6%			
ルール61	14.9 ↓	*	*	北	39.9 ↓	結晶片岩類 安山岩質凝灰岩 輝石安山岩	*	安全	201	175	26	366	87.1%	35.4%			
ルール8	*	*	*	*	34.9 ↓	*	240 ↓	安全	203	177	26	364	87.2%	35.8%			
ルール6	9.9 ↓	0.19 ↓	*	*	*	*	*	安全	205	179	26	362	87.3%	36.2%			
ルール48	*	*	*	北	39.9 ↓	結晶片岩類	*	安全	207	181	26	360	87.4%	36.5%			
ルール34	14.9 ↓	*	0.49 ↓	*	*	*	270 ↓	安全	208	182	26	359	87.5%	36.7%			
ルール113	*	*	*	*	44.9 ↓	結晶片岩類 安山岩質凝灰岩 輝石安山岩 第三紀層 下平閃緑岩類	*	比較的 安全	424	346	78	143	81.6%	74.8%			
ルール73	*	0.29 ↓	*	*	*	結晶片岩類 安山岩質凝灰岩 輝石安山岩 第三紀層 下平閃緑岩類	*	比較的 安全	463	373	90	104	80.6%	81.7%			
ルール128	*	0.39 ↓	*	*	*	*	270 ↓	比較的 安全	475	382	93	92	80.4%	83.8%			
ルール97	14.9 ↓	*	*	*	*	結晶片岩類 安山岩質凝灰岩 輝石安山岩 第三紀層 下平閃緑岩類 菱刈安山岩	*	比較的 安全	479	386	93	88	80.6%	84.5%			
ルール135	*	0.19 ↓	0.49 ↓	*	49.9 ↓	結晶片岩類 安山岩質凝灰岩 輝石安山岩 第三紀層 下平閃緑岩類 菱刈安山岩	*	比較的 安全	482	389	93	85	80.7%	85.0%			
ルール161	15.0 ↑	*	*	*	50.0 ↑	角閃石安山岩 その他	*	危険	492	397	95	75	80.7%	86.8%			
ルール171	20.0 ↑	0.20 ↑	*	*	*	菱刈安山岩 角閃石安山岩 その他	*	危険	498	403	95	69	80.9%	87.8%			
ルール189	*	0.10 ↑	*	*	40.0 ↑	*	*	比較的 危険	567	446	121	0	78.7%	100.0%			

※安全もしくは比較的 safety に該当するルールの場合、整合データ：非発生データ、矛盾データ：発生データ

※危険もしくは比較的危険に該当するルールの場合、整合データ：発生データ、矛盾データ：非発生データ

「危険」なルール 161 に該当する箇所を対象に現地の状況確認を実施した（写真-4.1～写真-4.5）。現地の状況としては、ルールに示されている通り、溪床勾配や流域内の傾斜が非常に急であることが特徴であった。また、小規模ながらも流域内に崩壊箇所や大きな転石も存在しており、雨の規模により大きな崩壊が発生した場合には、土石流化の可能性も十分あると考えられる。



写真-4.1 流域内の状況



写真-4.2 転石や堆積の状況



写真-4.3 斜面の崩壊状況



写真-4.4 斜面の崩壊状況



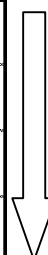
写真-4.5 転石の状況

また、各評価ルールに該当するデータ数と土石流の発生率との関係（表-4.6）に着目すると、「安全」と評価された箇所では土石流の発生率が12.5%と低いが、危険側に区分されるにつれて土石流の発生率は上昇しており、最も危険側となる「危険」と評価される箇所では土石流の発生率が87.5%にまで上昇している。そのため、設定した評価ルールは、実際の災害の発生・非発生に対して整合性が高く、対象箇所の危険度ランクを把握できる有効なルールであると考えられる。

表-4.6 各評価ルールに該当するデータ数と土石流の発生率との関係

危険度ランク	該当ルール数	①	②	③=①+②	①/③
		発生データ数	非発生データ数	該当データ数	災害発生率
安全	9	26	182	208	12.5%
比較的安全	5	67	207	274	24.5%
比較的危険	1	43	26	69	62.3%
危険	2	14	2	16	87.5%

災害発生率は上昇



本研究で提案する手法の有用性を確認するため、3章の検討との比較を行った。比較結果の概要を表-4.7に示す。

3章では7つの要因を用いた24個のルール（危険：3，比較的危険：4，比較的安全：5，安全：12）が抽出されており、ルールの適用精度は確信度90.0%，サポート56.6%であった。一方、本章では7つの要因を用いた17個のルール（危険：2，比較的危険：1，比較的安全：5，安全：9）となっており、ルール数の低減が図れた。確信度は78.7%と3章の結果より低下するものの、サポートは100.0%となっており、本章の検討テーマであったルールの汎用性については大きく向上させることができた。

