

都市中小河川実流域データセットを用いた 深層学習モデルによる 降雨流出ベンチマークテスト

藤塚 慎太郎¹・河村 明²・天口 英雄³・高崎 忠勝⁴

¹正会員 東京都 建設局 江東治水事務所 (〒124-0023 東京都葛飾区東新小岩一丁目 14-11)

E-mail: shintarou_fujizuka@member.metro.tokyo.jp (Corresponding Author)

²正会員 工博 東京都立大学教授 都市環境科学研究科 (〒192-0397 東京都八王子市南大沢一丁目 1)

E-mail: kawamura@tmu.ac.jp

³正会員 工博 東京都立大学助教 都市環境科学研究科 (〒192-0397 東京都八王子市南大沢一丁目 1)

E-mail: amaguchi@tmu.ac.jp

⁴正会員 東京都 土木技術支援・人材育成センター (〒136-0075 東京都江東区新砂一丁目 9-15)

E-mail: takasaki@doboku.metro.tokyo.jp

都市型水害が頻発しており、複雑な流出機構である都市中小河川流域において、簡易にモデルの構築が可能な機械学習モデルの洪水予測分野への適用に対する期待も大きい。そこで本論文では、大河川流域とは異なる 1 分ごとの短い観測間隔で作成された都市中小河川実流域データセットを用いて深層学習モデルを構築し、学習洪水数や中間層ノード数などのハイパーパラメータを変化させた場合のベンチマークテストを実施した。深層学習モデルと比較するため ANN モデルも構築し、深層学習モデルパラメータ数と観測データ数を PD 比という新たな指標で性能を評価した。その結果、同じ程度の PD 比では学習・検証洪水に対して ANN モデルよりも深層学習モデルの方が優位であり、特に ANN モデルでは検証洪水において実績に沿わない急激な変動がみられることが分かった。

Key Words: deep learning, Artificial Neural Network, urban runoff, benchmark data, Hyperparameter

1. はじめに

近年、都市型水害が喫緊の社会問題の一つとなっている。都市流域では不浸透域が多く存在するとともに、雨水排水施設などの流出経路や流出抑制施設や治水施設の存在により、非常に複雑な流出機構となっている¹⁾。また、都市域では降雨から流出までの時間が短く、事前に精度良く水位・流量を予測する手法の開発が重要であり、流域ごとの複雑な流出機構を簡易に精度良くモデル化できる都市流出予測モデルの開発が望まれている。

近年発展が目覚ましい人工知能 (Artificial Intelligence : AI) 技術²⁾を活用した機械学習は、学習に用いたデータに合うように自動で学習し、モデルが構築されるため、簡易にモデルを構築することが可能である。機械学習を洪水予測分野において活用すべく、ニューラルネットワークや深層学習を用いた予測モデルの構築が行われてきた^{3)~5)}。その多くは大河川流域を対象としており、1 時

間ごとの観測データを入力とし、モデルが構築されている場合が多く、また山地中小河川流域を対象とした深層学習の研究⁶⁾では、10 分データを用いてモデルが構築されている事例がある。一方、東京都などの都市中小河川では 1 分ごとと大河川流域よりも短い間隔で水位・雨量の観測が実施されている⁷⁾。特に 1 分ごとの雨量観測値についてはハイトグラフが鋸歯状になり、大河川流域の 10 分や 1 時間ごとのデータと傾向が異なる。また、ハイドログラフも同様に観測誤差と考えられる細かい変動が生じる。このような都市中小河川での 1 分ごとの観測値を活用した深層学習の事例はあまり見受けられない。

次に、著者らは都市流域を対象流域とした 1 分ごとの仮想の降雨 (バーチャルハイトグラフ : VH) と流出高 (バーチャルハイドログラフ : VQ) を作成し⁸⁾、それらを真値として深層学習モデルで学習することで誤差を排除した深層学習モデルの特性について研究している⁹⁾。また、著者ら¹⁰⁾は AI 降雨流出ベンチマークテスト

