

# 土壤雨量指数の予測学習を活用した洪水予測

石井 明<sup>1</sup>・宮崎 利行<sup>2</sup>・天方 匡純<sup>3</sup>

<sup>1</sup>正会員 八千代エンジニアリング株式会社 技術創発研究所 (〒111-8648 東京都台東区浅草橋5-20-8)  
E-mail:akri-ishii@yachiyo-eng.co.jp (Corresponding Author)

<sup>2</sup>非会員 八千代エンジニアリング株式会社 技術創発研究所 (〒111-8648 東京都台東区浅草橋5-20-8)  
E-mail:ts-miyazaki@yachiyo-eng.co.jp

<sup>3</sup>正会員 八千代エンジニアリング株式会社 技術創発研究所 (〒111-8648 東京都台東区浅草橋5-20-8)  
E-mail:amakata@yachiyo-eng.co.jp

洪水到達時間が 1 時間未満の小流域において十分なリードタイムを確保し、余裕をもった防災操作や避難行動を実施するためには、高精度で実用的な洪水予測モデルが必要である。本稿では予測情報として気象庁が配信している土壤雨量指数の予測値を用いること、予測学習により水位を予測する深層学習モデルを構築することを提案し、宮ヶ瀬ダム上流域にある中津川水位観測所地点（流域面積 42.37km<sup>2</sup>、洪水到達時間 1 時間未満）を対象として予測精度を検証した。その結果、土壤雨量指数の現況が予測値と大きく乖離しない限り、6 時間先までの予測水位を高精度で予測できることを示した。

**Key Words:** river water level prediction, prediction learning, deep learning, LSTM, soil water index

## 1. はじめに

近年、2018年7月西日本豪雨や2019年10月令和元年東日本台風等のように気候変動の影響で降雨が激甚化し、河川の施設能力を超える出水で大規模な浸水被害が頻発している。施設整備は時間とコストがかかり限界もあることから、ソフト対策と両輪で被害を最小にする取り組みが展開されている。そのソフト対策の取組みの一つとしてダムや河川の流量や水位の洪水予測があり、防災操作や避難に必要なリードタイムを確保するため精度の高い洪水予測技術が要求されている。現在多くの河川では予測先行時間を 6 時間として物理現象を簡便化した流出と河道のシミュレーションモデルによる洪水予測や、シミュレーションと多くの観測所の観測データとのデータ同化技術により洪水予測を高度化する取組みが行われている。しかし、水位流量観測所がない小流域では物理モデルの構築・検証ができず、また降雨時到達時間が 1 時間未満と短いこともあり、十分なリードタイムを確保した精度の高い洪水予測を実現させることができない。一方、近年の深層学習の機運の高まりもあり、人工知能を用いた統計的なアプローチによる洪水予測技術に関する研究も多く報告<sup>1)2)3)4)</sup>されている。これらの研究はモデルの違いはあるが、流域内に多くの水文観測所があり、流域面積も比較的大きい(100km<sup>2</sup>～)流域を対象として検

討されている事例が多く、また予測時間における雨量を完全予測できたと見立てた実測学習でモデルを構築しているため小流域への実運用に向けては課題がある。

そこで本稿では、小流域を対象に予測情報を用いた深層学習モデルを構築（以下、「予測学習」という。）し、6 時間先までの予測精度検証を行う。

## 2. 既往関連研究と本研究の位置づけ

筆者らは深層学習モデルを用いて現時刻までの観測データだけでは 6 時間先の予測精度が確保できないことを確認した<sup>5)</sup>。また多くの既往研究で試行されている予測時間における雨量を完全予測できたとして実測雨量を用いる実測学習では、モデル構築時と実運用時でデータの前提条件が異なり予測精度を低下させる可能性があることを指摘した<sup>6)</sup>。

洪水時の水位の立ち上がりタイミングや降雨終了後の水位低減過程は土壤水分飽和状態と関係があり、物理モデルでは降雨と流出量の応答関係をパラメータとして設定している。6 時間先までの代表的な予測情報として降水短時間予報<sup>7)</sup>があるが、洪水期間のみを抽出して検討を行う場合、予測雨量では洪水毎に異なる初期の土壤水分飽和状態を反映できないため、本稿では土壤雨量指数