表-4.7 既往研究との比較

		3章での結果(危険度ランク) [SVM+ラフ集合]	本章での結果(危険度ランク) [SVM+ラフ集合]	
データ使用	データ数	[長崎市] 567溪流 (発生150 非発生417)		
	要因	地形・地質16要因, 降雨 2要因		
SVMの利用	パラメータ	$C=500, r=4$	$C=500, r=4$	
	ルール抽出に用いるデータ	【全データを用いる】 危険度ランクを設定し、 $f(x)$ の値に基づき4つのランクに区分する。その全データを利用	【危険度ランクと実現象が整合するデータ】 代表データを用いて分離超平面を構築 ↓ 元データに対する $f(x)$ を算出 ↓ $f(x)<0$ となる発生データおよび $f(x)>0$ となる非発生データを抽出	
ラフ集合によるルール抽出	整合度の要求水準	95.0%以上	100.0%	
	確信度の要求水準	発生(危険, 比較的危険): 67.5%以上 非発生(安全, 比較的安全): 92.4%以上	発生(危険, 比較的危険): 75.0%以上 非発生(安全, 比較的安全): 75.0%以上	
	データ数	危険: 32 比較的危険: 118 安全: 260 比較的安全: 157	危険: 27 比較的危険: 70 安全: 189 比較的安全: 144	
	抽出された重要要因	最急渓床勾配	最急渓床勾配	最急渓床勾配
		溪流長		
		流域幅		流域幅
		流域形状比		流域形状比
		溪流方位	溪流方位	溪流方位
		流域最大傾斜	流域最大傾斜	流域最大傾斜
		源頭部面積		
地質小分類		地質小分類	地質小分類	
実効雨量(半減期72hr)	実効雨量(半減期72hr)	実効雨量(半減期72hr)		
ルール数	危険: 3 比較的危険: 4 安全: 12 比較的安全: 5	危険: 2 比較的危険: 1 安全: 9 比較的安全: 5		
適用精度	累計確信度	90.0% (289/321)	78.7% (446/567)	
	累計サポート	56.6% (321/567)	100.0% (567/567)	

4.4 まとめ

本章における主要な結論を以下に示す.

- ①汎用性の高いルールを設定するためには、危険度ランクと要因の組み合わせの関係はできるだけ矛盾が少ないことが重要となる. そのため、元データから代表的な危険データ・安全データを抽出し、それを元に設定した危険度から危険度ランクを区分することが重要である.
- ②汎用性の高いルールを作成するポイントとして、区分した4つの危険度ランクに対して分離性の良い要因を選別した. その結果、流域長、流域幅、流域最大傾斜、地質小分類の4要因が設定した危険度ランクに対して分離性が高い要因として抽出された.
- ③抽出されたルールとしては、「溪床勾配及び流域最大傾斜が緩くて流域幅が小さく、かつ安山岩質凝灰岩、輝石安山岩、第三紀層等が分布する北向きの溪流であり、かつ長期降雨が少ない地域」という特徴を有する溪流は危険性が低いと評価することができる.
- ④危険性が高いと評価される溪流は、「溪床勾配及び流域最大傾斜が急で流域幅の大きい溪流であり、かつ角閃石安山岩等が分布する地域の溪流」という特徴を有している.
- ⑤最終的に設定された17個のルールで、S57.7長崎災害の事例の全データ（サポート100.0%）に対して約8割の精度（確信度）で説明可能となった. このことから本研究によるルール抽出方法は、高い汎用性と精度を有するルールを抽出できることが確認できた.

参考文献

- 1) 小林央宜, 杉原成満, 大石博之, 荒木義則, 古川浩平: 既存データの活用による災害発生危険に応じた土石流危険溪流の評価ルールの設定, 土木学会論文集 F6, Vol.68, No.1, pp.1-12, 2012.
- 2) Cortes C. and Vapnik, V.: Support vector networks, *Machine Learning*, Vol.20, pp.273-297, 1995.
- 3) 河野浩之: データベースからの知識発見の現状と動向, 人工知能学会誌, Vol.12, No.4, pp.497-504, 1997.
- 4) 杉原成満, 小林央宜, 川崎雄大, 篠崎嗣浩, 大石博之, 古川浩平: 昭和57年7月長崎豪雨における土石流の発生・非発生ルールの抽出, 土木学会論文集 F, Vol. 66, No. 4, pp.593-598, 2010.

- 5) 篠崎嗣浩, 森田真人, 大石博之, 古川浩平: SVM とラフ集合を用いた土石流の発生・非発生ルールの作成方法に関する研究, 土木学会論文集 F, Vol. 65, No.4, pp.448-460, 2009.
- 6) 大石博之, 小林央宜, 尹禮分, 田中浩一, 中山弘隆, 古川浩平: サポートベクターマシンによる対策工効果を考慮した斜面災害危険度の設定, 土木学会論文集 F, Vol.63, No.1, pp.107-118, 2007.
- 7) 竹本大昭, 杉原成満, 佐藤丈晴, 荒木義則, 中山隆弘, 水山高久, 古川浩平: ラフ集合を用いた土石流発生・非発生ルールの抽出と危険度の設定手法に関する研究, 砂防学会誌, Vol.57, No.2, pp.4-15, 2004.
- 8) Greco, S., Slowinski, R. and Stefanowski, J.: An Algorithm for Induction of Decision Rules Consistent with the Dominance Principle, Lecture Notes in Artificial Intelligence, Vol.2005, Springer-Verlag, Berlin, pp.304-313, 2001.

