

機械学習による永瀬ダムへの流入量予測モデルの構築および有用性の検証

高知高専専攻科 学生会員 池田圭吾

1. はじめに

永瀬ダムでは、流入量が 500 m³/s 以上となる洪水時、洪水調節を実施して下流域の洪水被害を軽減している。そのためには、洪水前に予備放流を実施し、洪水調節容量を確保する必要がある。予備放流はダムへの流入量予測に基づいて計画され、実施前には関係機関および地域住民に周知する必要がある。そのため、予備放流が遅れると下流域での浸水リスクの増大やダムの貯水容量不足による緊急放流の実施可能性が高まる。

永瀬ダムの流入量予測に採用されている貯留関数法モデルは、流域を一様と仮定した集中型モデルであり、精度には限界がある。一方、近年注目を集めている機械学習モデルは、ビッグデータから人間には理解が困難な複雑な特徴まで抽出でき、モデルにデータを与えると自動で予測が行える利点がある。そのため、近年洪水リスクが高まる中、人員の少ない県管理ダムにおいては、機械学習モデルの適用可能性を検討すべきと考える。

そこで本研究では、洪水時における効果的な永瀬ダム運用支援ツールの構築を目的として、機械学習によるダムへの流入量予測モデルを構築し、その有用性を検討した。またモデルに用いる降雨量データに、ダム流域の地点雨量を用いた場合と 1km メッシュ解析雨量を用いた場合で比較し、予測精度への影響を考察した。

2. 研究方法

機械学習モデル構築には、永瀬ダム管理事務所より提供いただいたダムへの流入量および流域 5 地点観測所(別府、久保、笹、舞川、永瀬)の降雨量データならびに気象庁の 1km メッシュ解析雨量データのうち、2006 年から 2019 年の 14 年間の 1 時間毎、計 122,712 時間分(欠損時間を含む)のデータを用いた。

流入量および地点雨量データは欠損値を含んでおり、紙媒体のダム管理日報から手作業での補填を行った後、1 か所でも欠損値が残る時間は削除した。なお 9,904 時間あった欠損時間のうち 6,262 時間を補填した。

解析雨量データは、図-1 に示す永瀬ダム流域を囲む範囲で抽出した。またグリニッジ標準時になっており、-9 時間で日本標準時に変換した。また Max プーリング処理を行い、486 面(27×18)から 81 面(9×9)に圧縮した。

データのスケールに差異があると学習効率下がるため、Min-Max normalization による正規化を行った。

教師あり学習を行うため直近 24 時間の流入量および流域降雨量と 1 時間先の流入量のデータセットとした。また単純なホールドアウト法により、前半の 8 割を訓練データ、後半の 2 割をテストデータに分割した。なおテストデータ内には、既往最大規模の 2018 年 7 月洪水が含まれる。

既往洪水を参考に、6.64 m³/s を閾値として、閾値未満の流入量を含むデータセットは全て削除した。図-2 に示す構築モデルの概要は、直近 24 時間のダムへの流入量および流域降雨量データ(説明変数)を入力することで、1 時間先のダムへの流入量(目的変数)を出力する予測モデルである。なお流域降雨量データに地点雨量 5 地点を用いたモデル(Case1)と解析雨量 81 面を用いたモデル(Case2)を構築した。

開発言語には Python3.7(Anaconda3)を、機械学習モデル構築には Keras2.4(TensorFlow2.0)を用いた。モデルには一般的な DNN(Deep Neural Network)モデルを用いた。既往の研究¹⁾では、他モデルに比べ高精度とされている。

