

# Recurrent Neural Networkによる ダム流入量の予測

## APPLICATION OF RECURRENT NEURAL NETWORK FOR DAM INFLOW PREDICTION

谷口純一<sup>1</sup>・小島崇<sup>1</sup>・曾田康秀<sup>2</sup>・福元秀一郎<sup>3</sup>・佐藤尚<sup>4</sup>  
町田佳隆<sup>5</sup>・見上哲章<sup>6</sup>・永山正典<sup>6</sup>・錦織俊之<sup>7</sup>・渡邊明英<sup>1</sup>

Junichi TANIGUCHI, Takashi KOJIMA, Yasuhide SOTA, Shuichirou HUKUMOTO  
Hisa SATOU, Yoshitaka MACHIDA, Tetsuaki MIKAMI, Masanori NAGAYAMA  
Toshiyuki NISHIKOHI and Akihide WATANABE

<sup>1</sup>正会員 博(工) 株式会社東京建設コンサルタント環境防災研究所 (〒170-0004 東京都豊島区北大塚1-15-6)

<sup>2</sup>非会員 博(理) 株式会社東京建設コンサルタント環境防災研究所 (同上)

<sup>3</sup>正会員 学(工) 株式会社東京建設コンサルタント九州支社環境防災部 (〒812-0016 福岡市博多区博多駅南  
2-12-3)

<sup>4</sup>非会員 学(工) 株式会社東京建設コンサルタント九州支社環境防災部 (同上)

<sup>5</sup>非会員 修(工) 株式会社東京建設コンサルタント九州支社環境防災部 (同上)

<sup>6</sup>正会員 修(工) 株式会社東京建設コンサルタント環境防災事業本部環境防災部 (〒170-0004 東京都豊島区  
北大塚1-15-6)

<sup>7</sup>正会員 修(工) 株式会社東京建設コンサルタント東京本社河川計画本部 (同上)

We applied Recurrent Neural Network (RNN) and Long-Short Term Memory (LSTM) for flood forecasting problems and compared those models with Multi-Layer Perceptron (MLP) for the prediction of inflow to Yabakei Dam in the Yamakuni River system. We investigated RMSE value for the predictions of inflow against five validation floods. It was found that RNN is relatively reproducible model, although the variation due to the initial values is larger than that of MLP. We took the average for the predictions against the model parameters with different random seeds in ensemble model. Investigating the characters of these ensemble models, we got the three results: (1) The RMSE value was around or less than the median for all of the three models. (2) The average of the RMSE values against five validation floods for RNN was slightly smaller than that for MLP. (3) RNN was likely to be robust against each prediction period than MLP.

**Key Words:** flood forecasting, dam inflow prediction, deep learning, recurrent neural network

## 1. はじめに

洪水被害を最小限に抑えるためには柔軟なダム操作が重要な役割となり得る。本稿ではそのダム操作時の判断材料となるダム流入量の予測の高精度化のためにDeep Learningの手法の一つであるRecurrent Neural Network (RNN)を適応する。

近年様々な分野でDeep Learningの有用性が注目されているが、洪水予測の分野においてもMulti-Layer Perceptron (MLP)による河川水位の予測<sup>1)</sup>やダム流入量予測<sup>2)</sup>の適用事例が報告されている。Deep Learningは一般に大量のデータが必要となるが、洪水のデータは年に数回しか取れず、河川の形が変わることも考えられ、たくさん有意義なデータを用意することは難しい。その中で、洪水予測問題で扱われる特徴量は雨量や水位といっ

たデータでこれらは時系列データであるが、全層結合のMLPではその時系列のつながりを表現することはできない。そこで、本稿では限られたデータに含まれる時系列であるという情報を最大限に活かすためにRNNをこの問題に適用した。

またデータが少ない場合は、Deep Learningでは一般的なプロセスである、学習には用いないデータをモデルの汎化性を見るために使うということが難しい。そこで本稿ではモデル検証用データの振る舞いを見ずに学習をやめたとき、初期値によらず安定して良いモデルが得られかについても議論する。

