

レーダ雨量を用いた深層学習による ダム流入予測

DAM INFLOW PREDICTION BY DEEP LEARNING USING RADAR RAINFALL

一言正之¹・遠藤優斗²・島本卓三³・房前和朋³

Masayuki HITOKOTO, Yuto ENDO, Takuzo SHIMAMOTO and Kazutomo FUSAMAE

¹正会員 博士(工) 日本工営株式会社 中央研究所 (〒300-1259 茨城県つくば市稲荷原2304)

²学生会員 長岡技術科学大学 (〒940-2188 新潟県長岡市上富岡町1603-12304)

³正会員 国土交通省 九州地方整備局 九州技術事務所 (〒830-8570 福岡県久留米市高野1-3-1)

We developed the deep learning model for dam inflow prediction. The input data of the model was radar rainfall. We applied the model to the Shimouke Dam and confirmed the good result up to 60 minutes prediction. We compared the deep learning model with conventional 3-layer neural network, and confirmed the superiority of the deep learning. We also compared the developed model with the deep learning model trained with rain-gauge data, and confirmed the superiority of the radar rainfall. In the end, we applied the prediction rainfall data to the input data of the model. We confirmed that there was no effect of the rainfall prediction error up to 20 minutes prediction in this model. In 60 minutes prediction, effect of the rainfall error appeared especially in the rising limb of the flood.

Key Words : dam inflow prediction, real-time prediction, deep learning, artificial neural network

1. はじめに

洪水時のダム運用のために、ダム流入量予測の高精度化は重要である。洪水予測手法の一つとして、ニューラルネットワーク (Artificial Neural Network : ANN) を用いた手法がある。これまでの研究で、様々な河川において洪水時の水位予測へのANNの適用性が報告されている^{1), 2), 3)}。さらにANNに深層学習⁴⁾を適用することにより、洪水予測の精度向上が報告されている^{5), 6), 7)}。

ダム流入量予測においてもANNを適用した例が数多く報告されている^{8), 9), 10)}。またダム流入量予測に深層学習を適用した例として、Baiらは中国の三峡ダムを対象に日流入量予測を行っている¹¹⁾。しかしながら、Baiらの研究は日流入量のみを学習データに用いており、データの時系列的なトレンドから自己回帰的に予測を行うものである。そのため、降雨-流出応答を予測に反映できておらず、流入量ピークの予測が実績よりも小さくなる傾向が見られる。

本研究ではダム流域の降雨-流出応答に着目して、雨

量データを主な入力データとする深層ANNにより、ダム流入量予測モデルを構築する。なお、これまでの深層学習を用いた水位予測モデルでは、地上雨量データをモデルの入力に用いている^{5), 6), 7)}。地上雨量に代えて、時間・空間的に解像度の高いレーダ雨量を用いることで、さらなる精度向上の可能性がある。本研究では比較検討として、主な入力データにレーダ雨量を用いた場合と、地上雨量を用いた場合との精度検証を行う。また深層学習を適用しない従来型ANNをあわせて構築し、比較検討を行うことで深層学習の効果を検証する。これらの検討により、①従来型ANNに比べ深層ANNの予測精度が高いこと、②地上雨量に比べレーダ雨量を用いたモデルの方が予測精度が高いことを確認する。

2. 検討流域と対象洪水

対象流域は、筑後川水系の下笠(しもうけ)ダム流域とした(図-1)。流域面積は185km²であり、雨量観測所が4地点、水位観測所が3地点に設置されている(表-1)。



図-1 下笠ダム流域と観測所位置および小流域番号。

表-1 予測対象地点の概要。

予測地点名	流域名	流域面積	上流の水位 観測所数	雨量観測所数
下笠ダム	筑後川	158.0	3	4

また、ダムへの流入量は常時計測されている。

検討対象洪水として、レーダ雨量の入手できる2006年以降の22洪水を抽出した。対象洪水の一覧を表-2に示す。検討対象洪水について、地上雨量データ、レーダ雨量データ（国土交通省のCバンド同時刻合成レーダ雨量）を収集した。レーダ雨量は全て5分間隔、地上雨量は2015年6月半ば以降が10分間隔データ、それ以外は1時間間隔データである。以降の検討では、60分データは線形内挿により10分データに置き換えた。レーダ雨量は、前10分間のデータを平均することで10分データに換算した。ダム流入量データは全て10分間隔の実測値を用いた。