の予測情報に着目した。土壤雨量指数を用いた機械学習の検討事例は、村松ら<sup>8)</sup>の降雨後の流出量（河川総流出量）予測や坂本ら<sup>9)</sup>のダムの積算流入量予測があるが、河川水位予測に適用した事例は見当たらない。そこで本稿では降雨の累積効果を反映させた指標である土壤雨量指数の予測情報を説明変数として用い、洪水到達時間が1時間未満のダム上流山間域の小流域を対象に予測学習により1~6時間先の水位を予測するモデルを検証する。

### 3. 対象流域と使用データの概要

#### (1) 対象流域の概要

本稿では、国土交通省関東地方整備局相模川水系広域ダム管理事務所（宮ヶ瀬ダム）の中津川水位観測所地点を予測地点とした（図-1）。宮ヶ瀬ダムの流域面積（直接流域）は101.4km<sup>2</sup>であり、ダム上流域内には地上雨量観測所5箇所、水位計1箇所ある。水位観測所は宮ヶ瀬ダム上流の主要河川中津川にあり、水位観測所より上流の流域面積は42.37km<sup>2</sup>でダム上流域のほぼ中央に位置している。またダム上流域の最上流端から水位観測所までの洪水到達時間は概ね30分である。

#### (2) 土壤雨量指数

土壤雨量指数は、降雨が土壤中に水分量としてどれだけ溜まっているかを気象庁がタンクモデルを用いて数値化（各タンクの貯留高の合計）したものであり、各地の気象台が発表する大雨警報（土砂災害）や土砂災害警戒情報等の判断基準として利用されている。気象庁のタンクモデルは日本国内の陸上を対象に地表面を約5km四方の領域に分割し、それぞれの領域は直列3段タンクで、流出孔の高さ、流出係数、浸透係数の各パラメータは全国一律の値<sup>10),11)</sup>である。2009年12月から30分間隔で実測降雨（解析雨量）による現況の他、降水短時間予報を用いた6時間先までの予測値が配信されており、本稿時点では約11年分のデータが利用できる。2018年3月5日から1km格子で配信間隔を10分とした「高頻度化した土壤雨量指数」が提供されたことに伴い、今後、土壤雨量指数の配信は終了するが、約5km四方の格子領域に含まれる全ての1km格子には同じ値が格納されるため、今後も同質のデータが蓄積され学習データを増やしていくことが可能である。

宮ヶ瀬ダム上流域の約5km四方の格子分割は図-1のとおりであり、中津川水位観測所上流域にある4つの約5km四方の格子を検討に使用する。

#### (3) 検討に用いるデータと整理方法

中津川水位観測所の水位データおよび気象庁の土壤雨

量指数の現況と1~6時間先の予測値を検討に使用する。中津川水位観測所の観測水位データは水文水質データベース<sup>12)</sup>から取得した。

土壤雨量指数は2009年12月以降から整備されていること、中津川水位観測所では氾濫危険水位等の水防活動に係る基準水位が設定されていないことから、洪水時のデータとしては2010年~2020年で宮ヶ瀬ダムの洪水量である100m<sup>3</sup>/sを超過した出水を対象に、水位観測所の欠測データがある洪水を除いた29洪水を抽出した。また各洪水は洪水量超過期間の前12時間と後6時間を加えた期間を1洪水として抽出を行い、各洪水の抽出時間を均一にするため不足する期間のデータは0（ゼロ）とし、変数毎に全期間で正規化した。

学習データは2010年~2018年の26洪水、検証データは2019年の3洪水とする。なお対象とする29洪水の中津川水位観測所の観測水位は2019年10月台風19号による出水が既往最大水位（4.71m）となっており（図-2）、ピーク水位未経験洪水に対して検証をする。