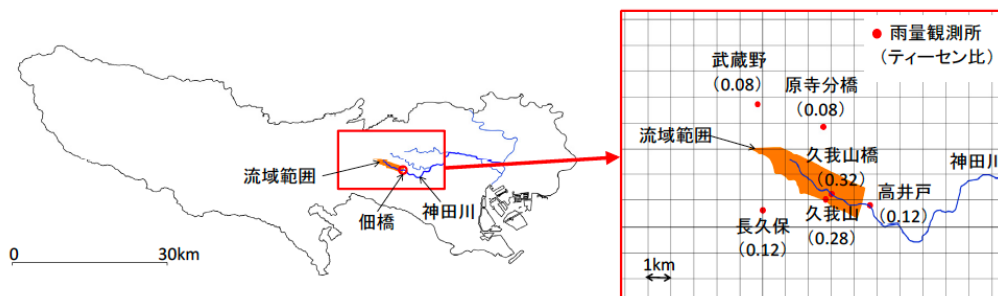


図-1 対象流域および水位・雨量観測所

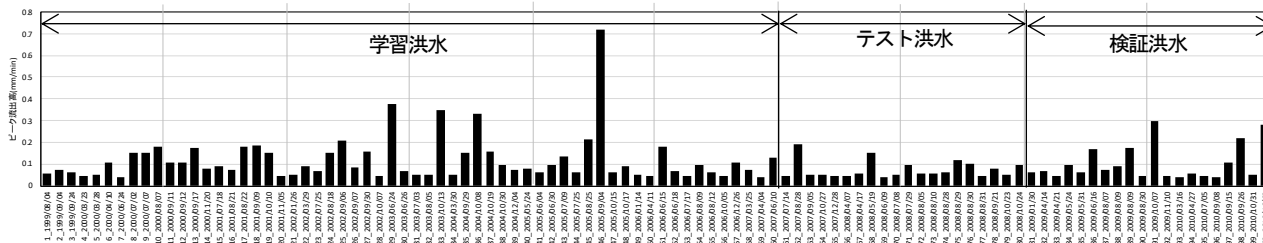


図-2 対象洪水のピーク流出高

に向けた都市中小河川実流域データセットを作成し、公開している。都市中小河川実流域データセットは東京都水防災情報システムによる1分ごとの水位・雨量観測値を基にベンチマークテストに活用しやすいように流域平均雨量や水位・流量換算式により変換された流量や流出高を付加し、各洪水のピーク流出高を基準に上位100洪水を抽出したデータである。

そこで、本論文では都市中小河川実流域データセットを活用し、流出機構が複雑な都市中小河川の実流域を対象とし、まずは流出解析モデルと同様の条件（入力層：1分ごとの雨量，出力層：1分ごとの流出高）の深層学習モデルを構築した。そして、観測データ数（学習洪水数）やモデルパラメータ数（入力層ノード数・中間層ノード数を変化させた場合）などのハイパーパラメータ（機械学習のパラメータのうち、人間が手で設定する必要のあるもの）を増減させた場合の深層学習モデルのベンチマークテストを実施した。これらのハイパーパラメータは他の流域においても同様の条件により検討を行うことが可能であり、本論文の検討結果はハイパーパラメータの最適値を探索する上でのノウハウとなると考えられる。深層学習モデルの性能は観測データ数とモデルの複雑さの関係を評価することを目的とし、モデルパラメータ数 N_p と学習に用いた観測データ数 N_b の比（以下、PD比と呼ぶ）という新たな指標で評価する。さらに中間層が1層の人工ニューラルネットワーク（以下、ANNと呼ぶ）モデルも構築し、比較を実施した。