洪水対象期間は、ピークの12時間前から8時間後までとした。したがって、22出水×20時間10分=443時間10分、10分データで2662組のデータを検討対象に用いた。ただし、表-2の19番目の出水の地上雨量データは未入手であったため、検討対象外とした。

3. ニューラルネットワークと深層学習

ANNおよび深層学習を用いた水位予測手法は、既往文献⁵⁾に詳述されている。本研究は、水位の代わりにダム流入量を予測対象としているが、基本的な計算手順は同様である。以下に計算手順の概要を記載する。

表-2 検討対象洪水のピーク水位一覧。

番号	ピーク時刻	ダム流入量 (m³/s)	レーダ雨量の 観測ピッチ	地上雨量の 観測ピッチ
1	2012/07/14 08:20	1464.94	5分	60分
2	2006/07/05 00:30	1287.38	5分	60分
3	2009/06/30 01:00	1156.18	5分	60分
4	2016/06/22 11:10	1116.77	5分	10分
5	2007/07/07 02:10	1098.16	5分	60分
6	2016/06/23 00:30	848.58	5分	10分
7	2006/07/23 21:40	816.09	5分	60分
8	2015/07/01 05:40	797.86	5分	10分
9	2007/07/04 11:30	791.48	5分	60分
10	2012/07/13 16:30	693.49	5分	60分
11	2011/06/11 08:50	674.52	5分	60分
12	2010/05/23 16:40	667.80	5分	60分
13	2013/08/25 10:40	595.83	5分	60分
14	2011/07/06 19:10	591.79	5分	60分
15	2015/06/11 13:30	543.99	5分	60分
16	2009/07/20 13:30	537.69	5分	60分
17	2011/06/20 05:00	535.93	5分	60分
18	2006/07/20 19:20	534.73	5分	60分
19	2017/07/06 04:30	532.43	5分	(未入手)
20	2006/07/20 04:20	527.00	5分	60分
21	2006/06/26 07:00	524.95	5分	60分
22	2012/06/24 20:10	513.31	5分	60分

(1) ニューラルネットワーク水位予測モデルの概要

本検討で用いたモデルは、入力層・中間層・出力層から構成される階層型ネットワークとした（図-2）。ネットワークを構成する各素子において以下のように計算を行う。

$$u = \theta_i + \sum_{i=1}^K w_i x_i \quad (1)$$

$$z = f(u) \quad (2)$$

ここで、 u は各素子の入力値、 x は入力値、 w は重み係数、 θ はバイアス、 K は各階層の構成素子数、 $f(u)$ は活性化関数、 z は素子の出力である。本稿で以下はバイアスも含めたパラメータベクトル w を改めて重み係数と呼ぶ。なお活性化関数には様々な関数が使われるが、本検討では最も適用事例の多いシグモイド関数を用いた。ネットワークの学習には確率的勾配降下法を用い、学習係数の設定にはAdaGradを適用した。重み更新の際にはモーメンタムを与えることで勾配降下法の収束性の向上を図った。誤差関数の勾配の算出には誤差逆伝搬法を用いた。学習データの関連性を十分に学習し、かつ過学習による精度低下を避けるため、ドロップアウトを適用した。ドロップアウトは、学習計算時にネットワークの素子を確率 p の割合で無効化することで、ネットワークの自由度を強制的に小さくし過学習を避ける狙いがある。

