

# 2016年8月の常呂川連続洪水における機械学習手法を用いた水位予測の研究

室蘭工業大学 ○学生員 岡崎亮太 (Ryota Okazaki)  
 室蘭工業大学 正員 中津川誠 (Makoto Nakatsugawa)  
 室蘭工業大学 非会員 小林洋介 (Yosuke Kobayashi)  
 室蘭工業大学 学生員 谷口陽子 (Yoko Taniguchi)

## 1. はじめに

2016年8月、北海道において4つの台風が連続して上陸及び接近し、広域にわたって甚大な被害をもたらした。道東の常呂川流域では、下流に位置する上川沿水位観測所において、計画高水位を超過し、観測史上最大の水位を記録した。このような被害を事前に予測し被害を軽減させるためには、水防活動、避難場所の開設、住民の避難等といった防災・減災対策に要する時間(リードタイム、以下、LTと称す)の確保が望まれ、LTが長く、ピーク水位について精度の良い河川水位予測が望まれる。

河川水位の計算手法としては、上流部の河川水位との相関をとる方法(以後、水位相関法)や、洪水流出モデルで流量を計算し、水位-流量曲線(H-Q曲線)から水位に換算する方法がある。本研究ではその考え方を拡張し、下流地点の水位に関係する説明変数の抽出に機械学習を用い、水位予測の精度向上を試みた。これを「関連要因相関法」と呼称する。

本研究では、説明変数の寄与度を算定できる機械学習法の一つであるRandom Forest法(以下、RF法)によって水位予測を試みる。さらに、現業での実用性を鑑み、寄与度の高い説明変数を用いて関連要因相関法による水位予測を行い、説明変数の選別、LTの延長といった面で検討を行った。以下に研究手順を示す。

- RF法を用いて、過去の洪水事例を学習させて作成した予測モデルより、2016年8月の洪水事例を予測した。ここでは、過去のデータのみで2016年8月の洪水事例を予測検証した。
- 1)において算出された寄与度から、上位の説明変数を用いて重回帰式を算定し、LT延長の検証を行った。

## 2. 研究方法 2.1 台風災害の概要

対象河川は北海道東部に位置する一級河川常呂川における下流の上川沿地点とした。図-1に対象河川の概略図を示す。この上川沿地点では、北見市街中心部付近の地点に位置しており、2016年8月の洪水時は、水位が下がりきらないうちに次々と台風が連続して接近したため、少量の雨量でも水位が上昇しやすい状態が続いた。図-2に今回の災害における上川沿地点における積算雨量-積算流出高の関係を示す。これより求めた各台風の期間における流出率は図-2に示すようになった。③の台風期間では、積算流出高が積算雨量を上回る形になっており、流出率は1.0を超えていることが分かる。これは、上川沿地点における流出量が降雨による表面流出に加え、土壌水分状態が関係しているためと考えられる。よっ

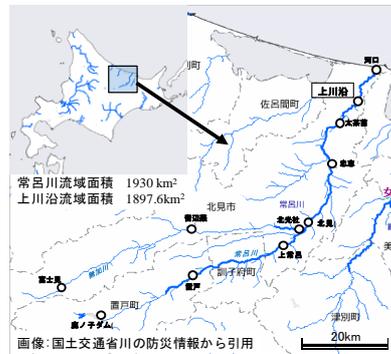


図-1 対象河川の概略図

表-1 関連要因の概要

観測項目	地点 流域	避難に必要な時間 (LT)
水位 (m)	上川沿 太茶苗	6-10時間前 の水位
地点観測水位 (テレメータ水位)	志志 北見 上常呂 置戸	
時間雨量 (mm)	上川沿 北見 置戸	
地点観測雨量 (テレメータ雨量)	常呂川流域	6-10時間前 の流域平均雨量
流域平均雨量 (レーダ解析雨量)		
SWI <sub>s</sub> <sup>*</sup> (mm)		
流域平均雨量(レー ダ解析雨量)を入力 した3段タンクモデ ルの1段目貯留高 <sub>1</sub>	常呂川流域	6-12,24,36,48,60, 72時間前のSWI <sub>s</sub>