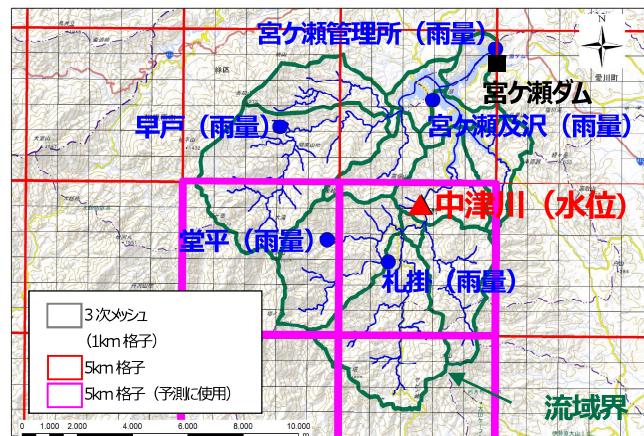


図-1 宮ヶ瀬ダム流域図

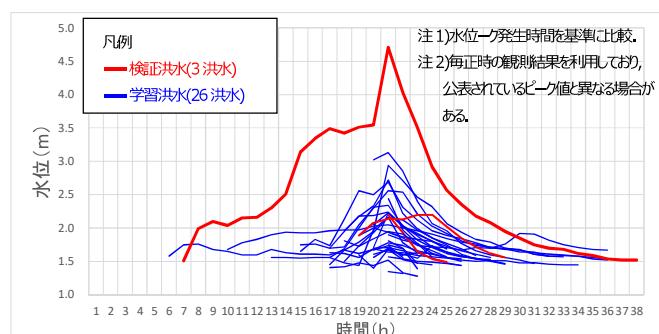


図-2 中津川観測所の洪水時の水位比較（2010年~2019年）

### 4. 深層学習モデルの構築

#### (1) モデルの構築

洪水予測では時系列データを扱うが、NN（Neural Network）やMLP（Multilayer Perceptrons），DNN（Deep Neural

Network) では入力データに時系列分の入力数を必要とし、時系列データの前後関係・連続性は保持されない。また中間層数や中間層のニューロン数は入力データ数に左右されることから<sup>13)</sup>、入力データ数を少なくすることは、ネットワーク構造の複雑化の回避になる。

そこで本稿では、時系列データを1つの入力ニューロンに対応付けることができる LSTM (Long Short-Term Memory)<sup>14)</sup>を用いて、現時刻(毎正時)から1~6時間のそれぞれの予測時間の水位を予測するモデルをTensorFlowとKerasのフレームワークを用いて構築する。LSTMを用いたモデルは図-3に示すようなネットワーク構成となり、NNやDNNの中間層のニューロンをLSTMブロックと呼ばれるメモリセルと入力、出力、忘却の3つゲートを持つブロックに置き換えることで長期依存性を高めることができる。

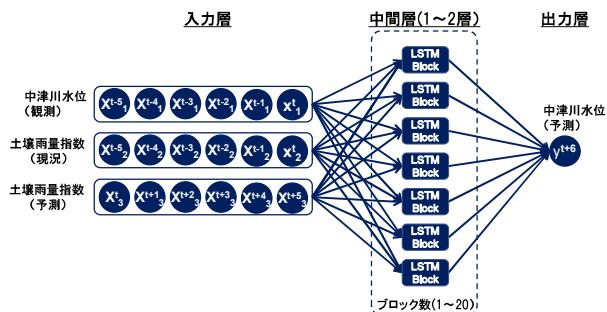


図-3 ネットワーク構成図例

## (2) 予測学習における検討ケース(予測データ設定)

説明変数に予測データを用いて予測学習を行うことが本稿の特徴でもある。現時刻において得られる予測データは1地点につき1~6時間先の6つであり、下記3ケ(図-4)で1~6時間先の予測検証を行う。なお各ケースの入出力層の変数の組み合わせを表-1に示す。

【Case1】1~6時間の連続性を重視

【Case2】予測時間毎のデータの性質を重視

【Case3】予測時間の予測データを重視

## (3) ハイパーパラメータの調整

本稿ではハイパーパラメータとして中間層数およびブロック数の組み合わせを調整する。中間層数が1~2、ブロック数が1~20の組み合わせの中から、検証データの2乗和誤差の損失関数(式1)を最小にするようにTPE(Tree-structured Parzen Estimator)で検索を行うものとした。なお中間層のブロック数は同一とする。

$$E = \frac{1}{2} \sum_{i=1}^n (Hc_i - Ho_i)^2 \quad (1)$$

ここで、 $Hc_i$ は予測水位、 $Ho_i$ は観測水位、 $i$ は時間ステップ数、 $n$ はデータ数である。

その他学習に関する各種設定は表-2に示すとおりとし、

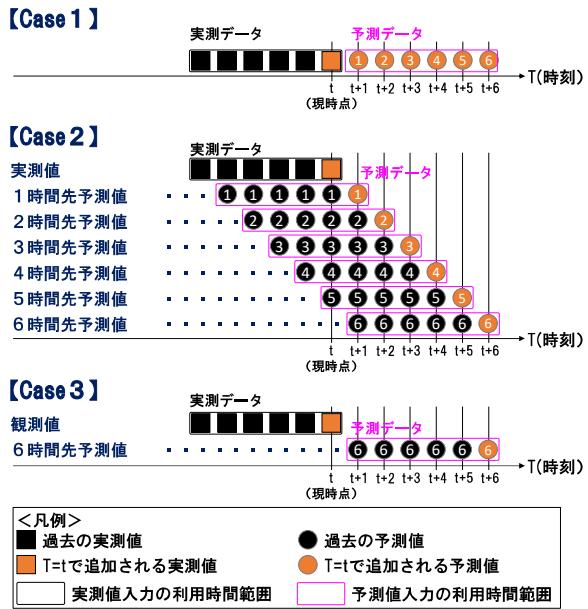


図-4 予測雨量のデータセットケース(6時間先予測の例)

