

深層学習モデルを活用した 融雪期のダム流入量予測 —複数流域への展開に向けて—

国立研究開発法人土木研究所 寒地土木研究所 水環境保全チーム ○山田 嵩
横山 洋
柿沼 孝治

現在、多くの分野で深層学習の普及が進んでおり、水文分野においても水位予測やダム流入量等で深層学習の活用が進められている。先行研究では山田ら(2022)が深層学習モデルを定山溪ダムの融雪期におけるダム流入量予測に活用している。本研究では山田ら(2022)にて、定山溪ダムのみであった深層学習モデルの他流域への適用を行い、汎用性向上を目指す。

キーワード：ダム管理、人工衛星画像、融雪進捗状況

1. はじめに

積雪寒冷地において積雪は非常に重要な水資源であり、春から夏にかけて長期間安定した水を供給している¹⁾。一方、積雪は融雪期の豪雨による融雪出水や土砂災害の要因となる場合もある、このため、山間部における積雪の量及びその分布を正確な把握は、水資源管理及び防災の両面で重要である。

近年の計算機性能の向上に伴い、人工知能(AI)の活用が進められており翻訳や画像認識の分野では実用化されている。水文分野においても、深層学習を河川水位やダム流入量予測に活用する研究^{2) 3) 4)}がなされている。

一方で、融雪期のダム流入量予測を実施した研究は事例が少なく、日単位での予測であるが滝口ら⁵⁾、込山ら⁶⁾の事例がある。また、山田ら⁷⁾ではダム管理への応用を目指して72時間先までの予測を行っているが、対象領域を定山溪ダムのみとしており定山溪ダム以外への深層学習モデルの適用性は不明である。

そこで本研究では、定山溪ダム以外のダムでの適用性について確認するために、山田ら⁷⁾のAIモデルを他のダムにて適用して、他流域での精度検証を行い、深層学習モデル汎用性の向上を試みた。さらに、山田ら⁷⁾にて課題としていた、中間層を増やした場合及び入力データ時間を増加させた場合の精度の検討を行った。

2. 対象領域

対象領域は北海道の中央部に位置する流域面積238.9 km²の忠別ダム流域である。対象領域周辺の標高分布を

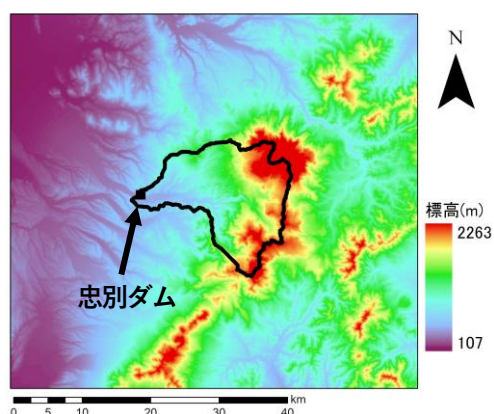


図-1 対象領域周辺の標高分布

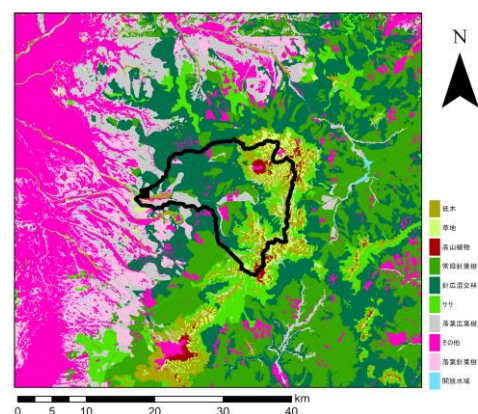


図-2 対象領域周辺の植生分布

図-1に、植生分布を図-2にそれぞれ示す。忠別ダム流域は標高1450 m付近が森林限界であり、植生分布が大きく変化する。忠別ダム域では森林域が大半を占める定山溪

ダム流域とは異なり、森林限界以上の高標高帯域が多く条件が異なる事から今回は忠別ダムを対象とした。

3. 使用データ及び解析手法

(1) 使用データ

本研究での使用データは忠別ダムにて観測された水文・気象観測データ（降水量、気温、積雪深及びダム流入量）である。

データの使用期間は2010年から2020年の各年における4月1日から6月30日でありデータの時間解像度は1時間であり、使用期間における欠測データは線形補間を行っている。入力データセットは降水量、気温、積雪深及びダム流入量として作成している。これらの条件は基本的に山田ら⁷⁾の研究と合わせたものである。