第5章 今後の防災事業計画への利用に関する一提案

5.1 緒言

第4章では、危険度ランクごとの汎用的なルールの設定に取り組んだ。その結果、既往研究¹⁾と同等の精度を維持した17個のルールで対象地域すべての箇所を特定できる汎用的なルールの抽出が可能²⁾となった。また、ここで得られたルールは4つの危険度ランク（「危険」、「比較的危険」、「比較的安全」、「安全」）に応じたものであることから、その有用性を大きく向上させることができた。例えば、将来的なハード対策の実施候補の絞り込みとして、最も上位ランクの「危険」に該当する箇所を優先することで効率的な防災事業の推進が可能と考えられる。逆に「安全」に該当する箇所に対しては調査実施の優先度を下げることが考えられる。

そこで本章では、抽出された各ルールに該当する箇所の持つ危険度に注目し、ルール毎の危険度の差異について分析した。これにより、特に危険なグループ（あるいは安全なグループ）を更に絞り込むことが可能となる。

また、本研究での一連の分析手法を活用した防災対策優先箇所の特定フローを検討し、本研究成果の実用化に向けた提案を実施した。

5.2 ルールに該当する箇所の危険度の確認

前章で抽出されたルール（表-4.4）に該当する箇所の危険度に着目し，SVMで設定した危険度（ $f(x)$ 値）の確認を実施した結果を表-5.1および図-5.1に示す。

ルールに該当するデータの $f(x)$ 値の平均値（ランク別）を確認すると、「危険」、「比較的危険」、「比較的安全」、「安全」の順に数値が大きくなっており，危険度ランクの傾向と合致していた。また個々のルールの平均値を確認すると，ルール161が平均値で-1.212と最も危険側の評価となっており，このルールに該当する全10箇所については，対象地域の中でも特に危険なグループであると考えられる。ここで絞り込まれた10箇所こそが，最優先箇所として詳細調査や対策検討を行うべき箇所と考えることができる。

一方，安全側のルールに着目すると，ルール31の $f(x)$ が1.504と最も高くなっており，このルールに該当する160箇所は，「安全」や「比較的安全」と評価される14個のルールに該当する箇所の中でも特に危険性が低いものと言える。このような箇所については，調査実施の優先度を下げることが可能と考えることができる。

表-5.1 各ルールに該当する箇所の $f(x)$ 値の整理

ランク	ルール	$f(x)$ 値		$f(x)$ 値		
		平均値 (ランク別)	データ数	最少	平均値 (個別)	最大
危険	ルール161	-1.078	10	-2.402	-1.212	-0.784
	ルール171		6	-2.430	-0.944	0.472
比較的危険	ルール189	-0.620	69	-1.900	-0.620	1.158
比較的安全	ルール135	0.240	3	-0.266	0.135	0.726
	ルール128		12	-1.281	0.206	1.103
	ルール113		216	-2.205	0.235	2.106
	ルール73		39	-1.000	0.241	1.944
	ルール97		4	-1.326	0.384	1.992
安全	ルール48	1.200	2	-0.461	0.525	1.510
	ルール61		5	-0.528	0.923	1.998
	ルール62		16	-0.136	1.191	2.515
	ルール47		14	0.699	1.231	1.812
	ルール6		2	1.095	1.298	1.500
	ルール8		2	1.240	1.304	1.367
	ルール16		6	0.465	1.332	2.137
	ルール34		1	1.494	1.494	1.494
	ルール31		160	-1.003	1.504	2.717