(2) 深層学習を適用した水位予測モデルの概要

深層ANNの学習時における勾配消失問題への対応として、本研究では自己符号化器による事前学習を行い、精度向上を図った。自己符号化器の模式図は図-2に示す。

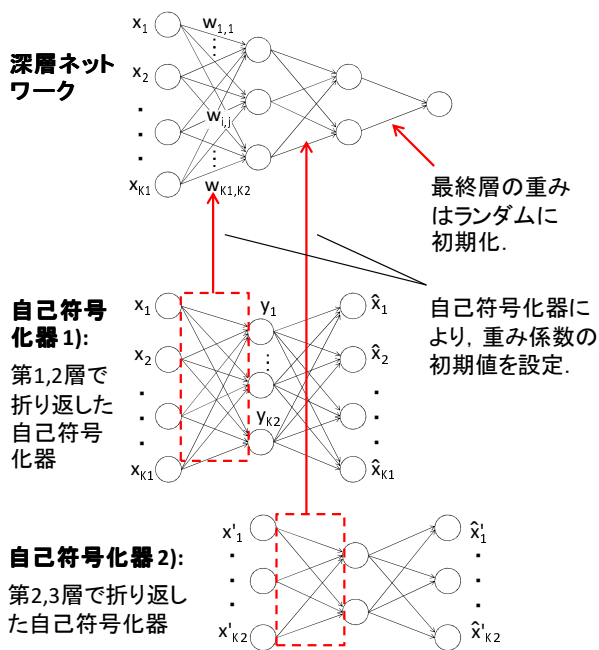


図-2 ANNの概念図。

通りであり、入力 x と出力 $\hat{x}$ ができるだけ同じになるよう、重みの調整（学習）を行う。図-2の例では、はじめに第1, 2層で折り返した自己符号化器1) を用いて学習を行い、 w を決定する。その際に得られた1, 2層間の w を深層ANNの初期値に用いる。続いて、自己符号化器1) の中間層からの出力を学習データとみなして、第2, 3層で折り返した自己符号化器2) の学習を行う。得られた自己符号化器2) の w を、深層ANNの2, 3層間の初期値に用いる。以下同様の手順を繰り返すことで、さらに階層が深いネットワークについても事前学習を行うことが可能である。事前学習を行ったあとは、従来型ANNの方法と同様に勾配降下法によって w を最適化する。以上の手順により、従来の方法に比べ効果的に学習が進むことが知られている。

4. 下笠ダム流入量予測モデルの構築

(1) ケース設定

下笠ダム流域の雨量データを入力データとして、10分～60分後までのダム流入量の変化を予測するANNモデルを構築した。モデルに入力する雨量データや学習手法の違いによる効果を検証するために、表-3に示す3つのモデルで比較検討を行った。

雨量データについては、地上雨量を用いた場合と、レーダ雨量を用いた場合とで2通りの入力データを検討した。レーダ雨量については、図-1に示すような下笠ダム流域内の11の小流域ごとに平均雨量を作成し、モデルの入力データとした。地上雨量については、4か所の地点雨量をそのまま入力データとした。なお分布型モデルや貯留関数などの流出解析モデルを用いる場合には、地

表-3 ANN流入量予測モデルの検討ケース。

No	モデル	モデルに入力する雨量データ
1	深層学習	レーダ雨量（小流域ごとの平均）
2	深層学習	地上雨量（地点雨量）
3	従来型 ANN	レーダ雨量（小流域ごとの平均）

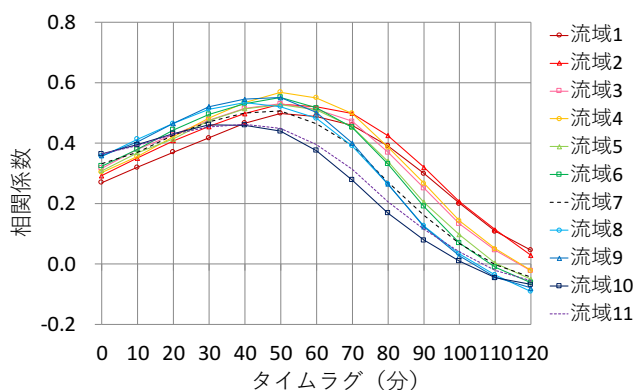


図-3 レーダ雨量（小流域ごとの平均）とダム流入量との相関。

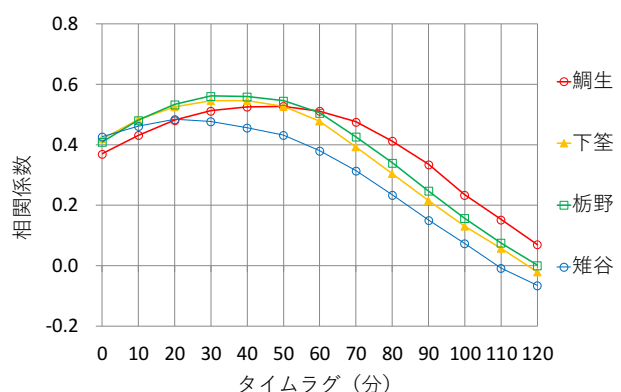


図-4 地上雨量（地点雨量）とダム流入量との相関。

点雨量をティーセン分割し、面的な降雨分布を明示的に近似することが一般的である。一方で本手法は、降雨分布の影響も含めて、降雨 - 流出の関係を過去の実績データから直接的に学習することを期待している。したがって、本研究では地上雨量に対するティーセン分割などの処理は行わない。