## 2. 対象流域

予測対象は大分県と福岡県の県境の山国川水系の耶馬

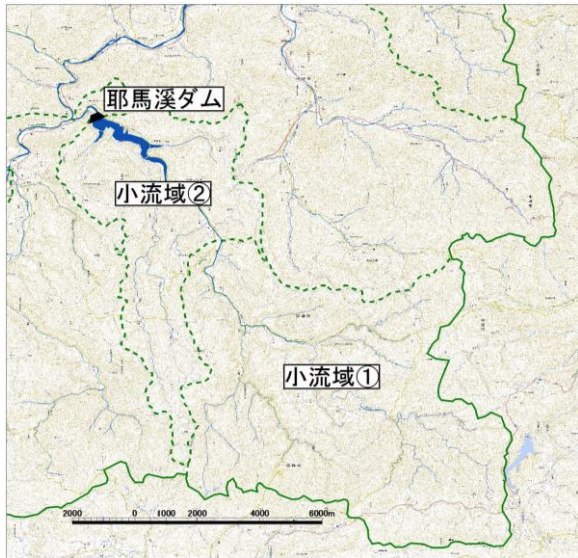


図-1 耶馬溪ダム流域の地図.

溪ダム（大分県）の流入量とした（図-1）．耶馬溪ダムより上流側を2つの小流域①（約60km<sup>2</sup>）と②（約28km<sup>2</sup>）に分割し、入力に用いる雨量はCバンドレーダ雨量の実績値から作成した小流域ごとの流域平均雨量とした．過去8時間の雨量，ダム流入量と将来6時間の予測雨量をもとに，将来6時間分のダムの流入量の予測を行う．なお入出力は10分間隔とし，予測雨量は完全予測を仮定して実績値を用いる．ダム流入量は観測データを用いた．表-1に収集したデータをまとめる．

### 3. 手法

本稿で検討した3つのモデルについて説明する．

#### (1) RNN

RNNモデルのネットワークを図-2に示す． $t = 0$ を現時刻として，過去のデータを入力するステップである  $-48 \leq t \leq 0$  のときは小流域①，②の雨量とダム流入量の3つの観測値を入力する． $1 \leq t \leq 35$  のときは雨量については予測雨量を入力し，ダム流入量は1ステップ前の予測値を用いる．時刻 $t$ の入力ベクトルを $\mathbf{x}(t)$ ，隠れ層の状態を $\mathbf{h}(t)$ とすると次式で表される．

$$\mathbf{h}(t) = \tanh(W_{xh}\mathbf{x}(t) + W_{hh}\mathbf{h}(t-1) + \mathbf{b}_h) \quad (1)$$

ここで， $W_{xh}$ ， $W_{hh}$ ， $\mathbf{b}_h$ は学習する重み行列とバイアスのベクトルである．また，次の時刻のダム流入量の予測値 $\mathbf{y}(t+1)$ は，既往研究<sup>2)</sup>を参考にダム流入量の変動量を予測するように以下の式で求める．

$$\mathbf{y}(t+1) = W_{hy}\mathbf{h}(t) + \mathbf{b}_y + \mathbf{y}(t) \quad (2)$$

$W_{hy}$ と $\mathbf{b}_y$ も学習する重み行列とバイアスのベクトルである．なお $\mathbf{y}(0)$ は現時刻のダム流入量の観測値とする．

表-1 収集した洪水データ．赤字は検証用洪水．

発生日	最大流入量(m <sup>3</sup> /s)	順位
2008_0611	122.95	18
2008_0619	267.16	6
2008_0621	107.33	21
2009_0630	190.82	9
2009_0725	202.65	7
2010_0523	104.59	23
2010_0714	116.01	19
2011_0630	14.55	30
2011_0711	4.27	31
2011_0725	43.07	29
2012_0616	102.04	24
2012_0624	315.2	5
2012_0703	1106.57	1
2012_0714	691.48	2
2013_0620	140.8	13
2013_0626	105.4	22
2013_0901	158.99	11
2013_0904	152.73	12
2014_0622	126.33	16
2014_0703	79.18	28
2014_0707	135.02	14
2015_0403	88.4	27
2015_0603	96.5	26
2015_0611	108.46	20
2015_0701	131.81	15
2015_0825	190.84	8
2016_0621	375.07	4
2016_0713	96.85	25
2016_0920	125.82	17
2016_0929	178.59	10
2017_0705	552.37	3