2. 実流域データセット

深層学習モデルの学習データとして著者ら¹¹⁾が作成したAI降雨流出ベンチマークテストに向けた都市中小河

川実流域データセットを用いる。都市中小河川実流域データセットは図-1に示す神田川流域の佃橋地点（流域面積5.2km²）の水位と周辺の6地点の地上雨量について、1999年4月～2010年12月を対象、洪水イベントを抽出し、各イベントのピーク流出高を基準に、100洪水整理したデータである。洪水イベントの抽出においては、6地点の地上雨量観測所すべてが3時間連続で無降雨となった場合を区切りとしている。なお、高崎ら¹¹⁾では6時間無降雨の場合を基準としているが、流域の規模が小さく流下するまでの時間が短いことを考慮し、3時間に修正したものを使用する。データセットには佃橋地点の流量（HQ式を用いて換算）と流出高も付加されており、6観測所ごとの地上雨量観測値をティーセン分割（図-1にティーセン係数を示す）することにより求めた流域平均雨量も付加されている。水位は式(1)に示すHQ式および式(2)より流出高に変換されたものを用いる。なお、HQ式は高崎ら¹¹⁾に記載されているものから、最低水位を見直し、修正したものを用いる。

$$Q = 2.6(H - 37.26)^{2.455} \quad (1)$$

$$q = \frac{Q * 60}{A * 1000} \quad (2)$$

ここに、 q ：流出高(mm/min)、 Q ：流量(m³/s)、 A ：流域面積5.2(km²)、 H ：水位(m)である。

神田川流域は早くから都市整備が進められており、2003年時点での市街化率は97%¹²⁾となっており、近年の大きな市街化率の変化は見られない。また、下水道は合流式下水道が普及しており、高水時の流出は下水道の影響を受けたものである。

データセットは時系列順に100洪水整理されており、これをQ(1)～Q(100)と呼ぶこととする。図-2に抽出した100イベントのピーク流出高を示す。ピーク流出高

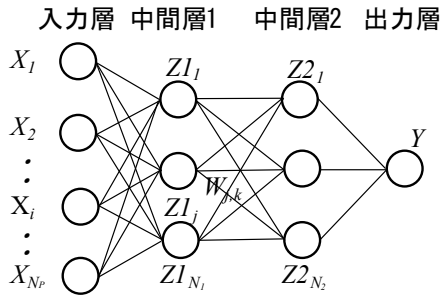


図-3 深層学習モデル構造

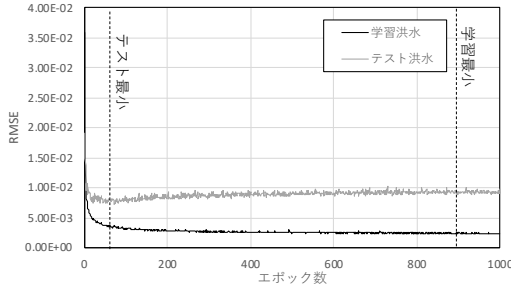


図-4 エポック数と RMSE の関係

0.2(mm/min)以下の洪水が多数を占めており、No.46 : 2005年9月4日洪水が最大規模でピーク流出高0.7(mm/min)程度であった。

3. 深層学習モデルの構築

(1) 適用した深層学習モデル

VH・VQを用いた藤塚ら^{8),9)}において中間層が2層の深層学習モデルでハイドログラフの再現が可能であったので、それを踏襲し、本論文でも図-3に示すその構造が最も単純な中間層が2層の深層学習モデル¹³⁾をまずは用いることとした。さらに、中間層が1層のANNモデルも構築し、深層学習モデルと比較した。

各ノードには重み係数とバイアスの2種類のパラメータが存在し、学習によりパラメータが更新され、精度が向上していく。各ノードの出力は活性化関数により出力値に変換され、次のノードに伝わる。一般に中間層の層数がLの深層学習モデルパラメータ数は次式で表される。

$$N_p = (N_a + 1)N_1 + \sum_{j=1}^{L-1} (N_j + 1)N_{j+1} + (N_L + 1)N_o \quad (3)$$

ここに、 N_p : 深層学習モデルパラメータ数、 N_a : 入力層ノード数、 N_j : 中間層jのノード数、 N_o : 出力層ノード数、 L : 中間層の層数である。