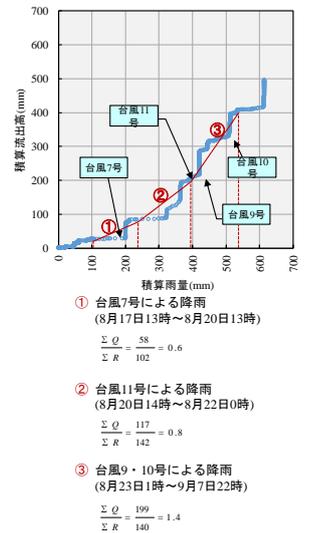


図-2 積算雨量-積算流出高の関係

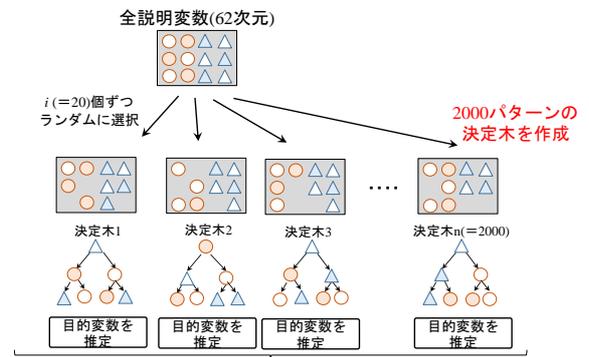


図-3 RF法の概略図

て本研究では、土砂災害発生リスクを判定する指標として用いられる土壌雨量指数(SWI)も水位を予測する上で重要な要素として考慮することにした。また、本研究ではこの2016年8月4日~9月14日までの台風接近期間を対象事例とし、上川沿地点でのピーク水位の予測精度向上およびLTの延長を目指した。こ

表-2 RF法で学習させた過去の洪水事例

事例期間	ピーク水位	総雨量
事例1 2001/09/06 1:00 - 09/20 24:00	8.48m	186.8mm
事例2 2006/10/03 1:00 - 10/17 24:00	8.41m	151.3mm
事例3 2011/08/30 1:00 - 09/13 24:00	7.22m	215.0mm

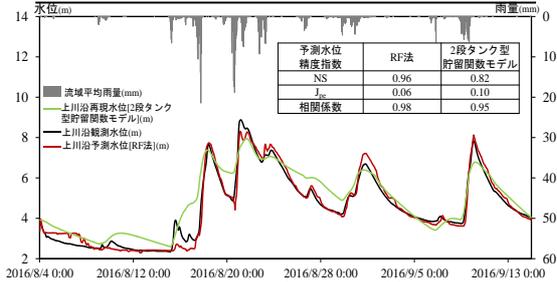


図-4 RF法の水位予測結果

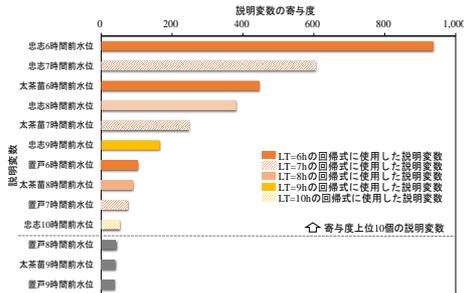


図-5 RF法から算出した寄与度

これらの予測を行う上で必要な要素として用いた62次元の関連要因の概要を表-1に示す。

2.2 RF法による水位予測

RF法とは、多数の決定木を組み合わせることで性能を高める機械学習アルゴリズムである。図-3にRF法の概略図を示す。決定木とは、全データの集合に対して、複数の変数を組み合わせた分割をデータから自動的に生成したものである。また、元データの分布を問わず外れ値に対して頑健な予測ができる特徴を持ち、非線形なデータ分析に多用される。RF法では、学習ごとにm本の決定木による予備実験によってi個の最適な値を決定し、1つの決定木の作成に全説明変数からi個ずつランダムに重複を許して選出を行うことで、n本の決定木を作成する。その後、各決定木の目的変数の推定値を平均して、最終的な目的変数の値を決定する。決定木は、各説明変数の寄与度を分割の良さを表すGini係数を用いて算出できる。したがって、RF法も各決定木の寄与度を統合して全説明変数の寄与度を分析できる。本研究では、m=200, n=2,000として、表-1の説明変数を使用することでRF法による上川沿の水位予測及び説明変数の寄与度の算定を行った。