表-1 T時間後予測の入出力層の変数

Case	入力(現時刻をtとする)	出力
Case1	・中津川水位(t-5~t) ・土壤雨量指数(現況)(t-5~t) ・土壤雨量指数(予測)(t-T-5~t+T)	T時間後 中津川 水位
Case2	・中津川水位(t-5~t) ・土壤雨量指数(現況)(t-5~t) ・土壤雨量指数(t~t+T予測)(t+T-5~t+T)	
Case3	・中津川水位(t-5~t) ・土壤雨量指数(現況)(t-5~t) ・土壤雨量指数(T時間後予測)(t-5~t)	

表-2 学習の各種設定

種別	設定内容
重み初期値	平均 $\mu=0$ 、標準偏差 $\sigma=0.01$ の正規分布 <sup>15)</sup>
最適化関数	Adam(学習率0.01)
活性化関数	ReLU関数
学習回数(Epochs)	200回
ミニバッチ	使用しない
ドロップアウト	なし

表-1の入出力層の変数で予測時間毎にTPEで80回検索した。表-3に学習回数が200回の中で最も損失関数が最小となった中間層数とブロック数の組み合わせを示す。また最適な中間層数とブロック数による2019年の検証3洪水(2019.5洪水、2019.8洪水、2019.10洪水)のCase毎の水位予測結果を、検討に使用している約5km四方の土壤雨量指数の4つを時刻毎に単純平均化した現況値および各予測時間の予測値と併せて図-5~図-7に示す。土壤雨量指数の現況値は観測水位と連動している状況も図より確認ができる。

2019.5洪水および2019.10洪水は土壤雨量指数の現況値と予測値に若干差異があるが、深層学習のネットワー

クで誤差を吸収し、どの予測時間も各 Case の予測水位は概ね再現性が良いことが確認できる。特に 2019.10 洪水のピーク水位は未経験洪水であったが、良好な再現結果を得られた点は有用である。一方、2019.8 洪水は 2~6 時間先の予測水位の再現性が劣っている。水位ピーク規模は 2019.5 洪水と同程度であるが、土壌雨量指数の現況値と予測値の乖離が他の 2 洪水よりも大きく、この予測誤差を深層学習のネットワークで完全に吸収できていないことが精度向上に繋がらなかった要因と推測される。2010 年～2020 年の 11 年間で洪水データは 29 洪水しかないため、今後もデータを追加して様々な洪水パターンを学習することで精度改善を検証していく必要がある。

また Case1 と Case3 に比べ Case2 では層数およびブロック数が多くなる傾向であり、入力データ数が多いほど深層学習のネットワークが複雑化することを確認できた。

## 5. 水位予測精度指標による予測結果の評価

検証洪水の予測時間毎に損失関数が最小になる各 Case について、水位予測の代表的な精度評価指標である RMSE (2 乗平均平方根誤差) (式 2a), NS (Nash-Sutcliffe 係数) (式 2b), ピーク水位相対誤差 Hp (式 2c), ピーク水位発生時間誤差 Tp (式 2d) の 4 指標で評価した結果を表-4 に示す。なお RMSE および NS は時間ステップ数の設定で値が異なるため、本稿では RMSE はダム流入量が洪水量を超えている期間、NS は洪水全抽出期間で評価を実施した。

表-3 ハイパー-パラメータの調整結果

予測時間	説明変数の数			Case1			Case2			Case3		
	Case1	Case2	Case3	損失関数	中間層数	ブロック数	損失関数	中間層数	ブロック数	損失関数	中間層数	ブロック数
1時間先	9	9	9	0.00086	1	16	0.00097	1	15	0.00097	1	15
2時間先	9	13	9	0.00214	1	9	0.00225	1	6	0.00209	2	19
3時間先	9	17	9	0.00258	2	16	0.00286	1	19	0.00219	1	11
4時間先	9	21	9	0.00242	1	11	0.00220	2	8	0.00200	1	12
5時間先	9	25	9	0.00263	1	11	0.00229	2	17	0.00180	1	20
6時間先	9	29	9	0.00258	1	11	0.00368	2	16	0.00223	1	8

