

AIを用いたダム流域の洪水予測

やまわき まさし みうら しん かねこ ひろふみ ふじつか しんたろう つじくら ひろき
山脇 正嗣¹・三浦 心²・金子 拓史¹・藤塚 慎太郎²・辻倉 裕喜³

¹(株)建設技術研究所 技術本部 (〒103-8430 東京都中央区日本橋浜町3-21-1)

²(株)建設技術研究所 東京本社 水システム部 (〒103-8430 東京都中央区日本橋浜町3-21-1)

³(株)建設技術研究所 大阪本社 水システム部 (〒541-0045 大阪府大阪市中央区道修町1-6-7)

洪水時のダム運用を的確かつ確実に実施するためには、ダム流入量を高精度で予測することが重要である。ダム管理者は貯留関数法等の洪水予測モデルを参考にしつつ、放流計画を立案しており、洪水予測モデルの予測精度の向上が課題となっている。

今回、洪水時の流入量及びダム下流河川水位の予測に、近年さまざまな分野で脚光を浴びている深層学習(ディープラーニング)を用いた洪水予測モデルの適用性について検討した。検討に当たっては、予測手法やハイパーパラメータ等を複数設定し、予測誤差が少ない手法を模索した。その結果、予測雨量が正確なものが得られれば、48時間先までのダム流入量とダム下流地点の河川水位を実務に適用できるレベルまで高精度に予測できる結果を得た。

Key Words : 洪水予測, 深層学習(ディープラーニング), ダム流入量, 河川水位

1. はじめに

昨今、IoTやAI(人工知能)のインフラ分野への活用が急速に進んでいる¹⁾。ダム・水路の水資源分野においても、AI技術の活用により、維持管理の高度化・効率化に向けたさまざまな側面からの検討²⁾がなされており、特にダム流入量や下流河川の水位予測精度の向上が望まれている。

そこで本研究では、深層学習の一種であるディープニューラルネットワーク(Deep Neural Network(以下DNN))を活用し、洪水時のダム流入量と下流河川水位の予測精度向上を目的とした検討を実施した。

2. 既往研究事例と本検討の方向性

(1) 国内・国外における既往研究事例

DNNとは、入力層・中間層・出力層の3層で構成される浅層型ニューラルネットワーク(Artificial Neural Network(以下ANN))の中間層を多層化した深層型モデルである。ANNよりも入力データの詳細な特徴抽出を行い、予測精度が向上する学習法であり、近年研究開発が急速に進められている。

流入量予測に係る既往研究では、従来式のANNを活用した研究³⁾⁴⁾が主となっているが、深層型ANN等のDNNの活用研究事例⁵⁾⁶⁾についてもいくつか報告さ

れている。特に、一言ら⁵⁾はDNNを基本とし、入力層に分布型流出解析モデルの計算結果を加えることで両モデルを融合し、元のモデル(深層ANN、分布型モデル)の予測精度を上回る結果を得ている。

一方で、これらの多くの既往研究では、入力値に実測値との誤差の大きい予測値(雨量等)を用いず、過去から現在までの実測値のみを用いて、1~6時間程度の短期間予測を行うことを目的としている。

(2) 本研究の特徴

本研究は、出水対応等における準備段階を含め洪水調節等の実務的なダム管理に必要な48時間先までの長期予測を実施することを目的としたものである。

本研究では、まずモデルダム(後述)における水文データを収集し、目的変数であるダム流入量及びダム下流河川水位と相関の高い説明変数を抽出した。次に、「48時間分の予測計算を一度に実施するモデル①」と、「1時間毎に計算を実施して個別の48時間分実施するモデル②」の2種のDNNモデルを構築し、各モデルに最適なハイパーパラメータ(ユーザーが定義するパラメータ)の設定を行った。