(2) データの前処理

忠別ダムにおけるダム流入量は水面変動によると思われる流入量の変動（ノイズ）が見られた。この影響を除外するために、24時間移動平均した値を本研究ではダム流入量としている。ただし、単純に移動平均を取るとダム管理上重要となるピーク値も大きく減少してしまう。それを防ぐため忠別ダムの洪水調節開始量である、移動平均前の値が100 m³/s以上の場合は、元の値を用いる事とした。2018年を対象にこの処理を行った結果を図-3に示す。

データセットは前述の線形補間及び移動平均に加えて、データの範囲をそろえるために、標準化処理を行っている。これらの処理は前処理と呼ばれ、深層学習をはじめとする機械学習では、一般的に必要とされるものである。標準化は平均を 0、分散を 1 にする処理であり式(1)で計

算される。

$$Y_s = \frac{X_s - \mu}{\sigma} \quad (1)$$

ここで Y_s は標準化後の値、 X_s は元データの値、 μ は平均、 σ は標準偏差である。標準化は同様に学習期間における各項目の平均値及び標準偏差により行っている。また、予測結果は標準化の逆処理を行い値を元に戻している。本研究にて標準化処理を採用したのは、山田ら⁷⁾では標準化処理を行った場合に最も高精度な予測となったためである。

(3) 深層学習モデル

深層学習モデルは山田ら⁷⁾の研究と基本的には同一であり、LSTM(Long Short-Term Memory)⁸⁾を用いている。モデル構造を図-4に示す。深層学習モデルの学習年は2010年から2018年までとし、2019年及び2020年を検証年としている。本来であれば、学習年以外の年を利用してハイパーパラメータの同定を行うが、本研究では汎用性検討のため中間層以外は山田ら⁷⁾の結果をそのまま利用した。中間層については山田ら⁷⁾の2層及び3層から、3層及び4

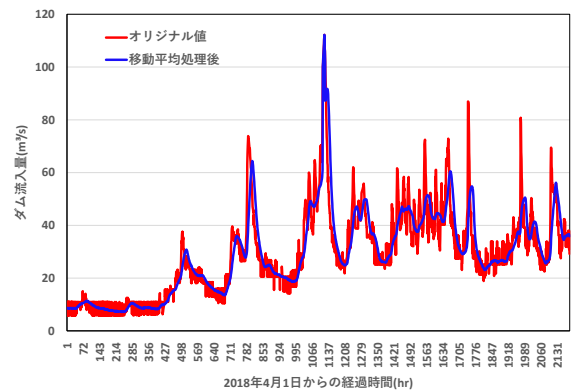


図-3 2018年のダム流入量

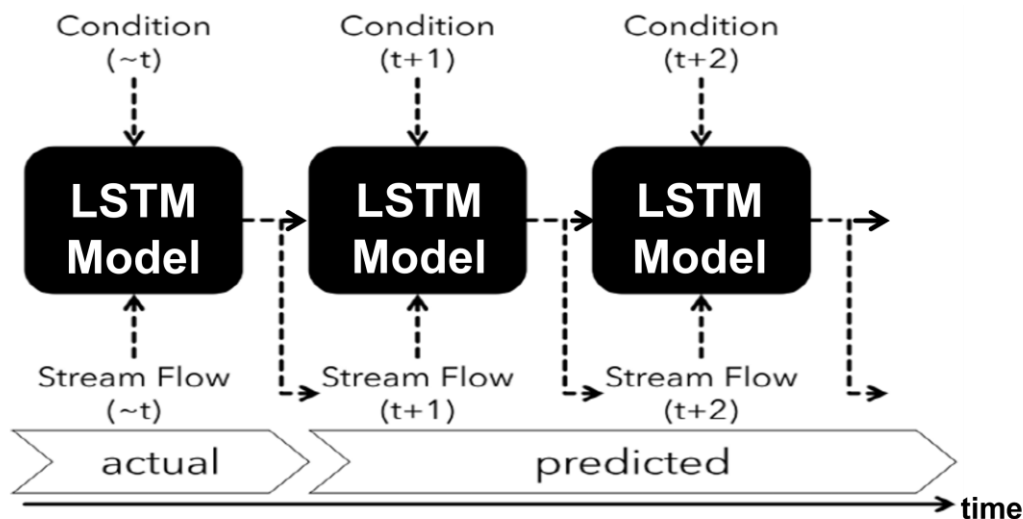


図-4 本研究におけるモデル構造

層に変更した。これは中間層を増やした場合の予測精度への影響評価を行うためである。ハイパーパラメータの一覧を表-1に示す。

予測に用いる入力データであるが、山田ら⁷⁾の研究では前24時間としているが、本研究では前24時間に加えて前36時間及び48時間とした場合についても検討を行った。尚、1時間先予測を行いその結果を24時間先予測に反映させる事、ダム流入量以外のデータについては完全予測値としている事は山田ら⁷⁾と同じである。