モデルについては、深層学習モデル、従来型ANNの2種類を検討した。深層学習モデルは、入力層・中間層2層・出力層の4層から構成され、学習時には自己符号化器による事前学習を適用した。ANN（従来型）は入力層・中間層・出力層の3層から構成され、事前学習は行わない。それぞれのモデルで、事前学習以外の学習手順や、学習・検証に用いるデータは同じとした。

(2) 相関分析による入力データの選定

モデルの入力データを絞り込むため、ダム流入量と雨量との相関分析を行った。相関分析の目的変数は、10分あたりのダム流入量変化 ΔQ とした。説明変数は、10分ごとの雨量強度 R とした。検討対象としている出水時の

2662組（地上雨量では2541組）のデータを対象に相関係数を算出した（図-3、4）。なお降雨とダム流入のタイムラグを考慮して、降雨データの時系列を0分～120分の範囲で前にずらしながら相関係数を算出した。

レーダ雨量による小流域ごとの平均雨量については、20分～70分のタイムラグで相関が高くなった。また、4か所の地上雨量については、ダム流入量変化とのタイムラグが20分～60分の範囲で相関が高くなった。これらの時間は、各地点・流域での降雨からダム流入までの流下時間とおおむね対応していると考えられる。

なお、流域の東側において、下流域よりも上流域の方がタイムラグが小さくなる逆転傾向が見られる。具体的には、図-3よりレーダ雨量における相関係数のピークを確認すると、小流域の10、11は他の小流域に比べタイムラグが小さい。同様に、図-4より地上雨量においては雉谷のタイムラグが小さい。これは、下笠ダム流域では雨域が西から東へと移動する場合が多いためと考えられる。すなわち、流域東側の上流部で雨が降り始めた際には、すでに西側の支川からダム流入が始まっているのだと推測される。こうした状況では、相関分析のみでは流域からの流下時間を適切に判断できない可能性があり、入力データの絞り込み方法としても改善の余地がある。

(3) 入力層の設定

入出力層の組合せは表-4の通りとした。ダム流入量変化の予測時間は10分～60分とし、予測時間ごとに6つのネットワークを構築した。入力データは、相関分析による洪水到達時間の検討より、水位予測に影響を及ぼしうる地点・時刻のデータをできる限り包含するように設定した。レーダ雨量を用いた場合については、各小流域でダム流入量変化との相関が高い20分～70分のタイムラグのあるデータを入力データとした。例えば現時刻から30分後までの流入量を予測するモデルでは、入力データは現時刻から40分前～10分後の雨量となる。また、雨量データのほかに、10分前～現時刻におけるダム流入量、20分前～現時刻における流入量変化を加えた。以上より、入力データ数は雨量が11流域×6データ=66、流入量データが2、流入量変化データが3であり、合計71個を入力データとした。

なお、30分～60分後の予測の際には、入力データに将来時刻の雨量データが必要となる。本研究では、まず実況雨量を予測雨量に見立てた完全予測データによりモデルの構築・精度検証を行った。その後、予測雨量を用いた精度検証により予測雨量誤差の影響を確認した。

地上雨量を用いた場合については、4地点での傾向の違いを踏まえて入力データを設定した。鯛生、下笠、桧野の3地点については20分～60分のタイムラグデータ、雉谷地点については0分～40分のタイムラグデータを入力データとした。その他はレーダ雨量を用いた場合と同様とした。以上より、入力データ数は雨量が4地点×5デ

表-4 現時刻からt分後の予測を行う場合の入出力層組合せ。

種別	入力地点	時刻	出力
ダム流入量	下笠ダム	-1, 0	
10 分間の流入量変化	下笠ダム	-2, -1, 0	下笠ダムの流入量変化
小流域ごとの平均雨量	11 流域	t-70, t-60, t-50, t-40, t-30, t-20	

表-5 学習の各種設定。

種別	設定内容
学習データ期間	過去 11 年間の上位 22 洪水
入力データ種類	ダム流入量, ダム流入量変化, 10 分雨量
データセット数	2662
バッチサイズ	100
予測時間	10 分～60 分
モーメンタム	0.5
学習率	初期値 0.1 として AdaGrad で自動設定
学習回数	10000