表-4 予測精度評価結果

洪水	予測時間	RMSE[m]			NS[-]			ピーク水位相対誤差[m]			ピーク水位発生時間誤差[h]		
		Case1	Case2	Case3	Case1	Case2	Case3	Case1	Case2	Case3	Case1	Case2	Case3
2019.5洪水	1時間先	0.08	0.10	0.10	0.95	0.94	0.94	0.02	0.02	0.02	2	0	0
	2時間先	0.14	0.15	0.20	0.83	0.84	0.75	0.06	0.03	0.02	1	1	1
	3時間先	0.17	0.26	0.17	0.79	0.59	0.77	0.10	0.09	0.05	1	1	1
	4時間先	0.27	0.22	0.21	0.60	0.70	0.74	0.05	0.08	0.00	1	1	1
	5時間先	0.26	0.22	0.14	0.65	0.74	0.86	0.04	0.04	0.06	1	1	1
	6時間先	0.26	0.29	0.23	0.58	0.56	0.70	0.06	0.03	0.06	-1	1	0
2019.8洪水 (台風10号)	1時間先	0.17	0.23	0.23	0.91	0.84	0.84	0.05	0.03	0.03	-1	-1	-1
	2時間先	0.32	0.31	0.36	0.63	0.65	0.53	0.15	0.08	0.06	-2	-2	-2
	3時間先	0.36	0.42	0.33	0.41	0.41	0.43	0.18	0.31	0.22	-3	-3	-6
	4時間先	0.36	0.42	0.46	0.36	0.32	0.27	0.22	0.18	0.30	-4	-4	-8
	5時間先	0.41	0.45	0.31	0.24	0.32	0.45	0.16	0.25	0.27	-6	-5	1
	6時間先	0.30	0.41	0.30	0.20	0.32	0.50	0.23	0.27	0.28	6	-5	0
2019.10洪水 (台風19号)	1時間先	0.20	0.21	0.21	0.96	0.96	0.96	0.07	0.07	0.07	-1	-1	-1
	2時間先	0.29	0.31	0.25	0.92	0.90	0.93	0.17	0.13	0.09	-2	-2	-1
	3時間先	0.26	0.31	0.23	0.92	0.90	0.94	0.09	0.10	0.09	0	0	0
	4時間先	0.21	0.22	0.19	0.94	0.95	0.96	0.09	0.09	0.01	0	0	0
	5時間先	0.23	0.24	0.24	0.94	0.94	0.94	0.11	0.02	0.10	-1	-1	-1
	6時間先	0.22	0.36	0.26	0.95	0.87	0.93	0.08	0.19	0.10	0	0	-1

\*太字は 3 ケースの中で一番精度が高い箇所

\*RMSE は 0.30(m) 以下、 NS は 0.7 以上、 ピーク水位相対誤差は 0.10(m) 以下、 ピーク水位発生時間誤差は -1(h) 以上 1(h) 以下を青く着色

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Hc_i - Ho_i)^2} \quad (2a)$$

$$NS = 1 - \frac{\sum_{i=1}^n (Ho_i - Hc_i)^2}{\sum_{i=1}^n (Ho_i - Ho_{ave})^2} \quad (2b)$$

$$Hp = \frac{|Ho_p - Hc_p|}{Ho_p} \quad (2c)$$

$$Tp = To_p - Tc_p \quad (2d)$$

ここで、  $Hc$  は予測水位、  $Ho$  は観測水位、  $Ho_{ave}$  は観測水位平均値、  $i$  は時間ステップ数、  $n$  はデータ数、  $Ho_p$  は観測水位ピーク値、  $Hc_p$  は予測水位ピーク値、  $To_p$  は観測水位ピーク発生時間、  $Tc_p$  は予測水位ピーク発生時間である。  $RMSE$ 、 ピーク水位相対誤差  $Hp$ 、 ピーク水位発生時間誤差  $Tp$  は値が小さいほど精度が高く、  $NS$  は 1.0 に近いほど精度が良いとされている。