#### (2) LSTM (Long-Short Term Memory)

RNNでは長期記憶が必要な問題が難しいことや，勾配消失が起こりやすい等いくつかの問題点が知られており，それらを解決するために開発されたのがLSTMである．モデル全体としては図-2と同じ形であるが，隠れ層 $\mathbf{h}(t)$ の計算がGers et al.<sup>3)</sup>の形になる．

$$\begin{cases} \mathbf{f}_t = \text{sigm}(W_f[\mathbf{h}(t-1), \mathbf{x}(t)] + \mathbf{b}_f) \\ \mathbf{i} = \text{sigm}(W_i[\mathbf{h}(t-1), \mathbf{x}(t)] + \mathbf{b}_i) \\ \tilde{\mathbf{C}} = \tanh(W_c[\mathbf{h}(t-1), \mathbf{x}(t)] + \mathbf{b}_c) \\ \mathbf{C}(t) = \mathbf{f} * \mathbf{C}(t-1) + \mathbf{i} * \tilde{\mathbf{C}} \\ \mathbf{o} = \text{sigm}(W_o[\mathbf{h}(t-1), \mathbf{x}(t)] + \mathbf{b}_o) \\ \mathbf{h}(t) = \mathbf{o} * \tanh(\mathbf{C}(t)) \end{cases} \quad (3)$$

ここでsigmはシグモイド関数， $[\mathbf{a}, \mathbf{b}]$ はベクトル $\mathbf{a}$ ， $\mathbf{b}$ の結合，\*は要素ごとの積を表す．式(3)中の $W$ や $\mathbf{b}$ を学習する． $\mathbf{C}(t)$ はCell Stateと呼ばれるものであり，長期記憶や勾配消失を避けるのに役立っているとされている．

RNN, LSTMの重み行列の初期値はGlorot and Bengio<sup>4)</sup>の方法を用いて決定した式(4)． $j$ 層目から $j+1$ 層目への順伝播に用いる重み行列を $W$ とすると，

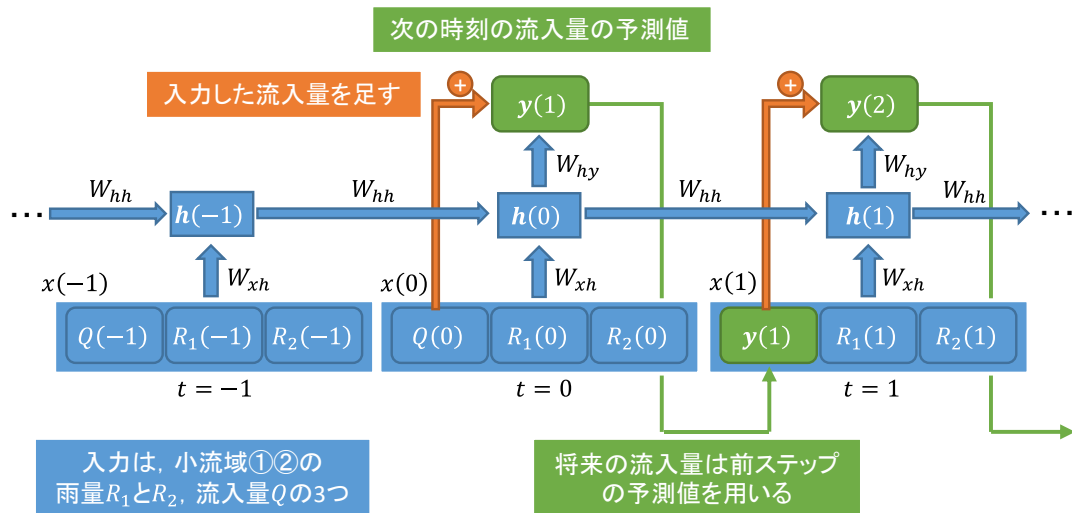


図-2 RNNによるダム流入量予測のモデル図。

表-2 既往のMLPモデルとの違いとその理由。

一言ら <sup>2)</sup>	本稿	理由
入力は相関分析である程度選ぶ。	RNNと同じ入力すべてを用いる。	Deep Learningでは多数の特徴量から必要な特徴量を獲得することができる(例えば <sup>例はあ</sup> )。
1時刻後,2時刻後...の予測のためにそれぞれ異なるモデルを作成。	出力層を複数にし、一つのネットワークで全時刻の予測を行う。	十分大きなネットワークであれば1時刻ごとの予測モデルも内包していると考えられるため。
学習の効率化のために初期値をオートエンコーダによって決定。	活性化関数にReLUを用いた。	効率的な学習が行えたとされるReLUを用いた。

$$W \sim U \left[ -\frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}, \frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}} \right] \quad (4)$$