学習洪水数はQ(1)からの洪水数を示し、例えば学習洪水数が1の場合はQ(1)のみを、学習洪水数が60の場合はQ(1)~Q(60)までの60洪水を学習データとして採用することになる。深層学習モデルの入力層、出力層の設定は、VH・VQを用いた藤塚ら^{8),9)}を踏襲し、まずは通常の流出解析と同様の条件を念頭においた設定とする。

表-1 モデル構造および固定するハイパーパラメータ

項目	内容
学習エポック数	1,000
学習洪水	Q(1)~Q(学習洪水数)
入力層	現時刻~(入力層ノード数-1)分前の1分間流域平均降雨量(mm/min)
中間層	・深層学習モデル: 2層(半分ずつ分割) ・ANN: 1層
出力層	現時刻流出高(mm/min)
活性化関数	ReLU関数
ミニバッチサイズ	60
最適化手法	Adam

表-2 検討ケース

項目	設定ケース	ケース数
学習洪水数	1,10,20,40,60	280ケース (5×8×7)
入力層ノード数	60,90,120,150,180, 210,240,270	
中間層ノード数	10,20,30,60,120,240,480	

入力層のノード数は現時刻~(ノード数-1)分前の1分間流域平均降雨量(mm/min)として設定し、出力層のノードは現時刻流出高(mm/min)とした。すなわち現時刻以前の雨量から現時刻の流出高を算出するモデルである。深層学習モデルの場合、中間層ノード数を中間層1と中間層2に半分ずつ分割することとした。表-1のようにハイパーパラメータ(学習エポック数、活性化関数、ミニバッチサイズ、最適化手法)を設定した。なお、入力層および出力層に与えるデータのスケールがそれぞれ揃っているため、正規化を実施していない。

本論文では機械学習モデルの学習・テスト・検証データをMaier¹⁴⁾らより以下のように定義付ける。

- ・学習データ: ネットワークモデルの学習(重み係数やバイアスの調整)に用いる
- ・テストデータ: 交差検証(過学習を防ぐため、目的関数の確認により、学習の停止を判断)に用いる
- ・検証データ: 学習や交差検証に用いず、構築したモデルにより精度を確認するために用いる

なお、データセットとして用いる100洪水のデータのうちNo.1~60を学習、No.61~80をテスト、No.81~100を検証用洪水として用いることを想定している。洪水予測等の流出解析を実施する条件と同様に、時系列的に過去のデータを学習データとして採用することとした。

(2) 検討ケースおよび評価関数

本検討ではそれぞれ表-2に示すように学習洪水数5ケースと入力層ノード数8ケース、中間層ノード数7ケースを組み合わせにより280ケースを実施することとした。

深層学習モデルのハイドログラフの再現精度は次式に示す二乗平均平方根誤差(Root Mean Square Error: RMSE)により評価した。

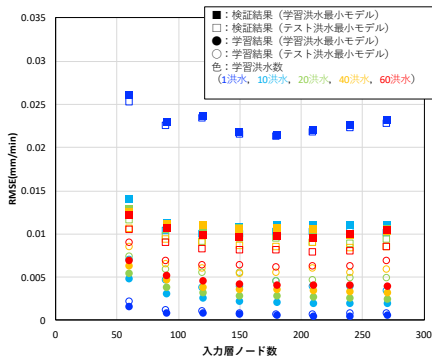


図-5 入力層ノード数と RMSE の関係

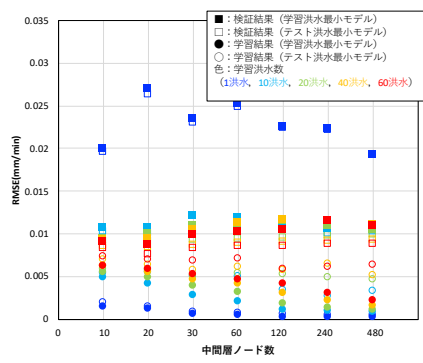


図-6 中間層ノード数と RMSE の関係

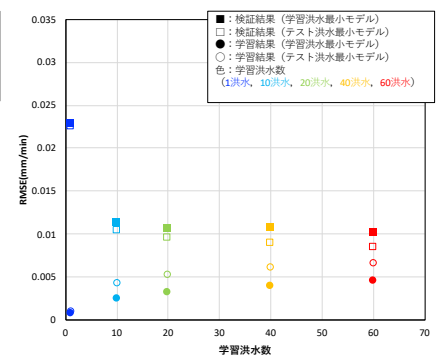


図-7 学習洪水数と RMSE の関係

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_{oi} - q_{ci})^2} \quad (4)$$

