

ランダムフォレスト法によるバーチャルハイドログラフ再現性に関する一考察

首都大学東京大学院 都市環境科学研究科 学生会員 ○竹内 泰裕
 東京都 土木技術支援・人材育成センター 正会員 高崎 忠勝
 首都大学東京大学院 都市環境科学研究科 正会員 河村 明
 首都大学東京大学院 都市環境科学研究科 正会員 天口 英雄

1. はじめに

都市域を流れる中小河川流域では、局地的な集中豪雨により河川の氾濫や内水による浸水被害が頻発している。特に東京においては、人口や資産の集積、都市化や気候変動等に起因した洪水流量の増大等の社会・環境条件により水害リスクが一層増大している。近年、土木分野においても機械学習やニューラルネットワーク等の AI 技術を用いた検討が数多く行われており、洪水予測では遺伝的アルゴリズム等が降雨流出モデルパラメータの同定に用いられている¹⁾他に、深層学習の適用も検討されている²⁾。降雨流出モデルとしての AI の適用に際しては、出力と関連性のある項目を入力に用いるが、深層学習では経験的に入力項目の決定を行っている。機械学習手法のひとつであるランダムフォレスト法（以下、RF 法）では、説明変数となる各項目の寄与度を算定できる特長を持つことから、入力項目を自動的に決定することが期待できる。本研究では、東京都の都市中小河川の洪水を想定したバーチャル洪水(以下、VF)に対して RF 法による降雨流出モデル(以下、RF 降雨流出モデル)を用いて流量を計算し、バーチャルハイドログラフ再現性を検証した。さらに、RF 法から得られる寄与度の高い説明変数のみを用いた計算を行い、説明変数を選別する効果を検証した。

2. 検証方法

まず、東京都の都市中小河川の洪水を想定した雨量・流量のデータセットである VF を作成した。VF の作成方法の詳細は文献^{3), 4)}を参照されたい。作成した 100 個の VF はピーク流量を基に並び替え、80 個の学習用洪水(VF(1)~VF(80))と 20 個の検証用洪水(VF(81)~VF(100))に分割した。

次に、学習用洪水を用いて RF 降雨流出モデルを学習させ、検証用洪水の流量を計算した。RF 降雨流出モデルは、入力を現在から n 分前までの 1 分雨量、出力を現在の流量としている。本研究の手法である RF 法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機械学習手法であり、寄与度は RF 法により求められる説明変数ごとの予測流量に与える影響を示す指標である⁵⁾。バーチャルハイドログラフ再現性の評価には誤差評価関数 RMSE(二乗平均平方根誤差)を使用した。

3. バーチャルハイドログラフ再現性の検証

表-1 に、本研究で検討したケースの各条件を示す。学習用洪水 80 個すべてを用いた学習を基本ケースとした。なお、入力項目となる説明変数の数を現在から 179 分前までの 180 個の 1 分雨量としている。図-1 に、検証用洪水の中で最もピーク流量が大きい VF(81)、最もピーク流量が小さ

表-1 検討ケース

検討ケース	入力		出力		学習用洪水		検証用洪水	
	内容	個数	内容	個数	内容	個数	内容	個数
基本ケース	R(t)~R(t-179)	180	Q(t)	1	VF(1)~VF(80)	80	VF(81)~VF(100)	20
ケースA-1	R(t)~R(t-179)	180	Q(t)	1	VF(1)~VF(40)	40	VF(81)~VF(100)	20
ケースA-2	R(t)~R(t-179)	180	Q(t)	1	VF(1)~VF(20)	20	VF(81)~VF(100)	20
ケースA-3	R(t)~R(t-179)	180	Q(t)	1	VF(1)~VF(10)	10	VF(81)~VF(100)	20
ケースB-1	R(t)~R(t-119)	120	Q(t)	1	VF(1)~VF(80)	80	VF(81)~VF(100)	20
ケースB-2	R(t)~R(t-89)	90	Q(t)	1	VF(1)~VF(80)	80	VF(81)~VF(100)	20
ケースB-3	R(t)~R(t-59)	60	Q(t)	1	VF(1)~VF(80)	80	VF(81)~VF(100)	20
ケースB-4	R(t)~R(t-29)	30	Q(t)	1	VF(1)~VF(80)	80	VF(81)~VF(100)	20
ケースB-5	R(t)~R(t-19)	20	Q(t)	1	VF(1)~VF(80)	80	VF(81)~VF(100)	20

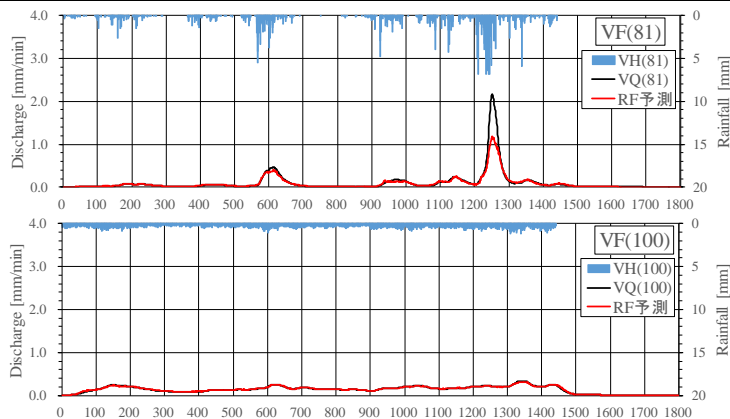


図-1 基本ケースにおけるバーチャルハイドログラフ再現性

いVF(100)のバーチャルハイトグラフとバーチャルハイドログラフを重ねて表示した。VF(81)において、予測流量はピークのタイミングは一致してものの、ピーク流量をやや過小評価していることが確認できる。VF(100)において、予測流量はバーチャルハイドログラフを精度良く再現しており、高い再現性を有していることが確認できる。