表-1 に Case1 および Case2 で統一した学習条件を示す。また表-2 にモデルごとに検討した学習条件を示す。



図-1 永瀬ダム流域図および解析雨量の抽出範囲

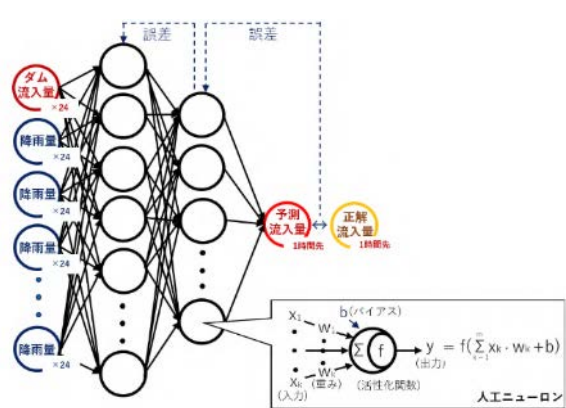


図-2 構築モデルの概要

表-1 両モデルで統一した学習条件

構築モデル	モデル	活性化関数 (中間層)	活性化関数 (出力層)	損失関数	最適化手法
Case1	DNN	ReLU関数	恒等関数	MSE (二乗平均誤差)	Adam
Case2					

表-2 モデルごとに検討した学習条件

構築モデル	中間層数	人工ニューロン数		バッチサイズ	エポック数
		1層目	2層目		
Case1	2層	256	128	512	278
Case2	2層	128	64	128	197

3. 研究結果

図-3 の Case1 モデルのテストデータによる 1 時間先の予測値と正解値の関係では、RMSE(二乗平均平方根誤差)は $13.80 \text{ m}^3/\text{s}$ となり、全体的に大きめに予測している。また $1500 \text{ m}^3/\text{s}$ 以上の洪水に対しては、精度のばらつきが大きい。図-4 の 2018 年 7 月洪水に対しては、ピーク付近でずれはあるが、規模、形状は概ね予測できている。欠損値の補填を行わない場合、RMSE は $16.58 \text{ m}^3/\text{s}$ となり、精度の低下を確認した。欠損時間による不自然な時系列が影響したと考える。また 2000 年から 2019 年までの 20 年間のデータを用いた場合、RMSE は $10.67 \text{ m}^3/\text{s}$ となり、精度の向上を確認した。2004 年の台風 16, 23 号および 2005 年の台風 14 号時の大規模洪水が、訓練データに含まれたことが影響したと考える。

図-5 の Case2 モデルのテストデータによる 1 時間先の予測値と正解値の関係では、RMSE は $17.22 \text{ m}^3/\text{s}$ となり、Case1 と比較して全体的に小さく予測している。 $1500 \text{ m}^3/\text{s}$ 以上の洪水に対しては、特に小さく予測しており、図-6 でもピーク付近を予測できていないことが分かる。洪水時には安全性の観点から、小さく予測することは避けるべきであると考え。また Max プーリング処理を行わない場合、RMSE は $23.25 \text{ m}^3/\text{s}$ となり、精度が著しく低下することを確認した。相関の高い説明変数を大量に用いたことが影響したと考える。

モデルの構築に用いていない 2020 年 7 月洪水を対象に Case1 モデルの予測精度を検証した。予測の流れは、①Case1 モデルに直近 24 時間の永瀬ダムへの流入量と 5 地点雨量データを入力し、1 時間先の流入量を予測する。②予測流入量と雨雲レーダーから読み取る予測雨量を用いて 24 時間を更新して、2 時間先の流入量を予測する。③②を繰り返して、16 時間先の流入量まで予測する。というものである。なお雨雲レーダーは、現時刻から 15 時間先までの予測結果が表示される。

図-7 は 2020 年 7 月洪水のピークの 2 時間前からの予測結果である。Case1 モデルによる予測が貯留関数法モデルによる予測を上回り、実績値に近い予測となった。ピーク規模は若干小さいが、形状を的確に予測しており、減水の予測も最後の 1 時間を除いて良好である。図-8 は同洪水ピークの 4 時間前からの予測結果である。図-7 と比較して予測精度が低下している。これは、予測時間が長くなることで、予測雨量の精度が落ちたことが要因であると考え。Case1 モデルによる予測に関しては、ピークが 1 時間後ろにずれ、規模も小さく予測している。しかし、貯留関数法モデルに比べて精度の低下割合は小さく、洪水の形状を概ね予測することができている。2020 年 7 月洪水は短時間でピーク ($1,229 \text{ m}^3/\text{s}$) に達する洪水であったため、貯留関数法モデルではピークの予測が困難であったと考えられるが、機械学習モデルでは、訓練データ内の洪水で規模、形状を予測できたと考える。また予測雨量の代わりに実績雨量を用いた場合、予測精度が改善されることを確認した。