最後に、構築したDNNモデルについて交差検証によるモデルの安定性評価及び予測精度評価を行い、DNNモデルによる洪水予測への適用性を確認した。

なお、現実には雨量の予測値と実績値には誤差があるが、本研究ではこの誤差を考慮せず実績雨量を予測雨量として使用し、DNNモデルの評価を行った。

3. 対象流域の概要と水文データの整理

(1) モデルダムと予測項目

本研究のモデルダムとして、水機構が管理するダムの中からAIの学習に必要な水文資料が充実している一庫ダム(図-1)を選定した。予測項目は①ダム流入量、②ダム下流10kmに位置する水位観測局水位(以下、下流水位)とした。下流水位地点は水系の中流部、ダム下流5kmのところでは本川流域が合流し、更に下流5kmに位置する地点である。下流水位地点は岩河床であり、河床変動はほとんどみられない。

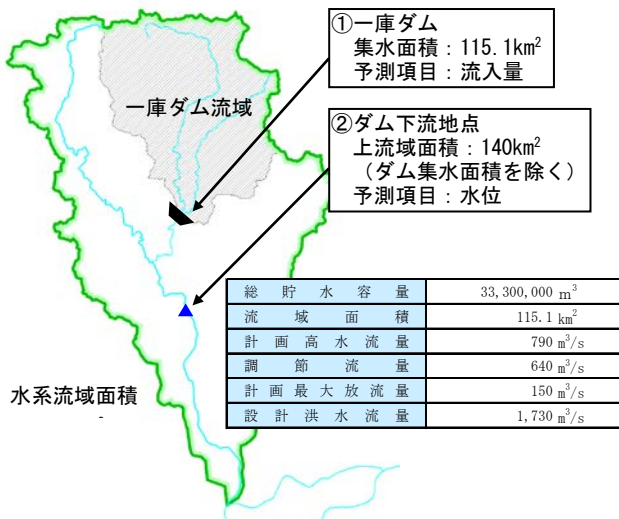


図-1 一庫ダム流域図とダム諸元

(2) 学習データとして使用した水文データ

a) 一庫ダムにおける出水データ

一庫ダム運用開始以降(1983年~2017年)の河川水位、雨量、ダム諸量データを収集し、過去に洪水調節を実施した15洪水に、流域平均の総雨量が30mmを超える25洪水を追加した合計40洪水をDNNモデルの学習データとして使用した。

b) 入力要素の相関分析

DNNモデルの構築に適した説明変数を選定するため、一庫ダムにおけるダム流入量、下流水位を目的変数とした各種水文データとの相関分析(図-2)を行い、流出特性を分析した。その結果、いずれも流域平均雨量及び流域平均累加雨量との相関が高い結果となった。これより、上記の40洪水データの中から後述の説明変数を選定した。

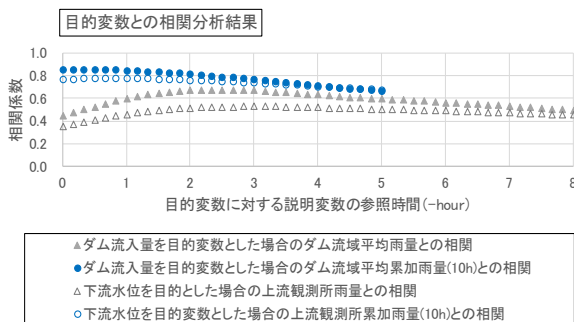


図-2 目的変数との相関分析

4. 深層学習モデルによる長時間予測手法の構築

(1) DNNモデルの検討

前述の学習データを用いて、DNNによるダム流入量、下流水位の48時間後予測モデルについて検討した。DNNモデルの予測精度は、入力に用いる説明変数やハイパーパラメータ(中間層の層数、活性化関数、損失関数、最適化関数等の人力で設定するパラメータ⁷⁾)に左右される。そのため、説明変数やハイパーパラメータの組合せ感度分析を行い、最も精度の良い組合せを選定した。