ここで、 $U[a, b]$  は  $a$  から  $b$  までの一様乱数を示し、 $n_j$  は  $j$  層目のノード数を示す。この時、 $j$  層目のノード数は、順伝播に関わる全てのノードの数となるため、例えばRNNの1層目の隠れ層の計算の場合、 $n_1$  は入力ノード数と隠れ層のノード数の和となり、 $n_2$  は隠れ層のノード数となる。

バイアスの初期値は、LSTMの $b_f$ は1、その他は0とした。また、RNN、LSTMの $h$ や $C$ は予測時に初期状態( $h(-49)$ や $C(-49)$ )を決める必要があるが、ここではその初期状態も学習させる。 $h$ 、 $C$ の初期状態は0で初期化した。

### (3) MLP

MLPは既往研究の一言らのモデル<sup>2)</sup>を参考にし、説明

表-3 モデル名とその詳細。MLPは中間層2層とした。

モデル名	モデル	ノード数	ドロップアウト率
A	MLP	200-200	0.0
B	MLP	200-200	0.2
C	MLP	200-200	0.4
D (MLP_S)	MLP	100-100	0.0
E	MLP	100-100	0.2
F	MLP	100-100	0.4
G	RNN	50	0.0
H	RNN	50	0.2
I (RNN_S)	RNN	5	0.0
J	LSTM	50	0.0
K	LSTM	50	0.2
L	LSTM	5	0.0

変数に10分ごとのダム流入量の変化量を加え、予測変数も10分ごとのダム流入量の変化量とした。一言らのモデルとの違いを表-2にまとめる。

重み行列の初期化は活性化関数がReLUのときに広く用いられているHe et al.<sup>9)</sup>の方法を用いた(式(5))。

$$W \sim N \left( 0, \frac{2}{n_j} \right) \quad (5)$$

ここで、 $N(a, b)$ は平均 $a$ 分散 $b$ の正規乱数である。

入出力データはすべての特徴量について平均0分散1になるように規格化した。学習はmini-batch勾配降下法で行い、バッチサイズは64、学習率の調整にはAdamアルゴリズム<sup>7)</sup>を用いた。またロス関数はMSEとした。

表-3に検討を行った各モデルの概要をまとめる。ドロップアウト<sup>8)</sup>はMLPについては一般的な方法をとるが、RNNやLSTMに対してはGal and Ghahramani<sup>9)</sup>の方法を用いる。また、RNN、LSTMのノード数5のモデルについては規模が小さなモデルであるためドロップアウトの適応はしない。学習はどのモデルも予備検討により十分な収束が確認できた500エポック行い、学習用洪水に対するロスが最も小さくなったタイミングでのモデルを採用す