表-6 学習の設定に関するスタディケース。

設定パラメータ	設定値
中間素子数 (第2層-第3層)	10-5 / 20-10 / 40-20
dropout 率	0.01 / 0.02 / 0.03 / 0.04
	0.05 / 0.06 / 0.07 / 0.08
	0.09 / 0.10 / 0.15 / 0.20
	0.25 / 0.30 / 0.25 / 0.40

表-7 中間素子数, dropout率の設定。

モデル	中間素子数	dropout率
1 (深層学習、レーダ雨量)	40-20	0.02
2 (深層学習、地上雨量)	10-5	0.04
3 (従来型ANN、レーダ雨量)	10	0.10

ータ=20、流入量データが2、流入量変化データが3であり、合計25個を入力データとした。

(4) モデルの設定

学習の基本設定は表-5の通りとした。ドロップアウト率(p)、学習回数、各層の素子数については、表-6のようにスタディケースにより設定した。検証対象のうち上位6出水に対してleave-one-subject-out交差検証を行った。具体的には、22事例の検討対象洪水のうち1出水を検証データ、残り21出水を学習データとした計算を1セットとし、同様の手順を全ての検証対象洪水について行うことで精度を評価した。なお精度評価の対象は、水位が顕著に上昇し始めた時刻からピーク6時間後までの期間における流入量のRMSE（Root Mean Squared Error）とした。

以上により、最も精度の高いケースをモデルの最終設定とした。各モデルで最終的に設定した中間素子数、dropout率を表-7に示す。

5. モデルの精度検証

(1) 実況雨量による予測計算

最適化したモデルにより、モデルの精度検証を行った。なお深層ANNモデルの学習開始時には、乱数によって重み係数を初期設定しているため、データや各種設定が同じでも、学習結果（最適化された重み係数）は計算のたびに異なる。一般に、異なる初期値から学習した複数のANNを用いて、複数の予測結果を算術平均することで、予測精度が向上する場合が多いことが知られている。これはモデル平均と呼ばれ、ANNに限らず機械学習全般において有効な方法とされている¹²⁾。本検討では10個のモデルを構築し、モデル平均を最終的な予測流入量とした。なお、10分～60分後までの予測計算には、予測時間ごとに10個のモデル平均を算出するため、合計60個のモデルを用いた。

予測流入量と実績流入量とのRMSEによる精度評価の結果を図-5に示す。なお、ここでは精度評価期間は流入量ピークの前後2時間とした。また、時刻ごとの予測流入量と実績流入量のグラフを図-6に示す。

図-5より、レーダ雨量を用いた深層学習によるモデル1が最も精度が良い。続いて、レーダ雨量を用いた従来型ANNによるモデル3、地上雨量を用いた深層学習によるモデル2の順となった。60分予測の精度を見ると、モデル1の予測流入量のRMSEが70 m³/s 程度である。図-6の洪水ピーク流量1465 m³/sに対しての5%程度の誤差にとどまっており、ダム操作などへの実用に向けても十分な精度だろうと考えられる。また、洪水波形の立ち上がりやピークの最大値など、防災上の観点で重要な部分も良く再現できている。

モデル1とモデル2の比較により、深層学習を用いたモデルでは、地上雨量よりもレーダ雨量を入力データとした方が精度が高くなる。これは、レーダ雨量を用いた場合の方が、流域内の降雨分布をより詳細に反映しているため、予測精度が高くなったものと考えられる。

モデル1とモデル3の比較より、深層学習モデルの方が従来型ANNモデルよりも精度が高いことが確認できた。

モデル2とモデル3の比較より、学習モデルの違いよりも、入力データ（レーダ雨量か地上雨量か）の違いの方が、予測精度への影響が大きい結果となった。ただし、本検討では地上雨量計が4か所しか使われていないが、より多数の観測所がある流域では、レーダ雨量と地上雨量との違いは小さくなるだろうと考えられる。

また、本検討では地上雨量の元データの多くが60分間隔である。5分間隔のレーダ雨量データに比べて、地上雨量データの時間解像度は低い。より時間解像度の高い地上雨量データが揃っていれば、さらに地上雨量とレーダ雨量との予測精度の差は小さくなると考えられる。