2.3 関連要因相関法

前節のRF法において表-2に示す過去3事例の洪水を学習させることにより算定した説明変数の寄与度から、上位10個の説明変数を抽出して重回帰式を作成した。その後、過去3事例の洪水の平均回帰係数を用いることで、2016年8月洪水事例における上川沿での水位予測を行った。

3. 結果と考察

3.1 過去の洪水事例の学習による水位予測結果

図-4よりRF法の予測結果は、ピーク水位に関して若

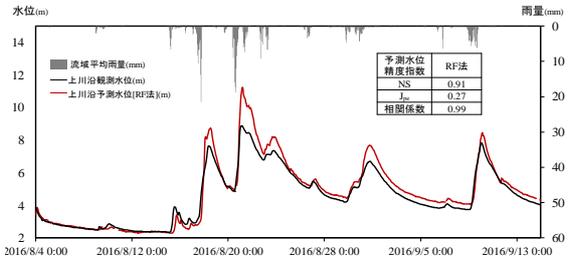


図-6 関連要因相関法での水位予測結果

干誤差があるものの、全体的に観測水位の挙動をよく予測している。同様に、常呂川流域において貯留関数モデルによる水位予測の結果を図-4に緑線で示すが、ピーク水位においてRF法が安全側に予測していることが分かる。これより、対象事例のような今までに経験したことのない洪水事例に対しても、RF法において過去のデータを学習することで予測が可能であることが示された。また、本研究で用いた精度指標は以下に示すNash-Sutcliffe係数(以下、NS係数)、ピーク水位相対誤差(以下、J<sub>pc</sub>)の2つを用いて行った。

$$NS = 1 - \frac{\sum_{i=1}^N (H_{oi} - H_{ci})^2}{\sum_{i=1}^N (H_{oi} - H_{oave})^2} \quad (1), \quad J_{pc} = \frac{|H_{op} - H_{cp}|}{H_{op}} \quad (2)$$

ここで、H<sub>oi</sub>は降雨開始i時間後の観測水位(m)、H<sub>ci</sub>は降雨開始i時間後の計算水位(m)、H<sub>oave</sub>は観測水位平均値(m)、H<sub>op</sub>は観測ピーク水位(m)、H<sub>cp</sub>は計算ピーク水位(m)、Nは総降雨時間数である。NS係数は、1.0に近づくほどモデルの精度が高く、0.7以上でモデルの精度が高いとされる。

3.2 RF法の寄与度算出結果

図-5に3.1節でのRF法で得られた寄与度の算出結果を示す。図には説明変数の寄与度の大きい順に、全説明変数の上位20%を示している。図-5から上位に各観測所の水位が上位を占めた。

3.3 関連要因相関法による水位予測結果

現業における実用性を鑑み、図-5の寄与度上位10個から説明変数を抽出し、関連要因相関法によりLTの延長を試みた。今回はRF法のLTが6時間であるので、LTを6時間で抽出した。予測結果を図-6に示す。RF法よりも説明変数を少なくした関連要因相関法においても、ピーク水位を安全側に予測する結果を得た。

4. 結論

本研究において得られた結果を以下にまとめる。

- 1) RF法による水位予測結果から、過去の情報により2016年8月の未曾有の連続洪水の予測が精度良く行えた。
- 2) 関連要因相関法では3つの説明変数により作成した重回帰式での水位予測結果から、この手法が現場で使用できる可能性が示唆された。

参考文献

- 1) 速水悟：事例+演習で学ぶ機械学習 ビジネスを支えるデータ活用のしくみ、第5章、森北出版株式会社、2016。
- 2) (財)北海道河川防災研究センター・研究所：実践流出解析ゼミ 講義テキスト編、第12回、2006。