(4) 検証手法

ダム流入量の再現性検証にはNash and Sutcliffe⁹⁾の効率係数（以後、NS値と言う）を用いた。NS値は1に近ければモデルの精度がよいとされ、0.7以上で再現性が高いとされている¹⁰⁾。NS値の計算式を式(2)に示す。

$$NS = \left(1 - \frac{\sum_{i=1}^n (Q_{pre}(i) - Q_{obs}(i))^2}{\sum_{i=1}^n (Q_{obs}(i) - Q_{ave})^2} \right) \quad (2)$$

ここで、NSはNS値、 $Q_{pre}(i)$ は*i*番目の予測ダム流入量(m³/s)、 $Q_{obs}(i)$ は*i*番目の実測ダム流入量(m³/s)、 Q_{ave} は計算期間中の実測ダム流入量の平均値(m³/s)、*n*はデータ個数である。

4. 結果と考察

ダム流入量予測のNS値を図-5に示す。図中の () 内の数値はNS値の平均値である。図-5を見るといずれの結果も概ね、10時間先予測まではNS値は0.9以上であり高い再現性がある事が確認できる。また、40時間先予測までは概ねNS値は0.7以上であり再現性は良好である。これらの結果は山田ら⁷⁾による定山溪ダムでの結果と傾向は概ね一致している。

中間層を増やした、モデル2と元々のモデル1で大きな

表-1 深層学習モデルのハイパーパラメータ
(山田ら⁷⁾より改変)

	モデル1	モデル2
エポック数	50000	
バッチサイズ	64	
最適化手法	Adam	
中間層数	3	4
ユニット数	64	
中間層の活性化関数	Tanh	
損失関数	MSE	
ドロップアウト率	0.1	

差は見られず、中間層数が増えれば予測精度が向上するという傾向はなかった。中間層数については今回は1層増やただけであり、今後更なる検討が必要ではあるが3層で十分であると考えられる。仮に中間層を増やしても表現力を向上させても、過学習の要因となり予測精度向上には繋がらない事も考えられる。

入力データ時間を前24時間から前36時間及び前48時間に増やした場合においても、大きな違いは確認できない。また、入力データ時間が増えれば予測精度が増加する傾向も確認できない。そのため、さらに入力データ時間を増やして検討が必要ではあるが、入力データ時間は前24時間で十分であると考えられる。尚、今回は入力データ時間を減らす検討は行っていないため、入力データ時間を減らした場合の検討も必要である。

中間層数及び入力データ時間の検討はあくまでも忠別ダムにおける事例であり、他のダムでは影響が異なる可能性があるため、より多くのダムでの検討が必要である。また、本研究では中間層数以外のハイパーパラメータを変更していないため、他のハイパーパラメータについても検討が必要である。