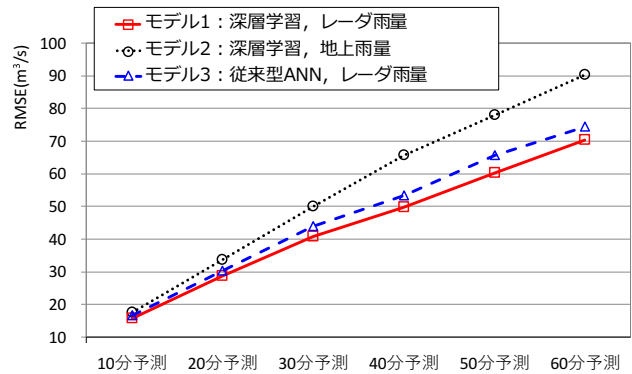


図-5 ダム流入量予測モデルの予測水位RMSE（21洪水平均）。

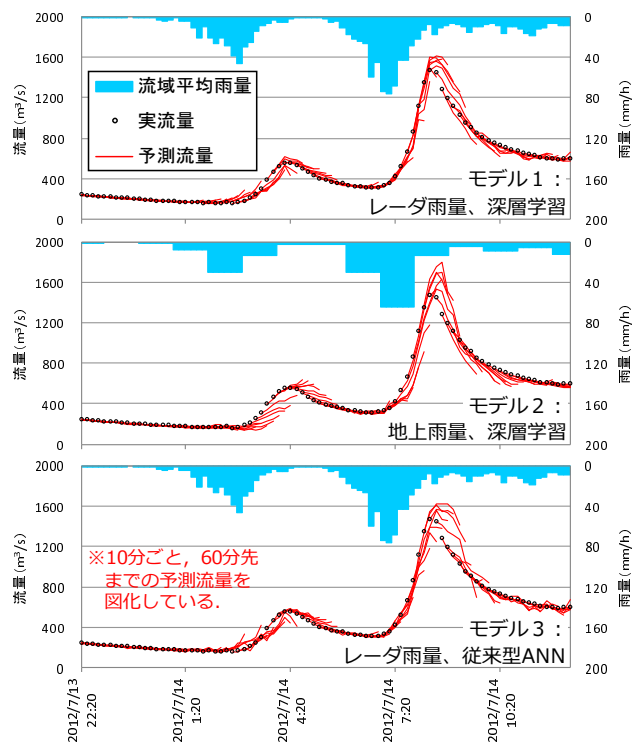


図-6 下笠ダム流入量予測（10分後～60分後）の結果。

(2) 予測雨量による予測計算

対象地点は流域面積が小さく流下時間が短いため、長い予測時間では予測雨量の精度に大きく左右される。本検討にて、予測雨量を用いた予測計算を実施し、実況雨量で計算した場合との比較を行った。なお、実況雨量は国土交通省のCバンドレーダ雨量（同時刻合成レーダ）、予測雨量は気象庁の降水ナウキャストである。図-7に、実況雨量と予測雨量との比較を示す。予測雨量のピークは実況雨量の半分程度であり、全体的にも予測雨量が大きく過小評価となっている。元となるレーダが異なるため一概に気象予測の問題とは言えないが、ここでは予測雨量誤差が大きい場合の事例と見なすことができる。流入量予測の結果を図-8に示す。10分～20分予測ではモデルの入力データに予測雨量を用いていない。したがって、実況雨量を用いた計算と、予測雨量を用いた計算と

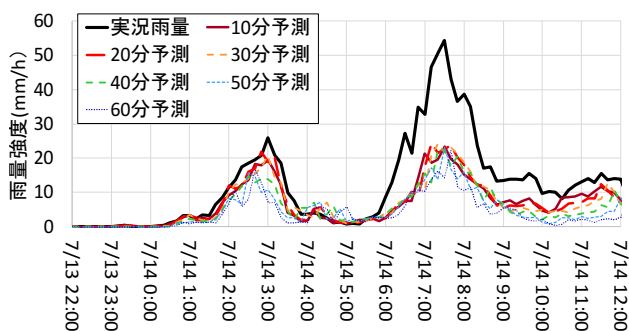


図-7 実況雨量と予測雨量の比較.