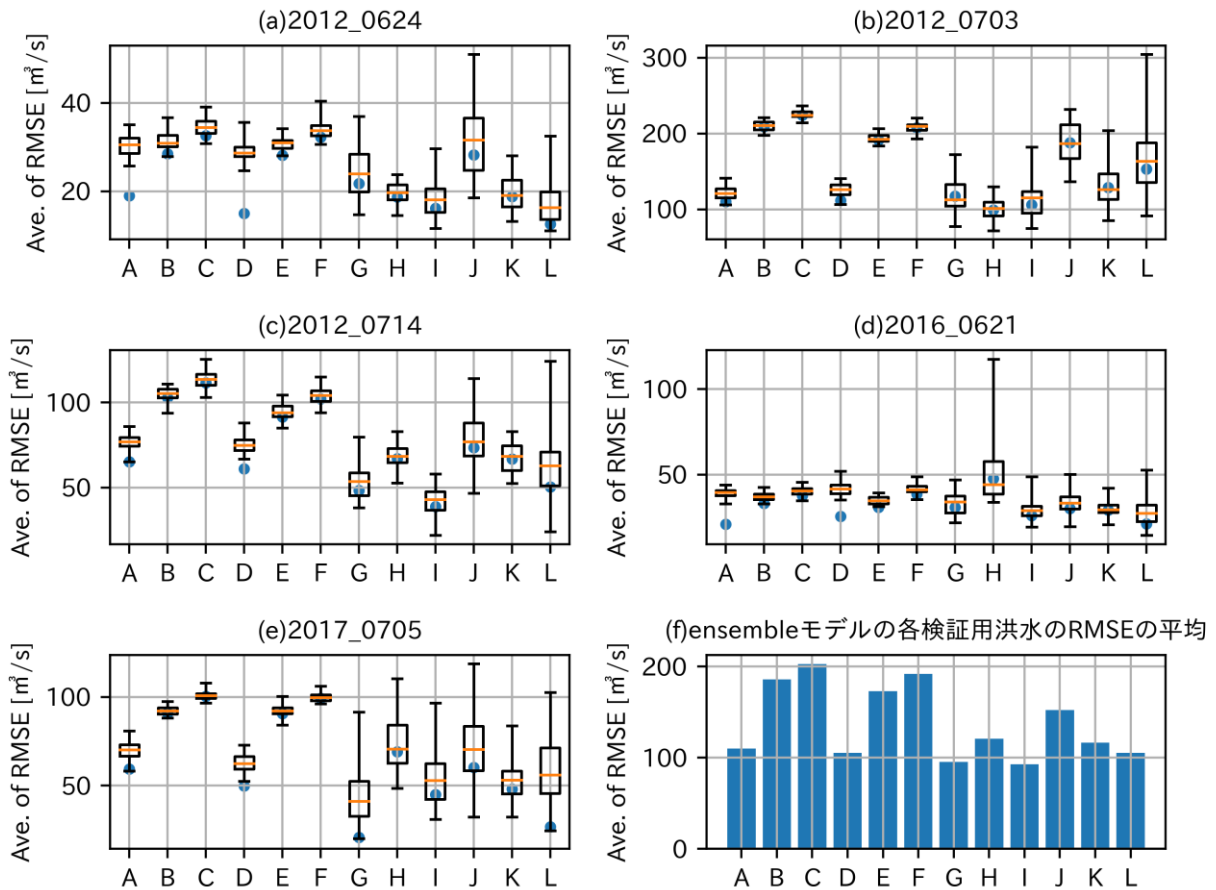


図-3 (a)~(e)は各モデルの各検証用洪水に対する各予測時刻のRMSEの平均値の箱ひげ図（最小値・第1四分位点・中央値・第3四分位点・最大値を示している）．図中の青丸は各モデル50種のランダムシードから得られたモデルによるアンサンブルモデルのRMSEの平均値．(f)はアンサンブルモデルの検証用洪水5洪水のRMSEの平均値（青丸の平均値）．

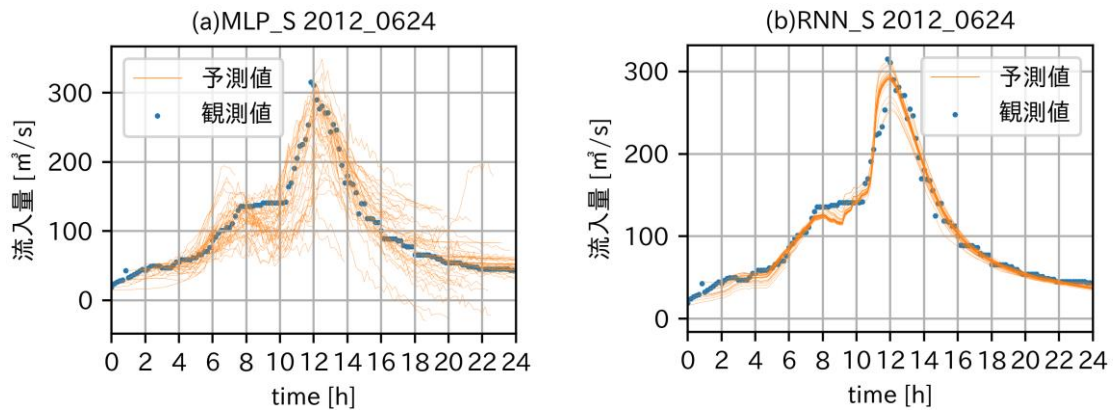


図-4 2012年6月24日出水に対してD(MLP\_S)でRMSEが最も小さいモデル(a)とI(RNN\_S)でRMSEが最も小さいモデル(b)の予測検証の結果．