ここに、 n ：サンプル数、 q_o ：実績流出高(mm/min)、 q_c ：深層学習モデルによる計算流出高(mm/min)である。

計算は1,000エポック（1エポック：入力データを一通り学習に使用）まで実施し、図-4に示すエポック数とRMSEの関係を基に、テスト用洪水のRMSEが最小となるエポックのモデルを最終モデルとした（以下、テスト洪水最小モデルと呼ぶ）。また、参考までに学習洪水でRMSEが最小となるエポックのモデル（以下、学習洪水最小モデルと呼ぶ）も保存し、その精度を確認した。学習したモデルによるハイドログラフの再現精度の確認は、学習用と検証用の入力データを機械学習モデルに与え、それぞれの出力ハイドログラフの結果を実績流出高と比較し、RMSEを算出した。

4. 結果および考察

(1) 入力層ノード数とハイドログラフの再現精度の関係

学習洪水数、入力ノード数、中間層ノード数を組み合わせにより280ケース学習させた結果について、それぞれの影響について評価する。学習洪水のRMSEは学習した洪水の計算ハイドログラフと実績ハイドログラフのRMSEの平均であり、検証洪水のRMSEは洪水No.81～100の平均RMSEである。

図-5に入力層ノード数とRMSEの関係を示す。洪水数と入力層ノード数が同じ場合で、中間層ノード数を変更したときの7ケースのRMSEを平均化している。中途り丸印は学習洪水最小モデルによる学習洪水のハイドログラフの再現結果を示し、中抜き丸印はテスト洪水最小モデルによる学習洪水のハイドログラフの再現結果を示す。色は学習洪水数を示しており、寒色系から暖色系の色になるに従い、学習洪水数が多くなる。また、四角印は検証洪水のハイドログラフの再現結果を示している。図-5より入力ノード数が90以下の場合、学習洪水、検証洪水問わず、入力層ノード数が少なくなるに従いRMSEが上昇し、ハイドログラフの再現精度が低下する。また、

それ以上の場合、RMSEは大きく変わらなかった。

(2) 中間層ノード数とハイドログラフの再現精度の関係

図-6に中間層ノード数とRMSEの関係を示す。洪水数と中間層ノード数が同じ場合で、入力層ノード数を変更したときの8ケースのRMSEを平均化している。図-6の検証洪水（図中、中抜き赤四角）において中間層ノード数が20の場合に、検証洪水の検討ケースの中で最もRMSEが小さくなった。これは中間層ノード数を少なくすることにより、入力層が持つ情報量が圧縮され、学習データの特徴を掴むように学習することから、検証洪水で精度が向上したものと考えられる。また、中間層ノード数が10の場合のRMSEの増加は、情報量が圧縮されすぎることによる表現力の低下が原因と考えられる。

(3) 学習洪水数とハイドログラフの再現精度の関係

図-7に学習洪水数とRMSEの関係を示す。洪水数が同じ場合で、入力層ノード数と中間層ノード数を変更したときの56ケースのRMSEを平均化している。図-7の学習洪水では学習洪水数が少ない場合に、RMSEが小さくなり、ハイドログラフの再現精度が向上していることがわかる。また、逆に検証洪水では学習洪水数が少ない場合（寒色系）、特に学習洪水数1洪水（青四角）の場合にRMSEが大きくなり、学習洪水数10洪水（水色）以上であればRMSEは大きくは低下しない傾向であった。