表-4 中の太字は各 Case の中で最も精度が高い箇所を示している。Case 毎の指標値の差異は僅かであり、データセットの違いによる顕著な差はないが、Case3 が比較的精度が高くなる箇所が多いことが確認できた。これは Case3 が Case1 や Case2 に比べて説明変数を持つ不確実性が少ない要因だと推定する。また表-4 では、  $RMSE \leq 0.30(m)$ ,  $NS \geq 0.7$ , ピーク水位相対誤差  $Hp \leq 0.10(m)$ , ピーク水位発生時間誤差  $-1(h) \leq Tp \leq 1(h)$  を実用的な精度条件として、精度評価指標毎に精度が確保できている箇所を青く着色した。2019.5 洪水および 2019.10 洪水では全予測時間、2019.8 洪水は 1 時間先予測に対して、実用に有用な精度を有していることが確認できた。

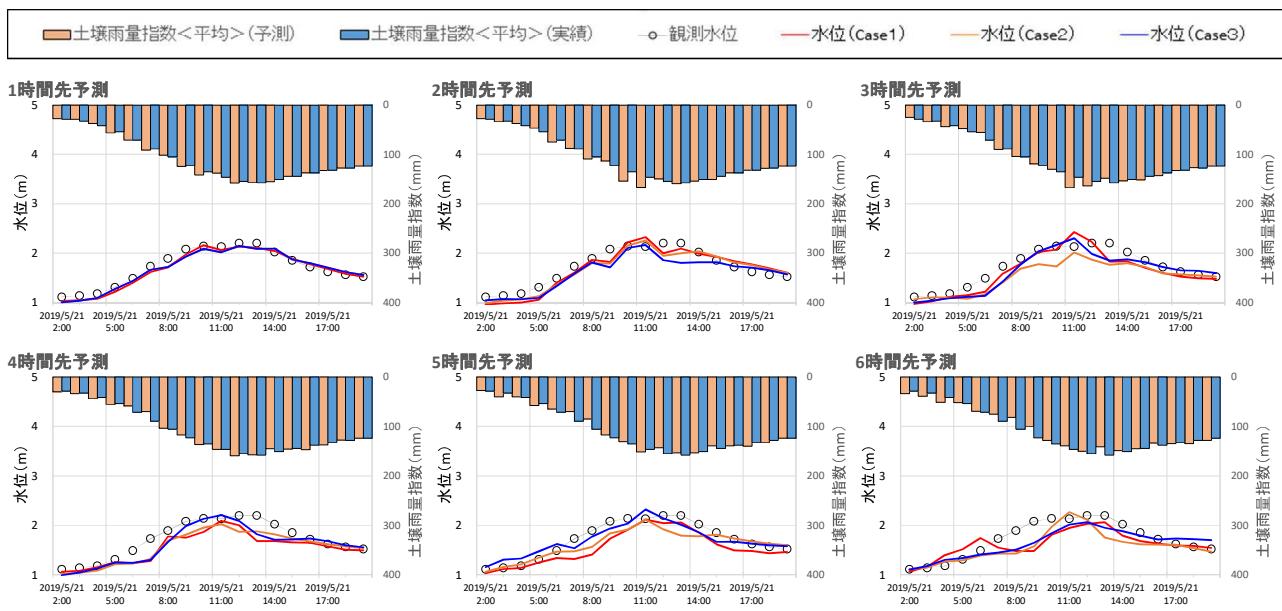


図5 予測結果の比較（2019.5洪水、前線降雨）