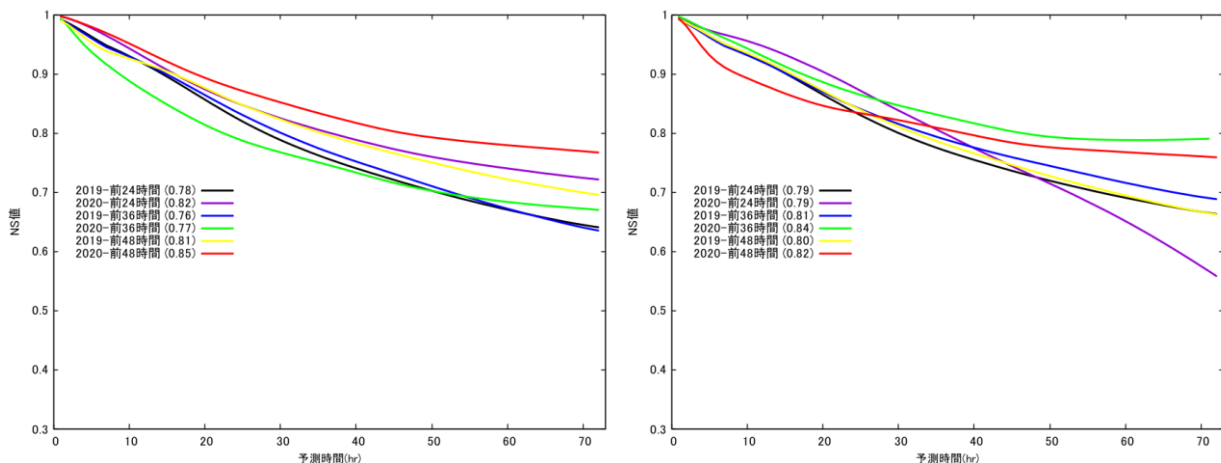


図-5 ダム流入量予測のNS値（左：モデル1 右：モデル2）

図-6には2019年にてNS値の平均が最も高かったモデル2にて前36時間のデータを入力データとした結果を、図-7には2020年にてNS値の平均が最も高かったモデル1にて前48時間のデータを入力データとした結果をそれぞれ示す。ただし、図-6及び図-7ともにダム流入量は日平均値に変換して作成している。図-6を見ると12時先予測までの結果は実測値とほぼ一致している。一方で、24時間先よりも後の予測では精度の低下が見られる。また、図-7に示した2020年の結果でも、2019年と同様の結果となっている。この結果は定山溪ダムでの結果と同様であり、実務上利用可能な予測は12時先から24時間先までと考えられる。今後、さらに適用するダムを増やす必要があるが、今回の深層学習モデルは融雪期のダム流入量予測を行うにあたり、一定の性能を有していると考えられる。

ただし、今回も学習年を機械的に設定しており多雪年・少雪年の検討は行っていない。これらの検討については今後の課題としたい。また、今回の事例では未経験規模の出水も含んでいないことから、未経験事例への適用

性についても検討が必要である。

5. まとめ

本研究では定山溪ダムにて適用した深層学習モデルを忠別ダムにて適用した。その結果、定山溪ダムの結果と同様の結果となり、実務上利用可能な予測時間は24時間先までと考えられる。また、中間層数及び入力データ時間を増やしても、予測精度は向上しなかった。

