

石狩川を対象とした RF 法による水位予測の研究

室蘭工業大学 ○学生員 坂本 莉子 (Riko Sakamoto)
 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)

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。このような水害を事前に予測し被害を軽減させるためには、防災・減災対策に要する予測先行時間（リードタイム、以下、LT と称す）の確保が望まれる。特に市街地に近い河川では洪水時における早めの適切な判断が住民の安全に影響するため、LT が長い水位予測が望まれる。

現在、水位予測で主に用いられているのは予測雨量を入力条件に、貯留関数法（たとえば、星ら¹⁾）のような流出計算モデルで流量を計算し、水位に換算する予測手法が主流である。しかし、多くの不確実な誤差要因があり、必ずしも満足な結果が得られない場合がある。そのため近年では、機械学習手法により水位そのものを使用し予測する方法が注目されている。岡崎ら²⁾は各説明変数の寄与度が算定できる機械学習手法の 1 つである Random Forest 法（以下、RF 法）による水位予測手法を提案した。また、RF 法により各説明変数の寄与度を算出し、それらから回帰式を作成し水位予測を行う関連要因相関法を提案している。本研究では石狩川の水位予測に岡崎ら²⁾の考え方を適用し、LT の延長を考慮といった面で検討を行った。

2. 研究方法

2.1 対象地点と対象事例

対象河川は北海道西部に位置する一級河川石狩川とした。その中で最も重要な基準点であり、わが国の洪水対策の要の一つといえる石狩大橋地点（河口より 26.60km）を対象に、水位の予測を行った。図-1 に予測対象地点の概略図を示す。対象事例は 2016 年

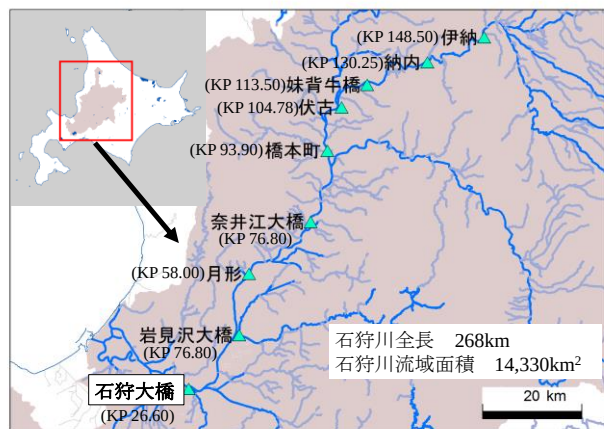


図-1 対象河川の概略図

8 月事例の北海道豪雨を対象とした。水位予測を行う上で必要な要素として国土交通省所管のテレメータ観測所の水位データを使用した。

2.2 RF 法による統計モデル

RF 法とは、多数の決定木を組み合わせることで汎化能力を高める方法である。図-2 に RF 法の概略図を示し、以下の方法で計算を行う。

- (1) 全説明変数 n 個から重複を許しランダムに 1 つの説明変数を抽出する作業を m 回行う。 n 及び m の値が十分大きい場合 n の約 64% ($=1-(1-1/n)^m$) の説明変数が抽出される。
- (2) 抽出された説明変数を用いて 1 本の決定木を作成する。決定木の構造は図-3 に示す平均二乗誤差（以下、MSE）を用いて作成する。これを全てのパターンで計算して、MSE の値が最も小さくなる x で説明変数を分割するような決定木を作成し、目的変数を推定する。
- (3) (1), (2)の作業を L 回行い L 本の決定木を作成する。各決定木で推定した目的変数を平均し、最終的な目的変数を決定する。
- (4) 本研究では、全説明変数 n は表-1 に示す通りであり、抽出回数 m は n と同数($m=n$)とした。決定木の本数 L は双方共に $L=2,000$ とした。