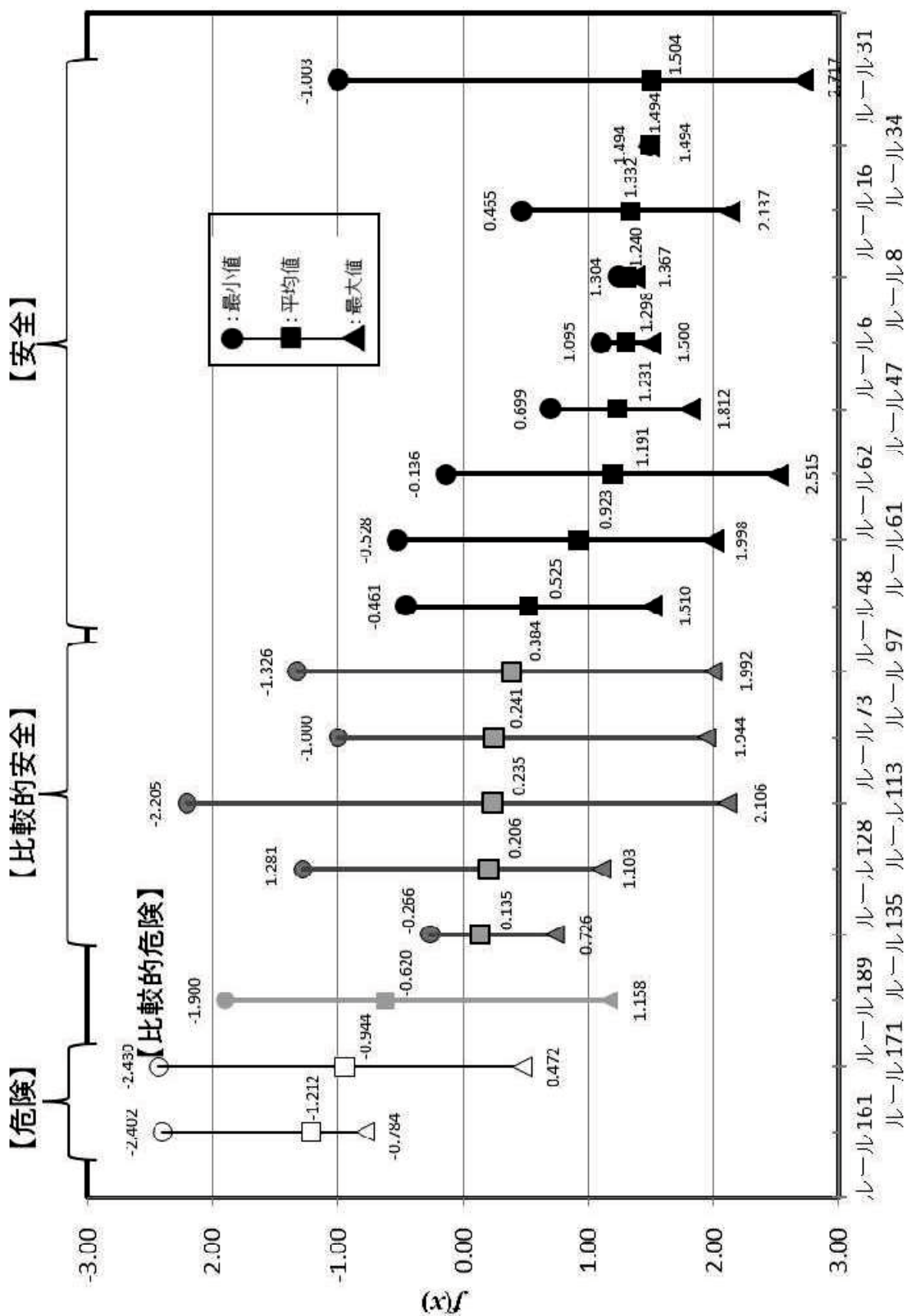


図-5.1 各ルールに該当する箇所の $f(x)$ 値の確認結果図

5.3 危険度評価システムの提案

本研究²⁾で検討したSVMによる危険度の設定とラフ集合を用いた危険度ランクに応じた評価ルールの設定の一連の検討方法を活用することにより、危険度評価システムの構築が期待できる。図-5.2に想定される危険度評価システムのフローを示す。また、次項に詳細な説明を示す。