る．これは、RNNやLSTMの学習では学習用洪水に対するロスはトレンドとしては減少するが、大きく跳ね上がることがあるなど単調減少とはならないためである．MLPについても同様のことが起こりうるため同じようにモデルを選択する．それぞれのモデルで50ケースずつランダムシードを変えて学習を行った．ランダムシードを変えることによって各パラメータの初期値や、学習時の

mini-batchの順番、ドロップアウトされるノードが変わるため、異なるモデルが得られる．

#### 4. 結果

現時刻がダム流入量のピーク前後12時間のときに注目

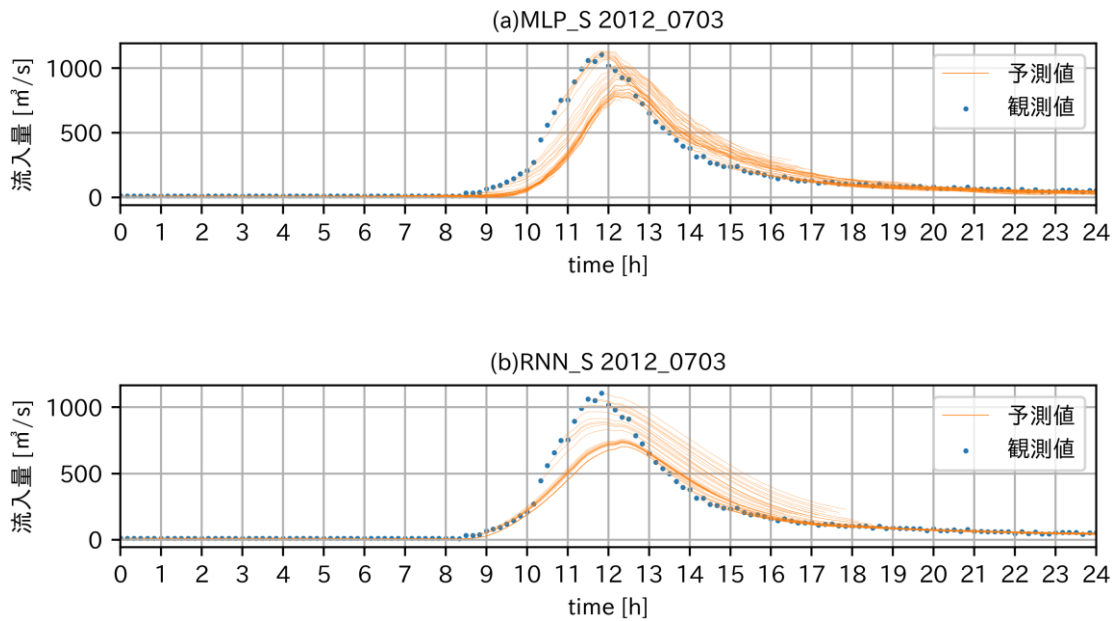


図-5 D(MLP\_S)のアンサンブルモデル(a)と, I(RNN\_S)のアンサンブルモデル(b)による2012年7月3日出水の予測検証の結果.