また、図-3 に示すような変数重要度（以下、 $\text{imp}[f_i]$ ）を決定木の分割ごとに計算を行い、全ての $\text{imp}[f_i]$ の総和で各説明変数の $\text{imp}[f_i]$ を除して正規化したものが寄与度となる。よって寄与度の値は 0~1 の範囲となる。この寄与度は、個々の説明変数と目

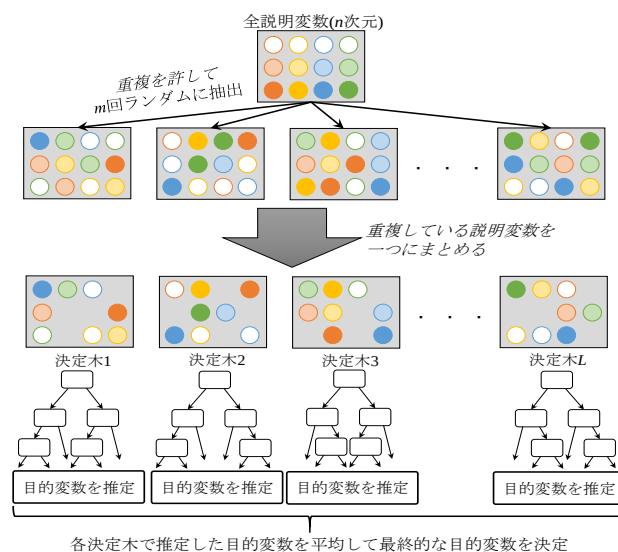


図-2 RF 法の概略図

キーワード 石狩川 水位予測 リードタイム Random Forest 法 関連要因相関法 寄与度
 連絡先 〒050-8585 北海道室蘭市水元町 27-1 国立大学法人室蘭工業大学 TEL 0143-46-5276

的変数の相関係数を求めた値と異なり、変数間の従属性の影響（多重共線性）を排除したよりロバストな説明力の指標となる。

2.3 関連要因相関法による予測

表-2 に示す過去 4 事例の洪水を学習させた統計モデルの全説明変数の寄与度から、上位の説明変数を選出して回帰式を作成した。作成する回帰式は過去 4 事例の各回帰係数の平均値を用いることで、石狩大橋地点での水位予測を行った。

3. 結果と考察

3.1 RF 法による統計モデルでの水位予測結果

RF 法において、表-2 に示す過去の洪水 4 事例を学習させ、水位予測を行った結果を図-4 に示す。図-4 に示す 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ているため LT=6h の時に 0.93、LT=12h の時に 0.86 と高く、RF 法は高精度な水位予測ができることが分かる。

3.2 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考え、図-5 の寄与度上位 10 個から LT を考慮し赤く囲った 2 個の説明変数を選出し、関連要因相関法による水位予測を行う。図-6 に関連要因相関法による予測結果を示す。この結果、LT=6h ではピーク水位や水位の立ち上がりを精度良く予測できた。また、LT=12h は多少過大にはなってい

るがピーク水位を安全側に予測できる結果を得た。

4. まとめ

本研究において得られた結果を以下にまとめる。

- (1) 長大な河川では上流部の水位データのみで下流水位の予測が可能であることが示された。
- (2) 関連要因相関法による水位予測結果より、少ない説明変数で精度の良い水位予測を行うことが可能であり、現業に実装しうる予測方法が提案できた。

謝辞

本研究を進めるにあたり、北海道開発局札幌開発建設部には洪水に関する資料を提供して頂いた。また、室蘭工業大学情報電子工学系学科の山田恒揮氏には解析プログラムの提供等ご協力をして頂いた。ここに記して謝意を表す。

参考文献

- 1) 星ら：小流域における総合貯留関数法の開発，第 31 回水理講演会論文集，pp.107-112，1987。
- 2) 岡崎ら：ランダムフォレスト法による洪水時の水位予測手法の提案，土木学会論文集 B1(水工学)Vol.74，No.4，I_1459—I_1464，2018。