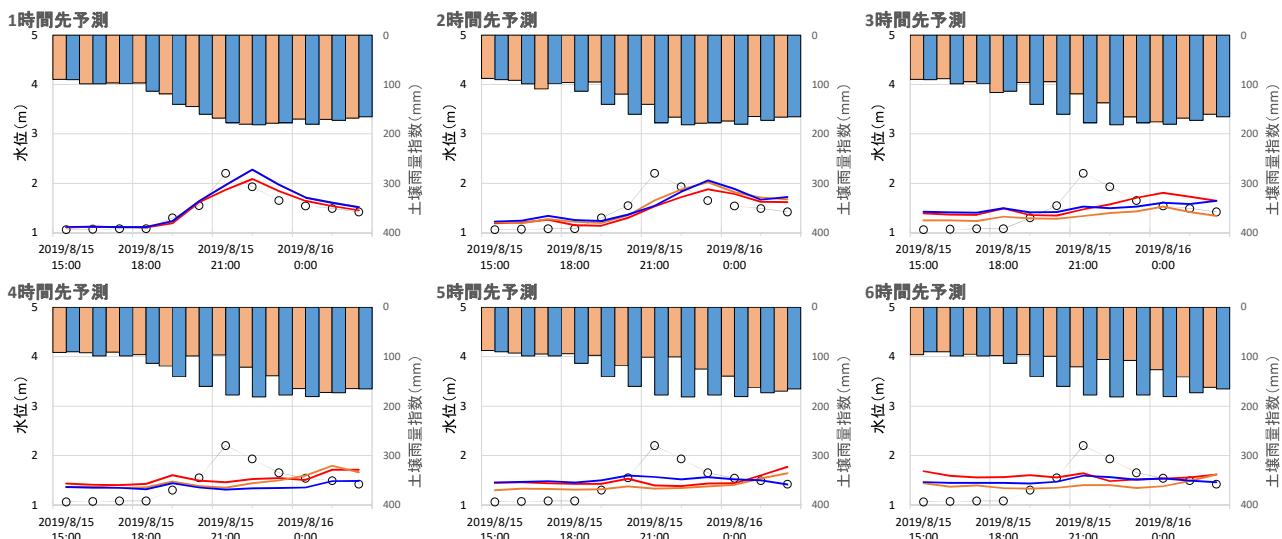


図6 予測結果の比較（2019.8洪水、台風10号）

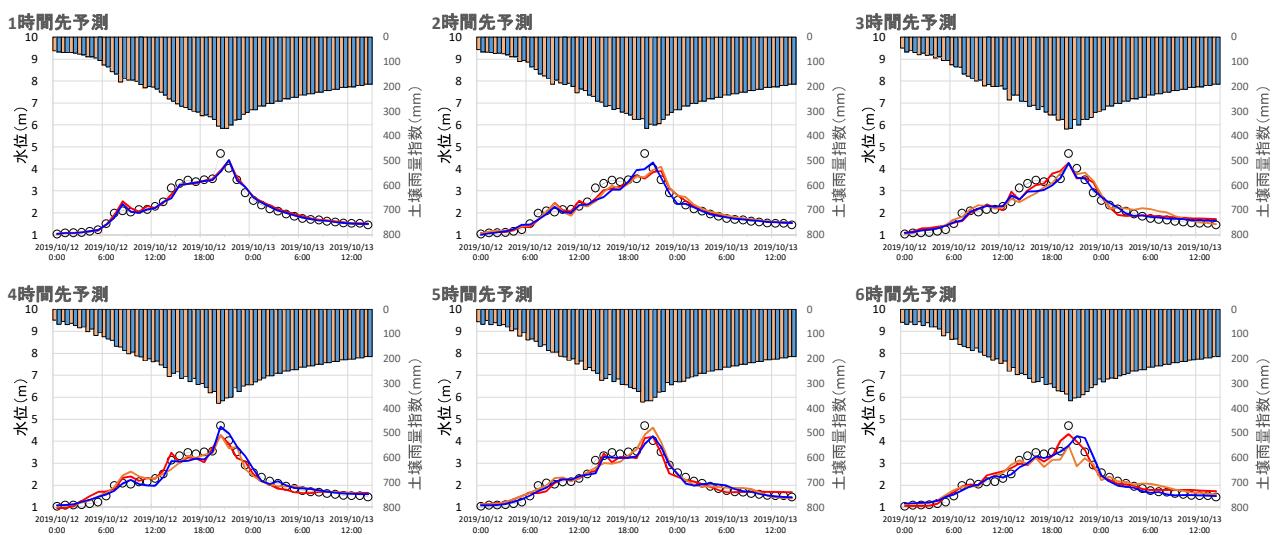


図7 予測結果の比較（2019.10洪水、台風19号）

## 6. おわりに

本稿では洪水到達時間 1 時間未満で流域面積が 42.37km<sup>2</sup> の中津川水位観測所地点の小流域を対象として、予測情報に気象庁の土壤雨量指数を用いた予測学習により 1~6 時間先の水位を予測するモデルを LSTM で構築した。その結果、土壤雨量指数を計算する降水短時間予報が解析雨量と大きな乖離がない場合は、6 時間先までの水位を高精度で予測できる結果が得られた。これより予測地点上流域内に水文情報がない小流域の場合においても、予測情報を取り込むことで 6 時間先までの水位を予測できる有用性を示した。また統計モデルの課題である未経験洪水に対しても再現性が高いことを示すことができた。