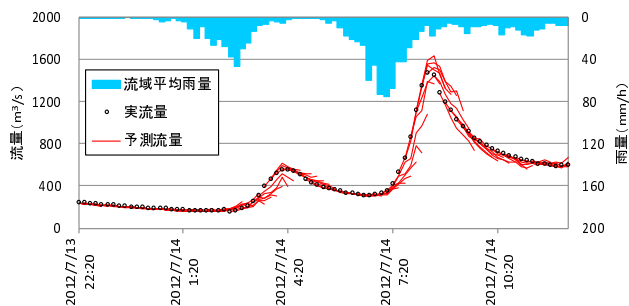


図-8 予測雨量を用いた下笠ダム流入量予測結果.

は完全に一致している。30分～60分予測については、特に出水立ち上りの部分で、予測雨量を用いた計算の方が流量予測が小さくなる傾向が見られた。出水低減の部分では実況雨量を用いた計算と差が小さい結果となった。

防災上、安全側のダム操作を考えるためには、このような予測雨量の誤差を踏まえたうえで、予測時間ごとの流入量予測精度を把握することが重要と考えられる。

なお予測計算時間は、Intel® Xeon®X5690 (6-Core, 3.46GHz) を搭載した計算機により1秒未満であった。

6. おわりに

本研究では、レーダ雨量を用いた深層学習によるダム流入量予測モデルを構築し、筑後川水系の下笠ダムを対象に適用性検証を行った。本研究の成果は次のようにまとめられる。

- レーダ雨量を用いた深層学習によるダム流入量予測モデルは、下笠ダムにおいて 60分先まで高い精度を示した。
- 深層学習によるダム流入量予測は、従来型ANNによる予測よりも高い精度となった。
- 同じ深層学習を用いた予測モデルでも、地上雨量を用いた予測よりも、レーダ雨量を用いた予測の方が高い精度となった。ただし、地上雨量計の設置数や、観測時間ピッチによって影響は異なるだろうと考えられる。
- 予測雨量を用いた検討では、20分予測までは雨量

誤差の影響は出なかった。予測時間が60分程度になると、特に出水立ち上がりで雨量誤差の影響が大きくなった。

深層学習によるダム流入量予測は十分な予測精度であり、今後はダム操作の最適化など、さらなるダム運用の効率化に向けた活用が期待できる。

参考文献

- 1) Dawson, C.W. and Wilby, R.L. : Hydrological modeling using artificial neural networks, *Progress in Physical Geography*, Vol.25, No.1, pp. 80-108, 2001.
- 2) Maier, H.R. and Dandy, G.C. : Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, *Environmental Modelling & Software*, Vol.15, pp. 101-124, 2000.
- 3) Maier, H.R., Jain, A., Dandy, G.C. and Sudheer, K.P. : Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions, *Environmental Modelling & Software*, Vol.25, pp. 891-909, 2010.
- 4) Hinton, G.E., Osindero, S. and Teh, Y.: A fast learning algorithm for deep belief nets, *Neural Computation*, Vol.18, pp. 1527-1544, 2006.
- 5) 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.187-192, 2016.
- 6) 一言正之, 桜庭雅明: 深層学習の適用によるニューラルネットワーク洪水予測の精度向上, 河川技術論文集, Vol. 22, pp.1-6, 2016.
- 7) 一言正之, 桜庭雅明: 多地点観測情報を活用した深層ニューラルネットワークによる河川水位予測の精度向上, 河川技術論文集, Vol. 23, pp. 287-292, 2017.
- 8) Raman, H. and Sunilkumar, N.: Multivariate modelling of water resources time series using artificial neural networks, *Hydrological Sciences Journal*, Vol.40, No. 2, pp. 145-163, 1995.
- 9) 飯坂達也, 松井哲郎, 植木芳照: ニューロ・ファジーによるダム流入量予測システムの開発, 電気学会論文誌B, Vol. 119, No. 10, pp. 1020-1025, 1999.
- 10) Valipour, M., Banihabib, M. E., and Behbahani, S. M. R.: Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir, *Journal of hydrology*, Vol.476, pp.433-441, 2013.
- 11) Bai, Y., Chen, Z., Xie, J. and Li, C. : Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models, *Journal of hydrology*, 532, 193-206, 2016.
- 12) 岡谷貴之: 深層学習, 講談社サイエンティフィック, 2015.

(2018. 4. 3受付)