モデルの評価指標としては、図-3に示すとおり、洪水予測分野の代表的な精度評価指標である①RMSE(2乗平均平方根誤差)、②Nash-Sutcliffe係数、③流出ボリューム誤差(※ダム流入量のみ)、④ピーク流入量・水位誤差を使用した。

$$RMSE(y_k, t_k) = \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - t_k)^2} \dots (1a)$$

$$NS = 1 - \frac{\sum_{i=1}^N \{q_o(i) - q_c(i)\}^2}{\sum_{i=1}^N \{q_o(i) - q_{av}\}^2} \dots (1b)$$

N : 計算時間数
 q_{av} : 実測値の平均値
 $q_o(i)$: i 時の実測値
 $q_c(i)$: i 時の計算値
 ※上りの精度を評価するため、ピーク前10時間の期間で評価

y : 予測値 t : 実績値 n : データ数

① RMSE(2乗平均平方根誤差)

$$E_v = \frac{\sum_{k=1}^n t_k - \sum_{k=1}^n y_k}{\sum_{k=1}^n t_k} \dots (1c)$$

y : 予測値 t : 実績値 n : データ数

② Nash-Sutcliffe係数

$$J_{pe} = \frac{|t_p - y_p|}{t_p} \dots (1d)$$

t_p : 実績ピーク値 y_p : 予測ピーク値

③ 流出ボリューム誤差

④ ピーク流入量・水位相対誤差

図-3 DNNによる洪水予測モデルの評価指標

また、学習の際は40洪水のうち39洪水を訓練データとしてモデルを学習させ、テストデータの1洪水を対象に、ダム流入量と下流水位の予測精度検証を行った。なお、テストデータは以下の観点より2017年10月洪水を選定している。

【テストデータの選定理由】

- 対象洪水の内、最大ピーク流量と最小ピーク流量の中間程度のピーク流量を持つ
- データの期間や精度を考慮し直近に生起
- 降雨及び洪水波形が1山(精度確認のため、2山洪水など複雑でない洪水を対象)

(2) DNNモデルの構築

48時間後までの予測結果として、次頁の図-4に示すとおり、1時間毎に48時間先まで予測するモデル①、現時刻から1時間後、計算結果により2時間後・・・を繰り返し48時間先まで予測するモデル②の2種類の深層型ANNモデル(DNNモデル)を構築した。

また、入力値である説明変数は、相関分析結果において相関の高かった次頁の表-1の項目を設定した。これらの説明変数を用いて48時間後までの予測値(目的変数)を算出し、損失関数(モデルの性能評価関数)による予測値との差異の算出、最適化関数(モデルの更新関数)によるパラメータ更新を行うモデルを構築した。

表-1 モデル入力値用の説明変数

| ダム流入量予測 | 下流水位予測 |
|---|---|
| 【説明変数】 ① ダム流入量 (実績値) ② 流域平均雨量 (実績値・予測値) ③ 流域平均累加雨量 (実績値・予測値) 【時間設定】 ①②: 0~6時間前 ③: 48時間後 (モデル①), 1時間後 (モデル②) | 【説明変数】 ① 下流水位 (実績値) ② ダム放流量 (実績値) ③ 流域平均雨量 (実績値・予測値, ダム流域を除く) ④ 流域平均累加雨量 (実績値・予測値, ダム流域を除く) 【時間設定】 ①②: 0~6時間前 ③: 48時間後 (モデル①), 1時間後 (モデル②) |
| モデル①: ダム流入量・下流水位の48時間予測を一度に行う | モデル②: ダム流入量・下流水位の1時間予測を48時間分行う |

(3) ハイパーパラメータの設定

入出力データ及びハイパーパラメータの540ケース (入出力データ: 18ケース×ハイパーパラメータ: 30ケース) の組合せの中から, 訓練データによる学習とテストデータによる予測精度検証を行い, 最も精度の高いモデル構造を暫定的に決定した。