(4) PD比による評価

本論文ではハイドログラフの再現精度を整理する指標としてモデルパラメータ数 N_p と学習に用いた観測データ数 N_D の比として以下のPD比を提案する。

$$PD比 = N_p / N_D \quad (5)$$

画像認識分野の深層学習モデルではパラメータ数の方がデータ数より圧倒的に多いモデル（PD比が1よりかなり大きい）が多い¹⁵⁾。一般的にPD比が大きいほどデータ数に対し複雑なモデルとなり、過学習が生じやすい。

学習洪水におけるPD比とRMSEの関係を図-8 a)に、検証洪水におけるPD比とRMSEの関係を図-8 b)に示す。

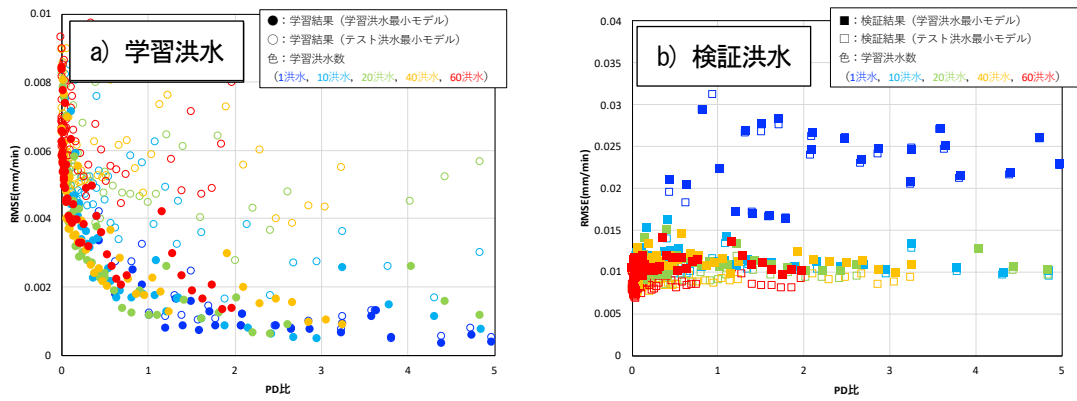


図-8 PD比とRMSEの関係 (a)学習洪水, b)検証洪水

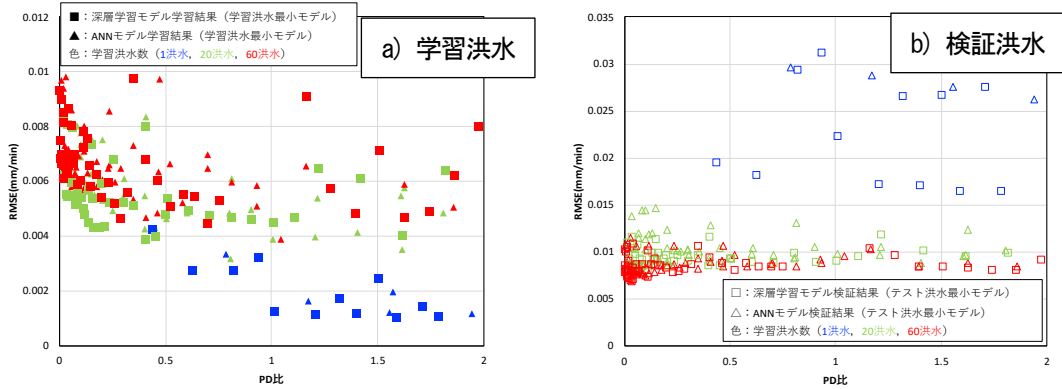


図-9 深層学習モデルとANNモデルの比較 (a)学習洪水, b)検証洪水

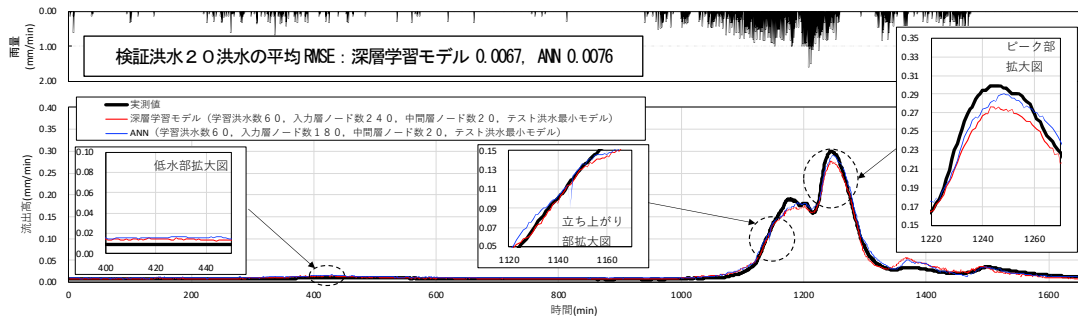


図-10 検証洪水 (No.91 : 2009.10.17) における深層学習モデルとANNモデルのハイドログラフの再現精度の比較

図-8 a)より学習洪水の学習洪水最小モデルにおいては、PD比が1以下(データ数>パラメータ数)の場合は急激にRMSEが大きくなり、ハイドログラフの再現精度が低下する。次に、図-8 b)の検証結果では、学習洪水の場合と異なり、PD比が1以下でも学習洪水数が1洪水の場合を除きRMSEの大きな増加は見られない。さらに、学習洪水数が増加するに従いRMSEが低下し、60洪水の場合にPD比が0.1程度で検証洪水で最もRMSEが小さくなった。これはパラメータ数の10倍データ数がある場合である。これより、検証洪水のハイドログラフの再現精度には学習洪水数が大きく影響しており、学習データの多様性が重要であることがわかる。

(5)深層学習モデルとANNモデルの比較

深層学習モデルとANNモデルによる学習洪水に対するPD比とRMSEの関係を図-9 a)に、検証洪水に対する