学習用洪水の数を変化させた場合の検討(A-1～A-3)を行い、バーチャルハイドログラフ再現性を検証した。図-2に、各検証用洪水のRMSEを示す。学習用洪水の数が多くなるほど高い再現性を有しており、また、バーチャル洪水を偏りがなく学習している検討ケース(A-1, A-2)のバーチャルハイドログラフ再現性は基本ケースとの差異がほとんどみられないことを確認した。

図-3に、基本ケースにおいて説明変数ごとの寄与度を示す。現在時刻を t 、現在の1分雨量を $VR(t)$ 、 n 分前の1分雨量を $VR(t-n)$ とする。図-3より、 $VR(t)$ から $VR(t-90)$ 程度までの寄与度が高いことが確認できる。

そして、基本ケースに対し、説明変数の数を変化させた場合の検討(B-1～B-5)を行い、バーチャル

ハイドログラフ再現性を検証した。図-4に、各検証用洪水のRMSEを示す。ケースB-2において、バーチャルハイドログラフ再現性が最も高くなり、寄与度の高い説明変数のみを用いることで、バーチャルハイドログラフ再現性が高まることを確認した。さらに説明変数の数を減らした場合の検討ケース(B-4, B-5)では、入力項目に寄与度の高い説明変数が一部含まれていないため、バーチャルハイドログラフ再現性が低くなることを確認した。

4. むすび

本研究では、東京都の都市中小河川を想定した洪水に対してRF法による降雨流出計算を行い、バーチャルハイドログラフ再現性を検証した。学習に用いる洪水の数を変えて計算した結果、学習用洪水の数が最も多いケースにおいてバーチャルハイドログラフ再現性が最も高くなることを確認した。そして、寄与度が大きい説明変数のみを用いて計算を行うことによりバーチャルハイドログラフ再現性がさらに高くなることを確認した。

参考文献

- 1) 高崎忠勝, 河村明, 天口英雄, 荒木千博: 都市の流出機構を考慮した新たな貯留関数モデルの提案, 土木学会論文集 B, Vol.65, No.3, pp.217-230, 2009.
- 2) 一言正之, 桜庭雅明, 清雄一: 深層学習を用いた河川水位予測手法の開発, 土木学会論文集 B1 (水工学), 72 巻, 4 号, pp.I_187-I_192, 2016.
- 3) 村井雅姿, 高崎忠勝, 河村明, 天口英雄: 東京都の中小河川を対象としたバーチャルハイトグラフの作成, 第46回土木学会関東支部技術研究発表会公演概要集, 2019.
- 4) 藤塚慎太郎, 河村明, 天口英雄, 高崎忠勝: AIを用いた都市流出モデルエミュレーションのためのベンチマーク模擬ハイドログラフの構築, 第46回土木学会関東支部技術研究発表会公演概要集, 2019.
- 5) 岡崎亮太, 中津川真, 小林洋介: 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), 74 巻, 4 号, pp.I_1459-I_1464, 2018.

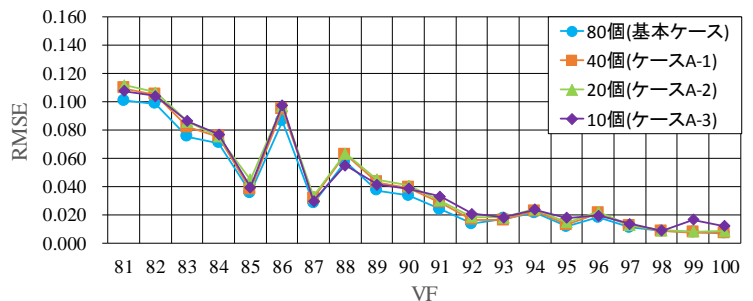


図-2 学習用洪水の数を変化させた場合のRMSEの値

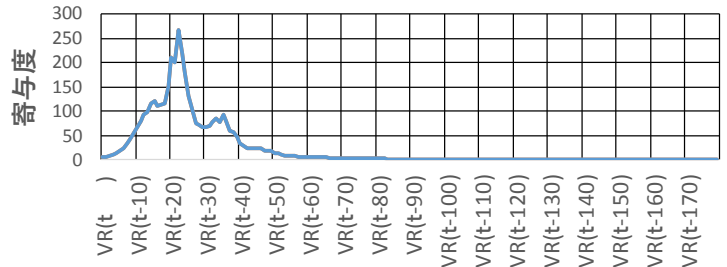


図-3 基本ケースにおける説明変数ごとの寄与度

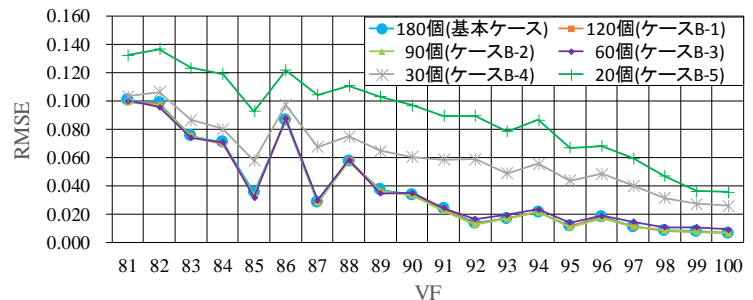


図-4 説明変数の数を変化させた場合のRMSEの値