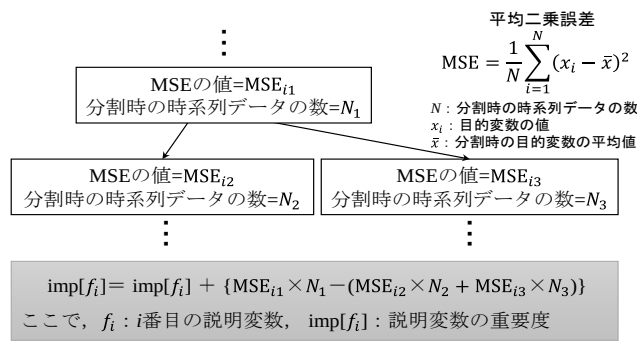


図-3 変数重要度算出方法の概略図

表-1 予測のための説明変数の諸元

観測項目	観測地点	データ概要	避難に必要な時間 (LT)
水位 (m)	石狩大橋	地点観測水位 (テレメータ水位)	予測リードタイム=6h 6~12時間前の水位 全説明変数n=63次元
	岩見沢大橋		
	月形		
	奈井江大橋		
	橋本町		
	伏古		
	妹背牛橋		
	納内		
	伊納		
			予測リードタイム=12h 12~24時間前の水位 全説明変数n=117次元

表-2 RF 法で学習させた過去の洪水事例

	事例期間	ピーク値
事例1	2001/09/06 01:00 ~ 2001/09/21 24:00	6.28m
事例2	2005/08/16 01:00 ~ 2005/08/31 24:00	4.03m
事例3	2011/08/30 01:00 ~ 2005/09/14 24:00	4.97m
事例4	1981/08/03 01:00 ~ 1981/08/12 24:00	9.23m

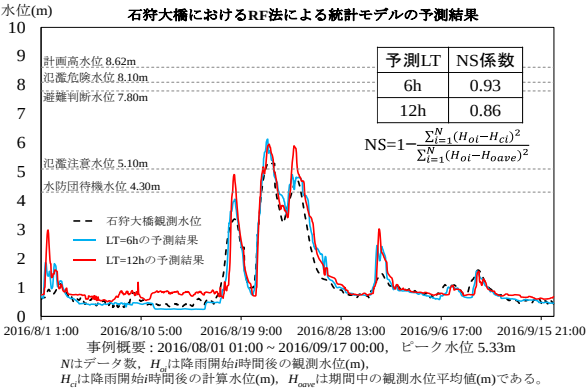


図-4 RF 法による統計モデルの水位予測結果

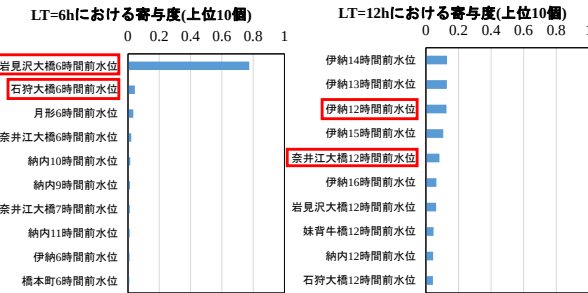


図-5 寄与度算出結果

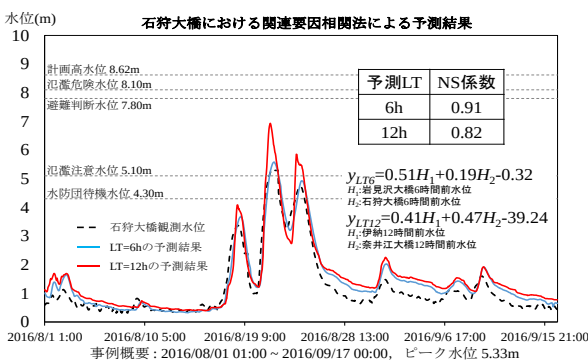


図-6 関連要因相関法による水位予測結果