Recurrent Neural Networkによる ダム流入量の予測 APPLICATION OF RECURRENT NEURAL NETWORK

FOR DAM INFLOW PREDICTION

谷口純一¹・小島崇¹・曽田康秀²・福元秀一郎³・佐藤尚⁴ 町田佳隆⁵・見上哲章⁶・永山正典⁶・錦織俊之⁷・渡邊明英¹ Junichi TANIGUCHI, Takashi KOJIMA, Yasuhide SOTA, Shuichirou HUKUMOTO Hisa SATOU, Yoshitaka MACHIDA, Tetsuaki MIKAMI, Masanori NAGAYAMA Toshiyuki NISHIKOHRI and Akihide WATANABE

 「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) ²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)
3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)
⁴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)
⁵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)
⁶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)
⁷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¹⁾やダム流入量予 測²⁾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 さんの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 耶馬渓ダム流域の地図.

渓ダム(大分県)の流入量とした(図-1). 耶馬渓ダム より上流側を2つの小流域①(約60km²)と②(約28km²)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 ≤ $t \le 0$ 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. $1 \le t \le 35$ 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x(t),隠れ 層の状態をh(t)とすると次式で表される.

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

ここで、 W_{xh}, W_{hh}, 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 y(t+1)は、既往研究²⁾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.

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

 W_{hy} と 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y(0)は現時刻のダム流入量の観測値とする.

最大流入量(㎡/s) 122.95	順位
122.95	10
	18
267.16	6
107.33	21
190.82	9
202.65	7
104.59	23
116.01	19
14.55	30
4.27	31
43.07	29
102.04	24
315.2	5
1106.57	1
691.48	2
140.8	13
105.4	22
158.99	11
152.73	12
126.33	16
79.18	28
135.02	14
88.4	27
96.5	26
108.46	20
131.81	15
190.84	8
375.07	4
96.85	25
125.82	17
178.59	10
552.37	3
	267.16 107.33 190.82 202.65 104.59 116.01 14.55 4.27 43.07 102.04 315.2 1106.57 691.48 140.8 105.4 158.99 152.73 126.33 79.18 135.02 88.4 96.5 108.46 131.81 190.84 375.07 96.85 125.82 178.59 552.37

(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h(t) の計算がGers et al.³の形になる.

$$\begin{cases} f_t = \operatorname{sigm}(W_f[h(t-1), x(t)] + b_f) \\ i = \operatorname{sigm}(W_i[h(t-1), x(t)] + b_i) \\ \widetilde{C} = \operatorname{tanh}(W_C[h(t-1), x(t)] + b_c) \\ C(t) = f * C(t-1) + i * \widetilde{C} \\ o = \operatorname{sigm}(W_o[h(t-1), x(t)] + b_o) \\ h(t) = o * \operatorname{tanh}(C(t)) \end{cases}$$
(3)

ここでsigmはシグモイド関数, [a,b]はベクトルa, bの 結合,*は要素ごとの積を表す.式(3)中のWやbを学習 する. 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 Bengio⁴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2)	本稿	理由
入力は相関分析	RNN と同じ入力	Deep Learningで
である程度選	をすべて用い	は多数の特徴量
ぶ.	る.	から必要な特徴
		量を獲得するこ
		とができると言
		われている ^{例えば5)} .
1時刻後,2時刻後	出力層を複数に	十分大きなネッ
…の予測のため	し,一つのネッ	トワークであれ
にそれぞれ異な	トワークで全時	ば1時刻ごとの予
るモデルを作	刻の予測を行	測モデルも内包
成.	う.	していると考え
		られるため.
学習の効率化の	活性化関数に	効率的な学習が
ために初期値を	ReLU を 用 い	行えるとされる
オートエンコー	た.	ReLU を 用 い
ダによって決		た.
定		

$$W \sim U\left[-\frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}, \frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}\right]$$
 (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2)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
А	MLP	200-200	0.0
В	MLP	200-200	0.2
С	MLP	200-200	0.4
D (MLP_S)	MLP	100-100	0.0
Е	MLP	100-100	0.2
F	MLP	100-100	0.4
G	RNN	50	0.0
Н	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**.^のの方法を用いた(式(5)).

$$W \sim N\left(0, \sqrt{\frac{2}{n_j}}\right)$$
 (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⁷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⁸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 Ghahramani⁹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.

time [h]



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.

1

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

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). 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 が起こりにくいためこういった点では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 機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

- 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.
- 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.
- Felix A. Gers, Jürgen Schmidhuber, Fred Cummins, Learning to Forget: Continual Prediction with LSTM, Neural Computation, Vol. 12, Issue 10, pp. 2451-2471, 2000.
- Xavier Glorot, Yushua 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.
- Diederik P. Kingma, Jimmy Lei Ba: Adam: A Method for Stochastic Optimization, Proceedings of the 3rd International Conference on Learning Representations (ICLR), 2015.
- 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.
- Yarin 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)