ここで, 表-2の各モデルにおいて最も予測精度が高くなるハイパーパラメータの組合せによる予測結果を図-5に示す。結果として, ダム流入量, 下流水位の両方においてテストデータの実測値に近接した48時間予測が実施できるモデル構造が得られた。

【各ハイパーパラメータが予測精度向上に与えた影響】

- ・ノード数が少ないと精度が低くなるが, 多すぎても向上しない。
- ・層数を多層化しても精度差はほとんど生じない。
- ・活性化関数はダム流入量ではRelu, 下流水位ではtanhが有効。
- ・ミニバッチサイズは少ない場合より多い方が予測精度が高くなる。
- ・学習回数は100回の精度が最も高くなった。
- ・最適化関数はAdam, RMSprop, SGDとも精度は同程度である。
- ・学習係数は0.001のものが最適となっている。
- ・データ正規化方法は標準化が最も優れている。
- ・Batch Normalization (過学習を防ぐテクニック) を実装した方が精度は高い。
- ・ドロップアウト率が低い方が精度が高い。
- ・従来型のANNと比較して, DNNの方が精度が高い。

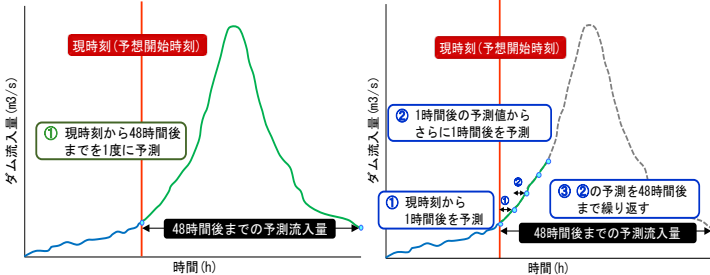


図-4 DNNによる2種類の洪水予測モデル

表-2 解析モデル①② (ダム流入量及び下流水位) において最も予測精度が高かったハイパーパラメータの組合せ

| 項目 | ハイパーパラメータ | 検討パターン | ダム流入量 | | | | 下流水位 | | | |
|------------------|---------------------------|--|-----------|------------------|------------------|------------------|------------------|-----------------|-----------------|-----------------|
| | | | モデル① | | モデル② | | モデル① | | モデル② | |
| | | | DNN | ANN | DNN | ANN | DNN | ANN | DNN | ANN |
| 中間層 | 重み初期値 | ①Heの初期値 (活性化関数がReluの場合) ②Xavierの初期値 (活性化関数がSigmoid, tanhの場合) | Heの初期値 | | Heの初期値 | | Xavierの初期値 | | Xavierの初期値 | |
| | 層数 | ①2層, ②3層, ③1層 (ANN) | 2層 | 1層 | 2層 | 1層 | 2層 | 1層 | 2層 | 1層 |
| | ノード数 | ①50, ②500, ③2500 | 500 | | 500 | | 500 | | 50 | |
| | 活性化関数 | ①Relu関数, ②Sigmoid関数, ③tanh関数 | Relu関数 | | Relu関数 | | tanh関数 | | Sigmoid関数 | |
| パラメータの更新方法 | 最適化関数 | ①Adam関数, ②SGD関数, ③RMSprop関数 | Adam関数 | | Adam関数 | | Adam関数 | | Adam関数 | |
| | 最適化関数の学習係数 | ①0.1, ②0.01, ③0.001, ④0.0001, ⑤0.00001 | 0.001 | | 0.001 | | 0.001 | | 0.001 | |
| | ミニバッチサイズ | ①10, ②30, ③100, ④300, ⑤1000 | 100 | | 100 | | 100 | | 100 | |
| その他 | 学習回数 (エポック数) | ①10, ②100, ③1000 | 100 | | 1000 | | 100 | | 100 | |
| | データ正規化 | 入力層 ①標準化, ②正規化, ③正規化処理なし 中間層 ①Batch Normalization, ②正規化処理なし | 標準化 | | 標準化 | | 標準化 | | 標準化 | |
| | ドロップアウト率 | ①0.3, ②0.5, ③0.8 | 0.3 | | 0.3 | | 0.3 | | 0.3 | |
| | テストデータに対する解析結果 (各ケース5回実施) | RMSE | 平均 (標準偏差) | 10.0319 (0.4115) | 14.2247 (0.5543) | 14.2351 (1.6951) | 16.9023 (2.1419) | 0.1652 (0.0160) | 0.1759 (0.0057) | 0.2587 (0.0534) |
| Nash-Sutcliffe係数 | | 0.858 (0.0651) | | 0.632 (0.0357) | 0.874 (0.1320) | 0.589 (0.1971) | 0.896 (0.0450) | 0.902 (0.0299) | 0.910 (0.0300) | 0.733 (0.3268) |
| ボリューム誤差 | | 0.0987 (0.0089) | | 0.1351 (0.0131) | -0.0323 (0.0844) | 0.2624 (0.0481) | - | - | - | - |
| ピーク流入量 | | 0.0651 (0.0223) | | 0.1310 (0.0267) | 0.1071 (0.0511) | 0.1251 (0.0326) | 0.0310 (0.0171) | 0.0389 (0.0169) | 0.0218 (0.0198) | 0.0229 (0.0214) |
| 水位相対誤差 | | - | | - | - | - | - | - | - | - |