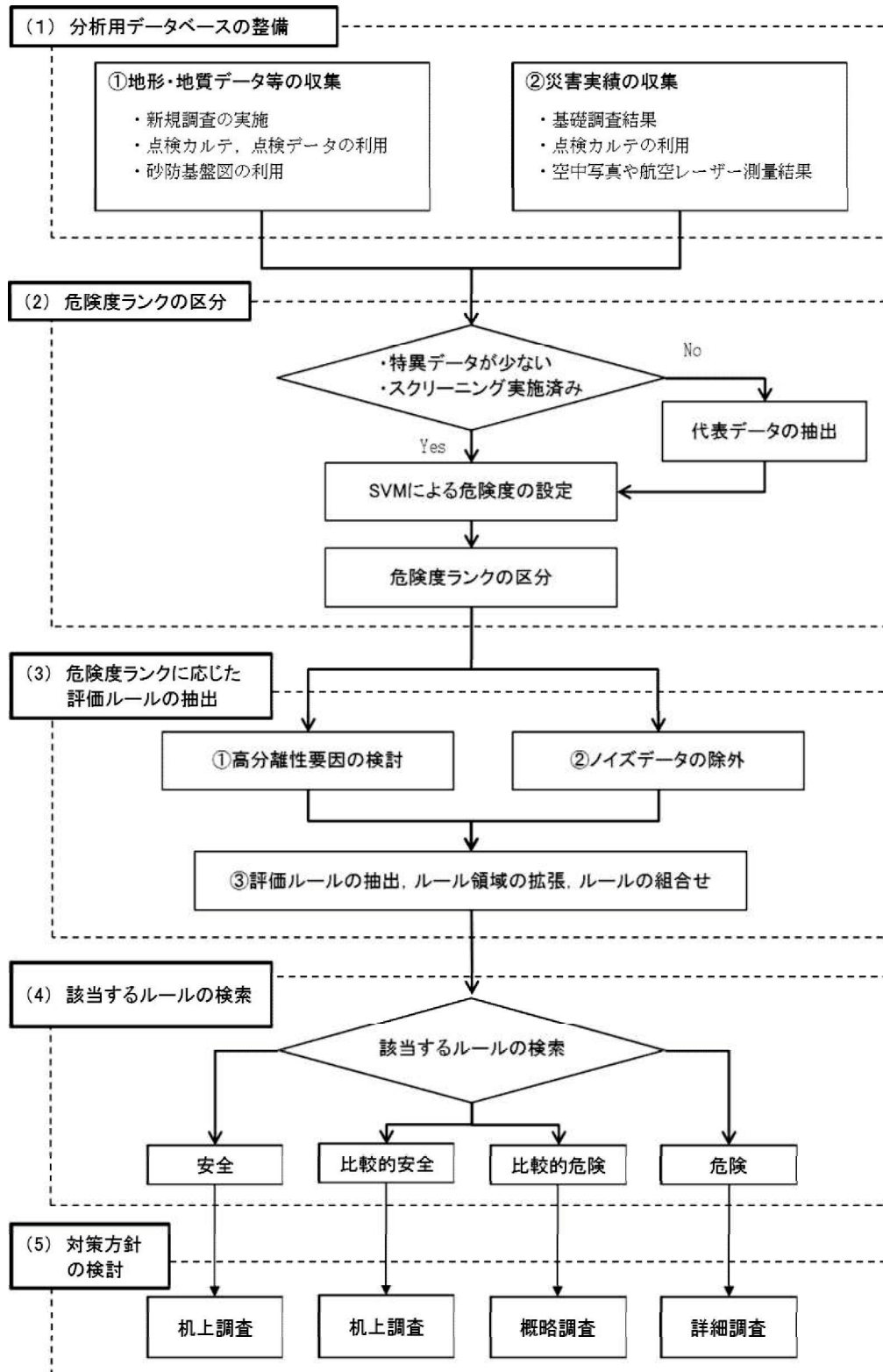


図-5.2 提案する危険度評価システムのフロー

(1) 分析用データベースの整備

① 地形・地質データ等の収集

まず、対象とする災害の危険性がある箇所に関する地形・地質等の情報を収集する。これは、新規に調査を行い必要な情報を収集しても良いし、既存の土石流危険渓流調査カルテのデータや道路防災総点検データ等がある場合は、これらを利用することも可能である。また、近年では、土砂災害防止法に基づく基礎調査が各自治体で実施されており、この動きは、平成26年8月豪雨による広島での災害を契機にそのスピードが増している状況である。この調査では砂防基盤図と呼ばれる数値地図情報が活用されるが、地形データに関しては、この砂防基盤図を利用することで容易に整理が可能である。

一方、地質区分に関しては、産業技術総合研究所 地質調査総合センターが公開しているシームレス地質図が入手可能である。これと流域のデータとをGIS上で重ね合わせることで、容易に区分可能である。

② 災害実績データの収集

地形・地質等のデータと併せて、過去の災害実績データの収集が必要となる。ここでは、基礎調査における災害履歴調査結果やカルテ等に記載の情報から抽出することも可能である。また、近年は比較的規模の大きな災害後には、空中写真の撮影や航空レーザー測量を実施することが多い。これらの資料から地形を判読する方法もある。

これらの地形・地質などの要因データと災害実績データを対応させ、データベースと整備することで、今後の分析の基礎資料とする。

(2) 危険度ランクの区分

次いでこれらのデータベースを用いてSVMにより危険度の設定を行う。本研究²⁾では対象地域における典型的な災害発生・非発生事例を代表するデータ（以下、代表データ）から構築した分離超平面に基づき危険度の設定を行ったが、特異データが少ないデータ群や予めスクリーニングが行われているデータ群を利用するのであれば対象とするデータ全てから構築した分離超平面に基づいて危険度を設定しても良いと考えられる。設定した危険度に対して危険度ランクの設定を行う。このランクについては、SVMの特徴を活かし、分離超平面 ($f(x)=0$) を境界として、その分離超平面を構成するのに寄与するサポートベクター ($|f(x)| \leq 1$) およびその外側のデータ ($|f(x)| \geq 1$) とに区分することが有効である。

(3) 危険度ランクに応じた評価ルールの抽出

代表データで構築された分離超平面に対し、代表データ抽出前の元データを再度入力する。ここからラフ集合を用いて当該危険度ランクに応じた評価ルールの設定を行う。この際、精度および汎用性を確保したルールを抽出するためには、以下の3点に関して

検討を実施することが有効である。

①高分離性要因の検討

ラフ集合を用いてルール抽出では、対象とする現象に対して分離性の高い要因を多数含む要因の組合せを採用することが有効³⁾である。そこで、カテゴリ値の増加に伴い、「危険」および「比較的危険」にランクされるデータの割合が単調増加する要因を高分離性要因とする。

②ノイズデータの除外

ノイズデータを多く含むデータベースからルールを抽出すると、その局所的かつ特例的なデータを説明するために、汎用性の低いルールが増大してしまう恐れがある。そこで、SVMを用いて設定した危険度ランクと実現象の関係が一致するデータのみをルール抽出に用いることとする。

③評価ルールの抽出、ルールの拡張、ルールの組合せ

ラフ集合で抽出される個々のルールでは、説明できる箇所が少ないという課題がある。この課題に対しては、ルール領域の拡張⁴⁾、ルールの組合せ³⁾が有効であることから、同様の考え方を用いて、精度と汎用性の高いルールの抽出を行う。

(4) 該当するルールの検索

この一連の流れによって得られる評価ルールに該当する箇所を検索することで、同時に危険度ランクを把握することが可能となる。この時、前節で検討したように、抽出したルールに該当するデータの危険度 $f(x)$ を確認することで、ルール毎に危険度を設定する事も可能である。なお、この作業については、分析に用いた母集団はもちろん、任意の危険箇所データに対しても検索が可能である。