PD比とRMSEの関係を図-9 b)に示す。図-9 a)より学習洪水において同程度のPD比の場合には、全体的に四角印の方が三角印よりもRMSEが小さく、深層学習モデルのほうがハイドログラフの再現精度が良いことがわかった。また、図-9 b)において検証洪水でも同様に深層学習モデルの方がハイドログラフの再現精度が良かった。

次に、各モデルで最も検証洪水のハイドログラフの再現精度が良かったケースの検証洪水最大規模のNo.91洪水におけるハイドログラフの比較を図-10に示す。深層学習モデル、ANNモデルともに実績流出波形の傾向変動は精度良く再現できている。ただし、拡大図を確認すると、両モデルによる計算結果で不規則変動が確認できるが、実測ハイドログラフにも不規則変動が確認できる。これは、大河川流域での研究の再現波形では目立って見られないことから短時間間隔の観測データに起因するものと考えられる。特にANNモデルでは立ち上がり部拡

大図で見られる実績に沿わない急激な波形の変動が見られ、これが再現精度低下につながっていると考えられる。

5. むすび

流出機構が複雑な都市中小河川の実流域を対象とし、1分ごとのデータセットを用いて深層学習によりハイパーパラメータを変更させた場合の降雨流出ベンチマークテストを実施した。その結果、以下のことがわかった。

- ・入力層ノード数が少ない場合に学習、検証洪水においてハイドログラフの再現精度が低下した。
- ・中間層ノード数を減少させた場合に、学習した洪水の特徴を掴むよう学習することで、検証洪水で最もハイドログラフの再現精度が良いケースが確認された。
- ・学習洪水数が少ない場合は過学習の状態となる。検証洪水の精度の向上には学習洪水数の増加が効果的である。
- ・PD比が1以下で学習洪水のRMSEが上昇するが、検証洪水では学習洪水数の方がハイドログラフの再現精度に対する影響が大きかった。
- ・ANNモデルは実績に沿わない急激な変動が見られるため、深層学習モデルに比べハイドログラフの再現精度が低いことがわかった。
- ・1分値を用いた今回の機械学習モデルの再現波形は、短い観測間隔の入力データに伴う不規則変動が見られる。今後は深層学習モデルの学習データの不均衡性に着目した検討を実施する予定である。