※各時刻における48時間予測の平均RMSEで評価 ※重み初期値がランダムで, 予測計算の差異が生じるため5回の平均値とした

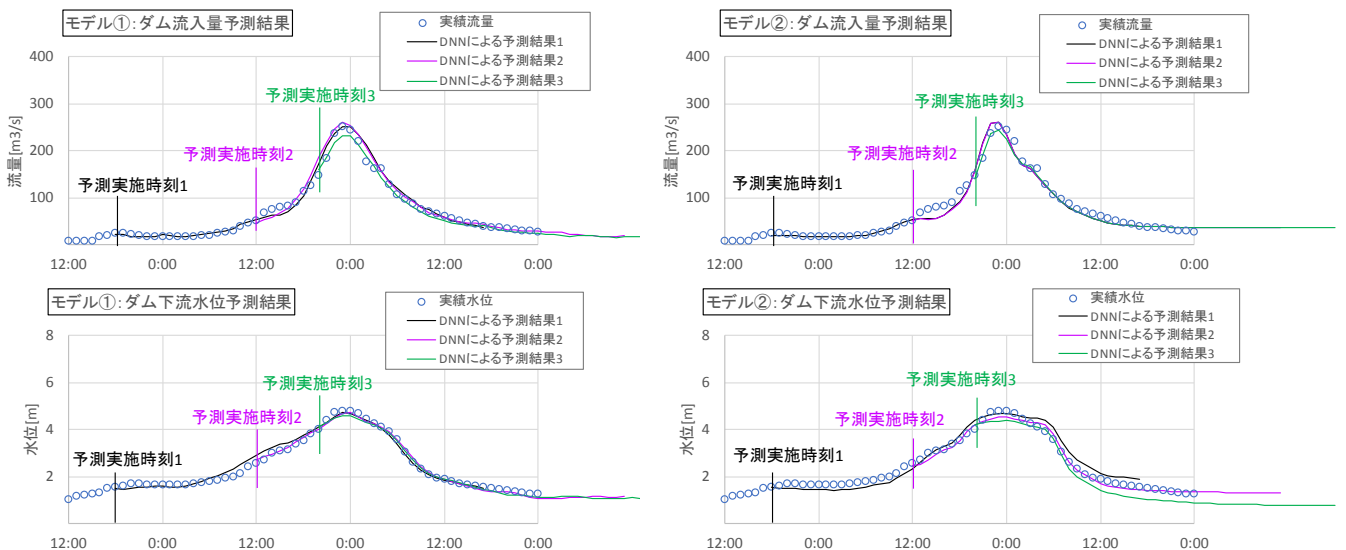


図-5 解析モデル①②のテストデータ (2017年10月洪水) に対する予測結果 (任意の3点からの48時間後予測結果)