一方、降水短時間予報と解析雨量の乖離が大きい場合は土壤雨量指数の現況値と予測値も大きく乖離し、深層学習のネットワークでも誤差を処理できず予測精度が低い点は課題である。学習データが不足していることが示唆され、様々な洪水の波形や予測情報の不確定性を学習させて精度改善を今後も検討していく必要がある。

また本稿でのアプローチについては他の小流域でも検証し、有用性を確認することも重要である。更に本稿ではハイパーパラメータの調整は中間層数とブロック数に対してのみ TPE で最適化検索を実施した。予測学習に関する手法や各種設定についてはまだ改良の余地があると考えられ、知見をやデータを蓄積し、今後も精度向上に取組んでいく必要がある。

**謝辞：**国土交通省関東地方整備局相模川水系広域ダム管理事務所の皆様には、データ提供等の協力を賜りました。深く感謝の意を表します。

## 参考文献

- 1) 一言正之, 桜庭雅明 : 深層ニューラルネットワークと分布型モデルを組み合わせたハイブリッド河川水

位予測手法, 土木学会論文集 B1 (水工学), 73(1), 22-33, 2017.

- 2) 田村和則他 : ダム流入量長時間予測への深層学習の提供—ダム防災操作の効率化を目指して—, 土木学会論文集 B1 (水工学), vol.74, No.5, 1\_169-I\_174, 2018.
- 3) 山田恒輝他 : リカレントニューラルネットワークを用いた 2016 年常呂川洪水事例の水位予測, 水工学論文集 B1 (水工学), vol.74, No.5, 1\_1369-I\_1374, 2018.
- 4) 谷口純一他 : RecurrentNeuralNetwork によるダム流入量の予測, 河川技術論文集, 第 25 卷, 2019.
- 5) 天方匡純他 : IoT を想定したダム流入量予測精度向上に関する研究, 水文・水資源学会誌, vol.32 No.6, pp.287-300, 2019.
- 6) 石井明, 宮崎敏行, 天方匡純 : ダム流入量予測実運用における一要因, 土木学会第 76 回年次学術講演会, II-55, 2021.
- 7) 気象庁, 降水短時間予報, [https://www.jma.go.jp/jma/kishou/know/kurashi/kotan\\_nowcast.html](https://www.jma.go.jp/jma/kishou/know/kurashi/kotan_nowcast.html) (最終確認日 : 2021 年 6 月 30 日)
- 8) 村松翼他 : 発電用ダム上流域における土壤雨量指数を活用した流量予測手法の開発, 電気学会研究会資料, PE-11-137, pp.81-pp.86, 2011
- 9) 坂本莉子他 : 異常洪水時のダム貯水位予測に用いる機会学習手法の比較, 土木学会論文集 B1 (水工学), vol.75, No.2, 1\_85-I\_90, 2019.
- 10) 気象庁, 土壤雨量指数, <https://www.jma.go.jp/jma/kis/hou/know/bosai/dojoshisu.html> (最終閲覧日 : 2021 年 6 月 30 日)
- 11) Ishihara, Y. and S. Kobatake: Runoff Model for Flood Forecasting, Bull.D.P.R.I., Kyoto Univ., 29, 27-43, 1979.
- 12) 国土交通省, 水文水質データベース, <http://www1.ri-ver.go.jp/> (最終閲覧日 : 2021 年 6 月 30 日)
- 13) 天方匡純, 石井明, 宮崎敏行 : ダム流入量予測等に活用された深層学習モデルの整理, 土木学会第 76 回年次学術講演会, II-54, 2021.
- 14) Hochreiter, S., Schmidhuber, J. : Long Short-Term Memory, Neural Computation, 9(8), 1997.
- 15) 巢籠悠輔 : 詳解ディープラーニング, pp.185-187, マイナビ出版, 2017.

(Received June 30, 2021)  
(Accepted September 3, 2021)

## FLOOD FORECAST USING PREDICTION LEARNING OF SOIL WATER INDEX

Akira ISHII, Toshiyuki MIYAZAKI, and Masazumi AMAKATA

A highly accurate and practical flood prediction model is required to carry out disaster prevention operations and evacuation actions with a margin in small watersheds with a flood arrival time of less than one hour. In this paper, we propose to build a deep learning model that predicts the water level by using prediction learning of the soil water index. The prediction accuracy was verified at the Nakatsu River water level observatory point (basin area 42.37km<sup>2</sup>, flood arrival time less than 1 hour) in the upstream area of Miyagase Dam. As a result, the predicted water level up to 6 hours ahead could be predicted with high accuracy unless the current state of the soil water index deviates significantly from the predicted value.