今後は、さらに適用するダムを増やして深層学習モデルの汎用性を検討する必要がある。

謝辞： 忠別ダムにおける気象観測データは国土交通省北海道開発局より提供を受けた。ここに記して謝意を表す。

参考文献

- 1) 小野延雄, 石川信敬, 新井正, 若土正暁, 青田昌秋: 雪氷水文現象 (基礎雪氷学講座VI), 前野紀一・福田正巳編, 古

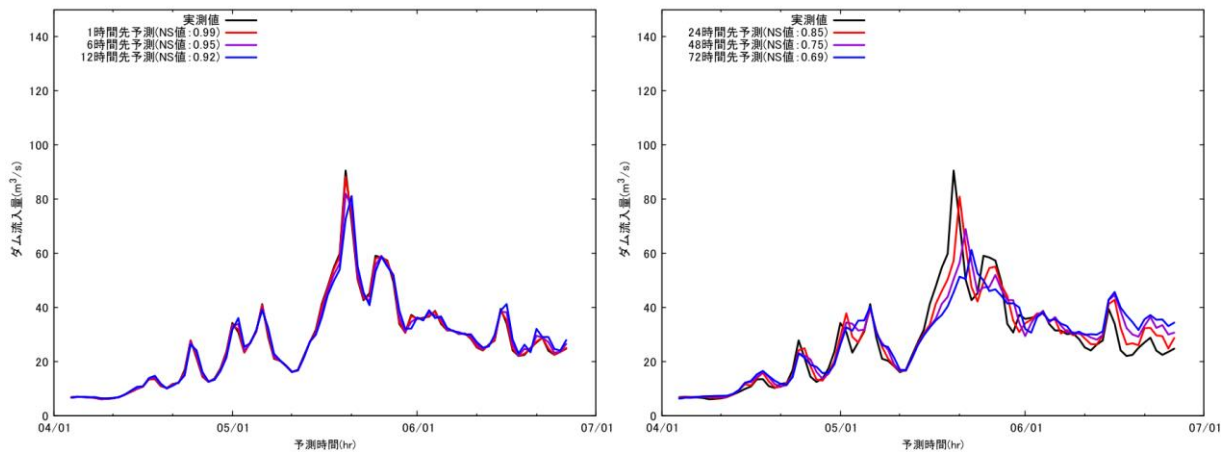


図-6 2019年におけるダム流入量の予測結果

左：1時間先予測から12時間先予測 右：24時間先予測から72時間先予測

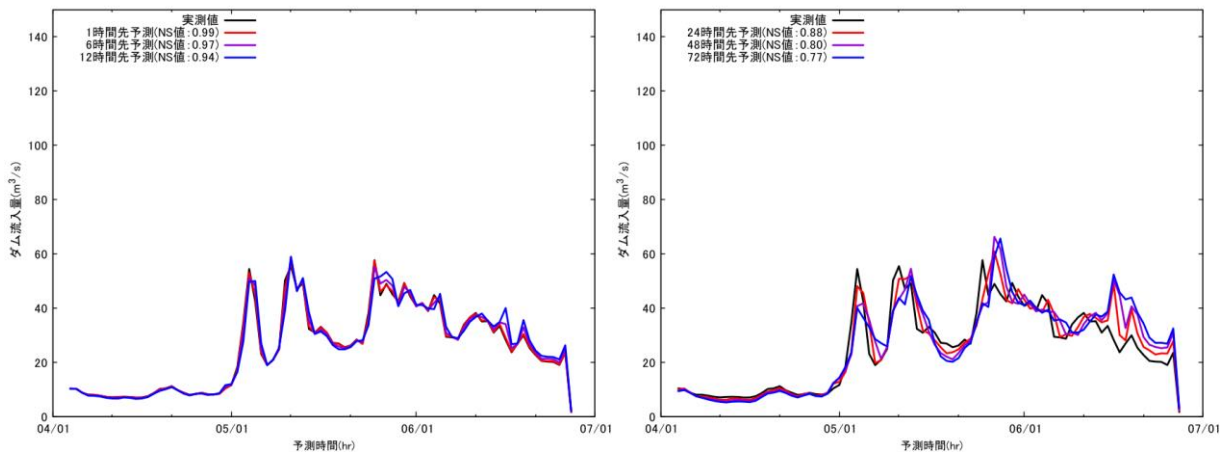


図-7 2020年におけるダム流入量の予測結果

左：1時間先予測から12時間先予測 右：24時間先予測から72時間先予測

今書院,p17,1994.

- 2) 一言正之, 櫻庭 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), Vol. 72, No.4, pp.1187-1192, 2016.
- 3) 田村和則, 加納茂紀, 三浦心, 山脇正嗣, 金子拓史: ダム流入量長時間予測への深層学習の適用ーダム防災操作の効率化を目指してー, 土木学会論文集 B1(水工学), Vol. 74, No.5, pp.11327-11332, 2018.
- 4) 樽利利晃, 萱場祐一: AI を用いた河川水位予測モデルの特異な洪水波形に対する適応性評価, 土木学会論文集 B1(水工学), Vol.77, No.2, pp.11237-11242, 2021.
- 5) 滝口修司, キムスンミン, 立川康人, 市川温, 萬和明: ニューラルネットワークを用いた積雪地域の河川流量予測における重要入力因子の抽出, 土木学会論文集 B1(水工学), Vol. 74, No.4, pp.877-882, 2018.
- 6) 込山晃市, 山本隆広, 武樋力: 深層学習を用いた融雪流出量の推定の試み, 土木学会論文集 B1(水工学), Vol. 77, No.2, pp.11225-11230, 2021.
- 7) 山田嵩, 阿部真己, 滝口大樹, 柿沼孝治: 深層学習を用いた融雪期のダム流入量予測, 土木学会論文集 B1(水工学), Vol.78, No.2, pp.1151-1156, 2022.
- 8) Hochreiter, Sepp, and Jürgen Schmidhuber, Long short-term memory, Neural computation 9.8, pp.735-1780, 1997.
- 9) Nash JE, Sutcliffe JV.: River flow forecasting through conceptual models part I-A discussion of principles, Journal of Hydrology, Vol.10, No.3, pp.282-290, 1970.
- 10) 日本学術会議: 回答 河川流出モデル・基本高水の検証に関する学術的な評価ー公開説明 (質疑) ー, p.10, 2011.