5. 深層学習モデルの評価

(1) 交差検証による解析モデルの安定性評価

構築したDNNモデルについて、テストデータ以外の39洪水についても同程度の精度予測が可能か検証するため、実績データ(40洪水)を9:1の割合で訓練データとテストデータに分類し、全ての組み合わせ(10ケース)を対象に安定性評価を実施した。

その結果、RMSEを評価指標とした場合、モデル①は2017年10月洪水のテストデータに比べて2割程度大きいRMSEが得られた。また、モデル②はテストデータに比べて約6割RMSEが増加した。この結果から、モデル①の方が安定性が高いといえる。

表-3 解析モデルの安定性評価結果の概要

| 予測項目 | モデル種別 | 安定性評価結果(各ケース5回実施) | | |
|---------------------------|-------|-------------------|-----------------|----------------|
| | | RMSE | ボリューム誤差 | ピーク流入量・水位相対誤差 |
| 平均値(標準偏差) | | | | |
| ダム流入量 (m ³ /s) | モデル① | 11.85 (0.44) | -0.15 (0.05) | 0.24 (0.05) |
| | モデル② | 22.07 (4.59) | -1.19 (0.53) | 0.67 (0.23) |
| 下流水位 (m) | モデル① | 0.210 (0.013) | — | 0.11 (0.02) |
| | モデル② | 0.343 (0.037) | — | 0.12 (0.02) |

※ 重み初期値がランダムで予測計算の差異が生じるため5回の平均値とした

(2) 解析モデルの予測精度検証

構築したDNNモデルについて、モデル①と②の予測精度を比較検討した。予測精度検証の対象洪水は、一庫ダムにおける現行の操作方式(150m³/s一定量)となる2000年以降の洪水の中から、最大流入量の上位5洪水(表-4)を選定した。

モデル①と②の予測結果(48時間先まで)に対して、同期間の実績値と比較し、各評価指標を算出した(表-5)。なお各指標は、予測結果の期間にピークを含むよう、ピークの40時間程度前から7時間毎に5時点の予測結果を対象に算出したものの平均値である。

これより、ダム流入量についてはRMSE及びNash-Sutcliffe係数、ボリューム誤差で評価した場合はモデル①の精度が高くなり、ピーク誤差で評価した場合はモデル②の精度が高くなった。

表-4 モデル①と②の予測精度比較のための対象洪水

| No | ピーク生起日 | ピーク流入量[m ³ /s] | 備考 |
|----|------------|---------------------------|------|
| 1 | 2004.10.20 | 411 | |
| 2 | 2013.9.16 | 468 | |
| 3 | 2014.8.10 | 440 | 2山波形 |
| 4 | 2014.8.16 | 338 | |
| 5 | 2015.7.18 | 313 | |

6. 結論

一庫ダムを対象に、深層学習の一種であるDNNモデルを活用したダム流入量及び下流水位予測モデルを構築した。ハイパーパラメータを適切に設定する

ことにより、48時間後までの予測について、予測雨量として実績雨量を与えた(予測雨量の誤差なし)場合には実務的に適用できるレベルの良好な予測精度が得られた。ただし、実際には予測雨量は誤差を含むため、それらの評価、取扱が今後の課題である。