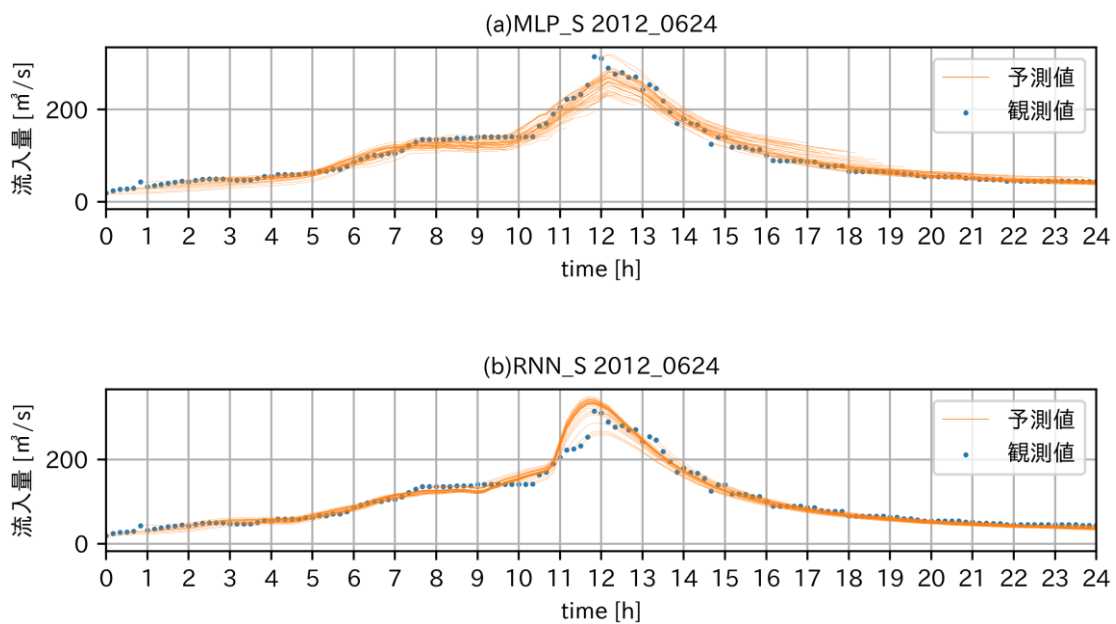


図-6 D(MLP\_S)のアンサンブルモデル(a)と, I(RNN\_S)のアンサンブルモデル(b)による2012年6月24日出水の予測検証の結果.

して解析を行った. 各モデル・各検証用洪水に対する学習したモデルの分布を図-3に示す. なお, 図中のAve. of RMSEは, 以下の式で表される.

$$\text{Ave. of RMSE} = \frac{\sum_{T=1}^{36} \sqrt{\frac{\sum_{t=1}^{145} (o_{T+t}^T - p_t^T)^2}{145}}}{36} \quad (6)$$

ここで,  $p_t^T$ は時刻 $T$ の $t$ 時刻後の予測値,  $o^T$ は時刻 $T$ の観測値である.  $t$ はピーク時刻の12時間前(72時刻前)を1として, ピーク時刻の12時間後の145までを考える.

MLPのRMSEの分布は全体的にばらつきが小さいが,

その中央値はRNNやLSTMより高い傾向にある. モデルI(RNN-S)のRMSEの最大値はMLPよりも大きくなることはあるが, 第3四分位点は悪いものでMLPの中央値程度の値である. したがってRNNが再現性の高いモデルを作りやすいといえる.

また, ランダムフォレストに代表されるように, 機械学習の手法は一般に様々なモデルの平均を取るといったアンサンブルモデルの方が精度が良いとされている. 本稿では各モデル50種類のランダムシードでモデルを作成しているので, 各モデルについてこれらの単純平均を取るモデルの精度についても調べた(図-3の青丸や棒グラフ

フ)．その結果どのモデルについても中央値程度かそれ以上の精度となった．これはRNNモデルのようにランダムシードによって得られるモデルにばらつきがある場合、それらの平均を取ることによってそのばらつきをある程度無視できる可能性があることを示唆している．またMLPの改善の幅が比較的大きい．これはMLPの予測は時系列を仮定していないためか、その予測波形が安定しないが(図4．各時刻から予測を開始しその波形を描いている)、ランダムシードごとの平均を取ることによって平滑化のような作用を及ぼしたことが要因であると考えられる．