(5) 対策方針の検討

危険度ランクに応じて、今後の対策方針の検討を実施する。ここでは一例として、以下を提案する。

- ・「危険」なルールに該当

⇒対象地域で最も危険性の高い事から、ハード対策の検討実施のための詳細調査を最優先で実施する候補とする。

- ・「比較的危険」なルールに該当

⇒将来的なハード対策を見据えた上で、現状把握のための概略調査を実施する。

- ・「比較的安全」、「安全」に該当

⇒安全側と判定されている箇所については、現地調査の優先度は低いと考え、机上

調査を基本とする。最新の地形データ等が入手できた場合には、机上調査を実施し、場合によっては現地調査を行う。

5.4 まとめ

本章では抽出された各ルールに該当する箇所を持つ危険度に注目し、ルール毎の順位を検討することを試みた。本章における主要な結論を以下に示す。

- ①抽出したルールに該当する箇所の危険度 $f(x)$ を確認することで、ルール毎の順位を確認することが可能である。
- ②「危険」、「比較的危険」に該当するルールの中でも、ルール 161 の危険度 $f(x)$ が最も小さいことから、特に危険なルールとして評価することができる。このルールに該当する箇所は、10 箇所のみであり詳細調査や対策検討を行うべき最優先箇所として絞込みが出来たことから、効率的な事業推進に寄与すると言える。
- ③「安全」、「比較的安全」に該当するルールの中でも、ルール 31 の危険度 $f(x)$ が最も大きいことから、特に危険性の低いルールとして評価することができる。このルールに該当する箇所は、160 箇所あり全体の約 3 割の箇所について、詳細調査を見送るなど優先度を落として評価することが可能となった。
- ④本研究で提案する手法を用いることにより、任意の溪流に対しても調査データ等から該当するルールを検索し、同時に危険度ランクを把握することが可能となる危険度評価システムの構築が期待できる。

参考文献

- 1) 杉原成満, 小林央宜, 川崎雄大, 篠崎嗣浩, 大石博之, 古川浩平: 昭和 57 年 7 月長崎豪雨における土石流の発生・非発生ルールの抽出, 土木学会論文集 F, Vol. 66, No. 4, pp.593-598, 2010.
- 2) 小林央宜, 杉原成満, 大石博之, 荒木義則, 古川浩平: 既存データの活用による災害発生危険に応じた土石流危険溪流の評価ルールの設定, 土木学会論文集 F6, Vol.68, No.1, pp.1-12, 2012.
- 3) 竹本大昭, 杉原成満, 佐藤丈晴, 荒木義則, 中山隆弘, 水山高久, 古川浩平: ラフ集合を用いた土石流発生・非発生ルールの抽出と危険度の設定手法に関する研究, 砂防学会誌, Vol.57, No.2, pp.4-15, 2004.
- 4) 佐藤丈晴, 河野泰典, 荒木義則, 中山弘隆, 水山高久, 古川浩平: 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, 土木学会論文集, No.721, VI-57, pp.13-25, 2002, 12.

第6章 結論

6.1 本研究のまとめ

本研究では、自然災害リスクへの取組みとして、効率的な事業推進を行うために、SVMによる災害の危険度ランクを設定し、ラフ集合を用いることでその危険度ランクに対応した評価ルールを明示することに取り組んだ。以下に各章の要旨と結論を示す。

第1章では本研究の背景を述べた上で、既往の危険度評価と災害の発生・非発生に関する素因・誘因の特定手法の現状を整理し、それらの手法に関する既往の研究成果と問題点を述べると共に、本論文の構成とその概要を述べた。

第2章では、対象地域における典型的な災害発生・非発生事例を代表するデータをSVMの利用により抽出し、この代表データを用いてラフ集合によるルールを抽出するという既往研究での手法の妥当性検証を行った。ここでは、他の地域の災害事例として、昭和57年の長崎豪雨における土石流災害を対象とし、この手法を適用し、その汎用性についての検証を行った。以下に主要な結論を示す。

- 1) 本研究で抽出したルールでの分析の結果、S57.7長崎豪雨災害における土石流発生・非発生事例の84.1%を、わずか7つのルールで説明できる非常に汎用性の高いルールを得ることができた。
- 2) 抽出したルールの確信度は78.8%と既往研究に比べやや劣る結果となったものの、全データの84.1%を説明できる汎用性の高いルールであることを勘案すると、精度としては十分な水準を有しているものと考えられる。
- 3) 抽出されたルールとしては、「溪床勾配が緩くて流域の規模が小さく、かつ結晶片岩類もしくは安山岩質凝灰岩、結晶安山岩が分布する地域の溪流」では土石流は発生していない。逆に「溪床勾配が急で流域の規模が大きく、かつ変朽安山岩もしくは角閃石安山岩が分布する地域」で土石流が発生したという特徴を掴んだ。
- 4) 篠崎らの提案するルール抽出方法を活用することで、要因やルール数を大きく削減しながらも、現象に対する高い精度と汎用性を有したルールを抽出できることが確認できた。

以上のように、第2章では他の地域においても、既往研究で提案されているSVMとラフ集合を組み合わせた評価ルールの抽出手法が有効であることが示される結果となり、手法としての汎用性を確認することができた。しかしながら、抽出されたルールは「発生」、「非発生」に関する2区分のみであり、対象箇所の持つ危険度に関しては考慮されていないことが課題として残っており、第3章においてこの課題について取り組むこ

ととした。

第3章では、前章で与えられた課題である対象箇所を持つ危険度を考慮した評価ルールの抽出方法の確立に主眼を置き、その抽出方法について論述した。

本来、土石流の災害発生や非発生に関する特徴は、対象箇所を持つ危険度に応じて差異があることが予想される。そのため、危険度に対応した類型的なルールを特定することができれば、優先的に防災事業を実施すべき箇所を容易に選定することが可能となる。そこで、SVMの基本機能を用いて、①危険、②比較的危険、③比較的安全、④安全の4つのグループに区分し、それぞれに対応したルールを検討することを試みた。以下に、第3章の主要な結論を列挙する。