表-5 モデル①と②の予測精度の比較結果

| 洪水 | ダム流入量 | | 下流水位 | |
|------------|--------------------------|------|--------------------------|------|
| | RMSE [m ³ /s] | | RMSE [m ³ /s] | |
| | モデル① | モデル② | モデル① | モデル② |
| 2004.10.20 | 20.6 | 27.6 | 0.36 | 0.88 |
| 2013.9.16 | 17.0 | 29.7 | 0.19 | 0.98 |
| 2014.8.10 | 28.7 | 41.9 | 0.38 | 1.03 |
| 2014.8.16 | 31.7 | 28.7 | 0.26 | 0.70 |
| 2015.7.18 | 17.8 | 27.6 | 0.18 | 0.82 |
| 平均 | 23.1 | 31.1 | 0.27 | 0.88 |

※値が小さいほど精度が高い

| 洪水 | ピーク誤差 [m ³ /s] | | ピーク誤差 [m] | |
|------------|---------------------------|------|-----------|------|
| | モデル① | モデル② | モデル① | モデル② |
| 2004.10.20 | 89.1 | 28.0 | 1.18 | 2.58 |
| 2013.9.16 | 31.5 | 29.5 | 0.19 | 1.63 |
| 2014.8.10 | 107.8 | 38.5 | 1.61 | 2.82 |
| 2014.8.16 | 82.8 | 24.0 | 0.72 | 0.60 |
| 2015.7.18 | 31.7 | 21.8 | 0.13 | 0.62 |
| 平均 | 68.6 | 28.4 | 0.77 | 1.65 |

※値が小さいほど精度が高い

| 洪水 | Nash-Sutcliffe係数 [-] | | Nash-Sutcliffe係数 [-] | |
|------------|----------------------|------|----------------------|------|
| | モデル① | モデル② | モデル① | モデル② |
| 2004.10.20 | 0.75 | 0.64 | 0.81 | 0.61 |
| 2013.9.16 | 0.88 | 0.77 | 0.92 | 0.71 |
| 2014.8.10 | 0.68 | 0.68 | 0.77 | 0.53 |
| 2014.8.16 | 0.48 | 0.44 | 0.76 | 0.54 |
| 2015.7.18 | 0.86 | 0.73 | 0.96 | 0.90 |
| 平均 | 0.73 | 0.65 | 0.84 | 0.66 |

※値が1に近いほど精度が高い

| 洪水 | ボリューム誤差 [×10 ³ m ³] | |
|------------|--|-------|
| | モデル① | モデル② |
| 2004.10.20 | 435 | 1,744 |
| 2013.9.16 | 601 | 1,068 |
| 2014.8.10 | 821 | 1,363 |
| 2014.8.16 | 1,049 | 911 |
| 2015.7.18 | 1,690 | 1,364 |
| 平均 | 919 | 1,290 |

※値が小さいほど精度が高い

謝辞：本研究を実施するにあたり論文を参考とし、関係者の方々にもご協力を戴いた。感謝の意を表す。

参考文献

- 1) 小国土交通省：経済産業省提出資料，首相官邸HP，政策会議構造改革徹底推進会合インフラ，ICT活用等による生産性と機能の向上について，資料5，第2回，2018.
- 2) 小西慎治，寺島善宏：都市トンネルの最新建設技術及び維持管理技術，地盤工学会誌，第66巻，第2号，pp.10-13，2018.
- 3) 関基，木村晃，古山一志，和田高宏，金子祐，穴水秀樹，眞間修一：河川特性を反映したニューラルネットワーク洪水予測の精度向上，土木学会 河川技術論文集，第19巻，2013.
- 4) 竹村仁志，川村一人：猪名川におけるニューラルネットワークモデルを用いた水位上昇部の予測精度向上，国土交通省近畿地方整備局 新技術・新工法部門，No.16，pp.1-5，2016.
- 5) 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.
- 6) Haytham Assem, Salem Ghariba: Urban Water Flow and Water Level Prediction based on Deep Learning, ECML PKDD 2017: Machine Learning and Knowledge Discovery in Databases, pp.317-329.
- 7) 斉藤 康毅:ゼロから作るDeep Learning, O'REILLY, 2017.