最後にモデルD (MLP\_S)とモデルI (RNN\_S)の予測波形の比較を行う．対象は予測のRMSEが最も大きかったピーク流入量1位の2012年7月3日出水(図-5)と、最も小さかったピーク流入量5位の2012年6月24日出水(図-6)である．2012年7月3日出水については、外挿領域となっていると考えられ、どちらのモデルも波形全体を表現することはできていないが、MLPではピーク2時間前からはピーク流量を捉えられている．一方で、RNNの方が立ち上がりの10hあたりまでの予測は高精度である．2012年6月24日出水については、全体のRMSEはMLPの方が小さい．一方で、RNNの方がピーク流量やその時刻は再現できている．さらに予測波形のばらつきは小さいため、RNNの方がMLPよりも予測開始時刻が変わったときに予測値が変動しにくいといえる．これは、MLPでは入出力を時系列データとして扱えないことが原因である可能性がある．しかし、RNNではピーク時刻の10~50分前の予測精度は良くない．これは、この時刻の波形はその前後の波形と比べて傾きが緩やかであるため、本項のRNNによる予測モデルでは急激な傾きの変化を表現できない可能性を示唆している．

## 5. まとめ

本稿ではダム流入量の予測にMLP, RNN, LSTMを適用した．その結果、RNNを用いると再現性の高いモデルが比較的作りやすいことが分かった．しかし、初期値によるばらつきがMLPよりも大きく、悪いモデルが選ばれた場合はMLPよりもRMSEが大きくなることもある．

各モデルについて異なるランダムシードで得られたモデルの単純平均を取るアンサンブルモデルを検証したところ、どのモデルについても中央値程度かそれ以下のRMSEとなることが分かった．特にMLPでは精度の改善幅が大きかった．このアンサンブルモデルを比較したところ、MLPよりもRNNの方がややRMSEは小さくなった．予測波形を比較すると、既往最大洪水の再現性はどのモデルも良くないが、MLPはピークの2時間前程度からはピーク水位を予測することができた．しかしMLPは予測開始の時刻が変わるごとに予測結果も変わりやすい傾向

があることが分かった．RNNは構造的にそのようなことが起こりにくい点でこういった点ではRNNの方が優位である．

LSTMは一般的にはRNNより高性能であるとされているが、本稿ではその優位さは確認できなかった．これは本稿で取り上げた問題が、LSTMが良く用いられている機械翻訳などのタスクに比べてシンプルで、それほど高性能なモデルが必要ないためであるかもしれない．

今後の課題として、シード値によってばらつきが大きくなならないようなモデルや学習の仕方の検討があげられる．RNNではアンサンブルを取る際にRMSEが大きなモデルも含んでいるため、得られるモデルの誤差のばらつきが小さくなればより良いモデルが作れるようになる可能性がある．また、本稿では流域分割した小流域の平均雨量を入力に用いたが、レーダ雨量をそのまま用いることで精度が向上する可能性がある．他に、RNNの入力にもMLPと同様に流入量の変化量を用いる等の工夫により、より大規模な洪水を表現できる可能性がある．

## 参考文献

- 1) 一言正之, 桜庭雅明, 清雄一: 深層学習を用いた河川予測手法の開発, 土木学会論文集B1(水工学), Vol.72, No.4, I\_187-I\_192, 2016.
- 2) 一言正之, 遠藤優斗, 島本卓三, 房前和朋: レーダ雨量を用いた深層学習によるダム流入量予測, 河川技術論文集, Vol.24, pp.403-408, 2018.
- 3) Felix A. Gers, Jürgen Schmidhuber, Fred Cummins, Learning to Forget: Continual Prediction with LSTM, Neural Computation, Vol. 12, Issue 10, pp. 2451-2471, 2000.
- 4) Xavier Glorot, Yoshua Bengio: Understanding the difficulty of training deep feedforward neural networks, Journal of Machine Learning Research, Vol. 9, pp.249-256, 2010.
- 5) 麻生英樹: 多層ニューラルネットワークによる深層表現の学習, 人工知能学会誌, Vol.28, No. 4, pp649-659, 2013.
- 6) Kaiming He, Xiangyu Zhang Shaoqing Ren Jian Sun: Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, 2015 IEEE International Conference on Computer Vision, pp.1026-1034, 2015.
- 7) Diederik P. Kingma, Jimmy Lei Ba: Adam: A Method for Stochastic Optimization, Proceedings of the 3rd International Conference on Learning Representations (ICLR), 2015.
- 8) Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research, Vol. 15, pp.1929-1958, 2014.
- 9) Yarın Gal, Zoubin Ghahramani: A Theoretically Grounded Application of Dropout in Recurrent Neural Networks, 30th Conference on Neural Information Processing Systems, pp.1027-1035, 2016.

(2019. 4. 2受付)