4. おわりに

本研究の結論として、永瀬ダム流域においては流入量が $1500 \text{ m}^3/\text{s}$ 以下の中小規模の洪水に関して機械学習モデルによる 1 時間先の流入量予測が有用であると確認できた。また機械学習モデルの構築に用いる流域降雨量データは、解析雨量よりも地点雨量の方が効果的であると確認した。また欠損補填やプーリング処理など適切な前処理により機械学習モデルの精度が向上することも確認できた。なお今年度は大規模洪水が少なく、精度検証に用いた洪水は、2020 年 7 月洪水のみであったが、予測精度にばらつきが生じているため、複数の洪水で検証すべきであると考え。

今後の展望として、本研究では DNN モデルを採用し、1 時間先の流入量予測を行ったが、CNN(Convolutional Neural Network)モデルや LSTM(Long short-term memory) モデルを用いた場合および長時間先までの予測を行うモデルを用いた場合の予測精度も確認すべきと考える。また前処理の工夫で精度の改善が期待できると考える。

5. 参考文献

- 1) 斎藤治秀, 出井章裕, 安田晃昭, 今井裕也, 岡田将治: 機械学習による流域降雨量を用いた栃木県塩原ダムへの流入量予測モデルの構築と今後の活用に関する検討, 河川技術論文集, 第 26 巻, pp.277-282, 2020.

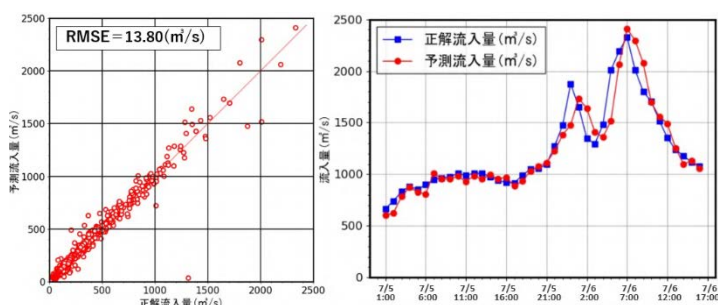


図-3 予測と正解の関係(Case1)

図-4 2018 年 7 月洪水(Case1)

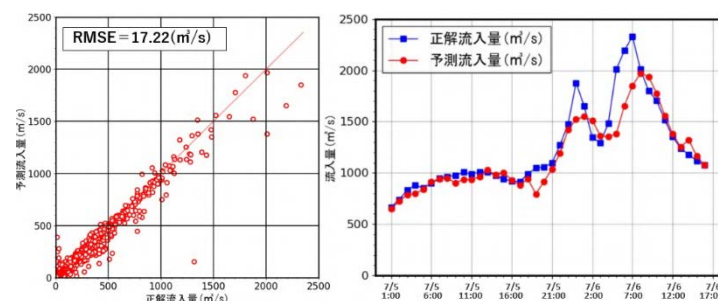


図-5 予測と正解の関係(Case2)

図-6 2018 年 7 月洪水(Case2)

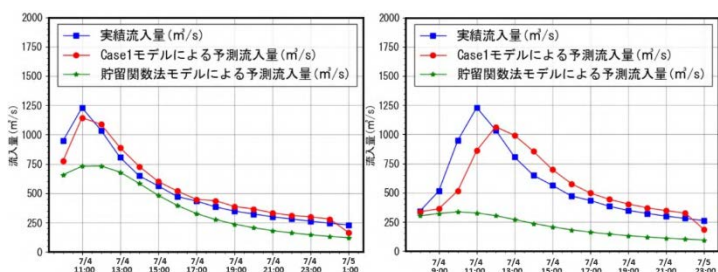


図-7 2020 年 7 月洪水(2h 前)

図-8 2020 年 7 月洪水(4h 前)