- 1) 区分した4つの危険度ランクに対応したルールを抽出することを試みた結果、最急渓床勾配、溪流長、溪流方位、流域最大傾斜、源頭部面積、地質小分類、実効雨量（半減期72hr）の7要因で構成される計24個のルール（危険：3、比較的危険：4、比較的安全：5、安全：12）を抽出することができた。
- 2) 危険側のルール（危険、比較的危険）に該当する箇所の発生率は、70.0%以上と高い値を示しており、実現象に相違ない結果となっている。
- 3) 危険度ランクで「比較的安全」に該当する箇所は、全部で157箇所存在したが、同ランクに対応したルールでは、そのうち7箇所しか説明することができていない。
- 4) 今回抽出したルールは、精度こそ90.0%と高いが、そのルールで説明できる範囲が非常に狭く（サポート56.6%）汎用性の低いものとなってしまった。

以上のように、第3章では、既往研究と同程度の精度およびサポートで危険度ランクに対応した評価ルールの設定が可能であった。しかしながら、汎用性の高いルールが抽出できているとは言えず、その原因に対する更なる検討が必要となり、第4章においてこの課題について取組むこととなった。

第4章では、前章で抽出した危険度ランクに対応した評価ルールの汎用性向上に対する検討を実施し、その検討内容に関して論述した。

汎用性の低いルールとなった原因として、危険度評価の段階で、SVMの優れた判別能力を利用するために、データを災害発生・非発生の2群に完全分離する分離超平面（危険度評価の基準となる面）を採用したことにあると考えた。母集団のデータを完全分離させた場合の分離超平面は、一般の傾向とは異なる特殊な事例を含んだまま判別してしまうことになる。このようなデータを含んだ形でルールの抽出を行うと、局所的な条件の組合せのルールができることにより、確信度は向上するが、そのルールの汎用性は低

下してしまうことが予測される。

そこで、ノイズデータに引きずられない分離超平面の構築のために、典型的な災害発生・非発生事例である代表データの抽出を行い、そのデータを用いて再構築した分離超平面に対して危険度ランクを設定し、汎用的なルールを抽出することに取り組んだ。以下に第4章の主要な結論を列挙する。

- 1) 汎用性の高いルールを設定するためには、危険度ランクと要因の組み合わせの関係はできるだけ矛盾が少ないことが重要となる。そのため、元データから代表的な危険データ・安全データを抽出し、それを元に設定した危険度から危険度ランクを区分するための分離超平面を得るものとした。
- 2) 代表データから得られた分離超平面を利用し、代表データ抽出前の元データを危険度ランクに対応した4つのグループに分類した。各危険度ランクに該当するデータグループからそれぞれルールの抽出を行うが、この時ランクの表す現象に対して矛盾するデータは対象から除くものとし、ルールの精度確保に努めた。
- 3) 抽出されたルールとしては、「溪床勾配及び流域最大傾斜が緩くて流域幅が小さく、かつ安山岩質凝灰岩、輝石安山岩、第三紀層等が分布する北向きの溪流であり、かつ長期降雨が少ない地域」という特徴を有する溪流は危険性が低いと評価することができる。
- 4) 危険性が高いと評価される溪流は、「溪床勾配及び流域最大傾斜が急で流域幅の大きい溪流であり、かつ角閃石安山岩等が分布する地域の溪流」という特徴を有している。
- 5) 最終的に設定された17個のルールで、S57.7長崎災害の事例の全データ（サポート100.0%）に対して約8割の精度（確信度）で説明可能となった。このことから本研究によるルール抽出方法は、高い汎用性と精度を有するルールを抽出できることが確認できた。

以上のように、第4章では危険度ランクに対応した評価ルールの汎用性向上に取り組んだ結果、17個のルールで対象地域全ての箇所を説明できることが確認され、当該地域の典型的なルール抽出ができたと言える。これにより、将来的なハード対策の実施候補の絞込みとして、最も上位ランクの「危険」に該当する箇所を優先することで、効率的な防災事業の推進に寄与することが可能となったと考えられる。逆に「安全」に該当する箇所に対しては調査実施の優先度を下げることができ、防災事業全体の効率化にも寄与するものと考えられる。

そこで第5章では、より一層の効率化を目指すために、第4章で抽出された各ルールに該当する箇所の持つ危険度に注目し、ルール毎の危険度の差異について分析した。そ

の結果、「危険」、「比較的危険」に該当するルールの中でも特に危険なルールの特定が可能となった。これにより、そのルールに該当する 10 箇所が、対象地域の中で最優先箇所として詳細調査や対策検討を行うべき箇所として抽出できた。一方、安全側のルールに関しても同様に、特に危険性の低いルールを特定できた。予算や時間に制約のある公共事業においては、全ての箇所を同列で事業を進めていくことは不可能であることから、このようなルールに該当する箇所は、調査実施の優先度を下げるなどの対応を取る必要がある。また、本研究の結果を基に一連の分析手法を活用した防災対策箇所の特定フローを検討し、本研究成果の実用化に向けた提案を行った。

6.2 今後の課題と将来への展望

本論文では、SVM とラフ集合といった数理的な手法を用いることで、土石流に対する危険度ランクに対応した汎用的なルールの設定を試みた。その結果、対象箇所全てを約 8 割の精度で説明可能となる汎用的なルールの設定が可能となった。

第 2 章では、既往研究でその有用性が確認されていた SVM とラフ集合を組み合わせた手法の汎用性検証を行った。本論文では対象地域を変えて検証を行ったが、災害の形態としては、既往研究と同じ「土石流」を対象としている。そのため、手法の適用性をさらに拡大していくためには、対象地域を変えた異なる災害形態（急傾斜地の崩壊、落石など）への適用を検証する必要があると考えられる。

第 3 章、第 4 章では、災害の「発生」・「非発生」に関するルール抽出のみではなく、危険度に対応したルールとして、①危険、②比較的危険、③比較的危険、④安全の 4 つの危険度ランクに対応した評価ルールを設定した。この区分数に関しては、SVM で出力される $f(x)$ 値の持つ意味合い（サポートベクターとなる領域とその外側の領域）に応じて設定したが、更に細かく区分することも可能と考えられる。しかしながら、母集団のデータ数によっては、危険度ランクを細かくしすぎると、そのランクに属するデータ数が少なくなり、有効なルール抽出ができない可能性も考えられる。

これらの検討は、昭和 57 年の長崎豪雨における災害を対象としていることから、土石流の発生データが非常に多い事が特徴であった。しかしながら、この手法を他地域への適用を検証する場合には、災害の発生事例が少ないもしくは発生事例がない地域も考えられる。その場合は、地質的な条件等が似ている地域での分析結果を学習データとし、当該地域に適用する方法が考えられる。しかしながら、直接的な分析ではないため、現地確認等の検証等を積み重ねる事で、手法の汎用性の確認が必要であると考えられる。

第 5 章では、本研究での成果を基に一連の分析手法を活用した防災対策優先箇所の特定フローを検討し、本研究成果の実用化に向けた提案を行った。現在実施されている土砂災害防止法に基づく基礎調査結果には、人家戸数・公共施設・災害時要援護者関連施設・防災拠点等の防災観点上の重要度に関する情報も整理されていることから、この情報と本研究で提案する危険度評価システムを連携させることで、その判断材料とするこ

とも可能と考える。ただし、実際のハード対策事業を実施するに当たっては、危険度のみの指標ではなく保全対象などの重要度も考慮した上で、事業化の可能性を判断する必要がある。よって、今後は「危険度」と「重要度」の両視点からの事業優先度評価に関する検討が望ましいと考える。

本研究の成果は、防災事業に対して、特に危険な箇所の特定制および安全な箇所の特定制が可能となることから、大幅な効率化を図ることが可能と考えられる。今後は予防保全的な管理を実施する必要がある社会資本の維持管理等に対しても、適用の可能性について検討を進めていく所存である。

謝 辞

本論文は、筆者が平成 24 年 10 月に入学した山口大学大学院理工学研究科博士後期課程システム設計工学系専攻における研究成果を取りまとめたものです。本論文をまとめるにあたり、多くの方々より多大なる御指導、御協力を頂きましたことを深く感謝し、ここに謝意を表します。

本研究の機会と便宜を与えていただいた山口大学 古川浩平名誉教授には深甚なる謝意をここに表します。古川浩平名誉教授には、学部 4 年生の研究室配属時から博士前期課程を含む計 6 年間に渡り、指導教官として御指導いただきました。その間、土砂災害に関わる研究についてはもちろんの事、(有)山口ティー・エル・オーを通じた特許技術の開発や技術移転に関する知識の習得など、技術者としての視野を広げる環境を与えていただきました。この経験は、設計コンサルタントに勤める技術者として貴重な経験となり、今後の大きな糧になるものと感じています。また、本論文の執筆期間においては、幾度となく会社に足を運んでいただき、論文の構成や取りまとめ方など、熱心に御指導いただきました。退官された後の貴重なお時間を筆者の論文執筆に割いていただきましたことに深く感謝申し上げます。

学位審査におきましては、主査を務めていただきました山口大学理工学研究科 松田博教授には、種々のご配慮をいただき本当にありがとうございました。さらに、予備審査員を務めていただきました兵動正幸教授、羽田野袈裟義教授、清水則一教授、鈴木素之准教授からはより良い論文にするための的確な御意見・御指導を頂戴しました。研究という会社では経験できないような視点とわかりやすい論文への配慮など今後の実務にも活用可能な貴重な体験をさせていただきました。心より感謝の意を表します。

中電技術コンサルタント株式会社 荒木義則博士、倉本和正博士、杉原成満博士、八千代エンジニアリング株式会社 菊池英明博士、岡山理科大学 佐藤丈晴博士、株式会社エイト日本技術開発 海原荘一博士、竹本大昭博士、株式会社四電技術コンサルタント 松本幸太郎博士、大日本コンサルタント株式会社 篠崎嗣浩博士の皆様には、常に温かい御助言・御助力を頂きました。古川研究室の先輩である皆様のお力添え無くして、この研究をやり通すことはできなかったと思っています。ここに心より深く感謝の意を表します。

また、西日本技術開発株式会社 森二郎土木本部長、西園幸久部長、小林博昭部長、大石博之部長代理、澤田尚課長、円田竜太課長補佐には、筆者の大学入学から卒業に至るまでに御尽力頂いたのを始め、本研究を行う機会と共に御支援頂きました。中でも大石博之部長代理におかれましては、業務と論文執筆の間で悪戦苦闘している筆者に対して、発注者を含めた社内外へのフォローや論文全般に対するアドバイスなど、御配慮頂きました。ここに心より深く感謝の意を表します。

最後に、私事ではございますが、常に筆者を支えてくれた家族に深く感謝いたします。生活面、精神面で大きな支えとなってくれた妻麻美と明るい笑顔で癒しを与えてくれた娘滯央、息子悠真に心から深く感謝の意を表します。

平成 27 年 8 月

小林 央宜