参考文献

- 1) 河村明:都市流域における洪水流出解析の現状と将来展望, 水文水資源学会誌, Vol.31, No.6, pp.219-234, 2018.
- 2) 深層学習は土木を変える? -AI活用の可能性を探る-, 土木学会誌, Vol.103, No.2, pp.7-37, 2018.
- 3) 磯部勇, 大河戸輝夫, 羽生田英彦, 小田誠一, 後藤祐輔:ニューラルネットワークによる水位予測システムの開発, 水文・水資源学会誌, Vol.7, No.2, pp.90-97, 1994.
- 4) 一言正之, 桜庭雅明, 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), 72巻, 4号, p. I_187-I_192, 2016.
- 5) 田村和則, 加納茂紀, 三浦心, 山脇正嗣, 金子拓史: ダム流入量長時間予測への深層学習の適用-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), 74巻, 5号, p. I_1327-I_1332, 2018.
- 6) 房前和朋, 島本卓三: AIを活用した洪水予測技術の開発について, 平成30年度九州国土交通研究会, 2018.
- 7) 東京都建設局河川部防災課: 東京都水防計画2020.
- 8) 藤塚慎太郎, 河村明, 天口英雄, 高崎忠勝: AIを用いた都市流出モデルエミュレーションのためのベンチマーク模擬ハイドログラフの構築, 第46回土木学会関東支部研究発表会講演集, CD-ROM版(II-20), 2019.
- 9) 藤塚慎太郎, 河村明, 天口英雄, 高崎忠勝: ベンチマークバーチャルハイエト・ハイドログラフを用いた深層学習による都市流出モデルのエミュレーション, 土木学会論文集G (環境), Vol.75, No.5, pp. I289I296, 2019.
- 10) 藤塚慎太郎, 河村明, 天口英雄, 高崎忠勝: ニューラルネットワークおよび深層学習による都市流出モデルのエミュレーション性能評価, 土木学会論文集B1 (水工学), Vol.75, No.2, pp. I229I234, 2019.
- 11) 高崎忠勝, 河村明, 天口英雄, 石原成幸: AI降雨流出ベンチマークテストに向けた都市中小河川実流域データセットの作成, 水文・水資源学会2019年度研究発表会要旨集, pp.28-29., 2019.
- 12) 東京都: 荒川水系神田川流域河川整備計画, 2016.
- 13) 岡谷貴之: 深層学習, 講談社サイエンティフィック, 2015.
- 14) Maier, H.R., 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.
- 15) 岡谷貴之: 画像認識のための深層学習の研究動向, 一畳込みニューラルネットワークとその利用法の発展一, 人工知能 Vol. 31 No. 2, pp.169-179, 2016.

(Received June 30, 2020)

(Accepted August 28, 2020)

RAINFALL RUNOFF BENCHMARK TEST BY DEEP LEARNING MODEL USING URBAN MEDIUM AND SMALL RIVER BASIN DATASET

Shintaro FUJIZUKA, Akira KAWAMURA, Hideo AMAGUCHI and Tadakatsu TAKASAKI

Urban floods occur frequently, and there are great expectations for the application of machine learning models that can be easily constructed to flood forecasting fields in urban medium and small size basin, which have complicated runoff mechanisms. Therefore, in this paper, we constructed a deep learning model using an actual river basin dataset of urban small and medium rivers, which was created at a short observation interval of 1 minute, which is different from that of large river basins. Benchmark tests were performed when the parameters were changed. An ANN model was also constructed for comparison with the deep learning model, and the performance was evaluated by a new index called the PD ratio, which is the number of parameters of the deep learning model and the number of observed data. As a result, it was found that the deep learning model is superior to the ANN model for the learning/verification flood at the same PD ratio, and in particular, the ANN model shows rapid fluctuations in the verification flood that do not follow the actual results.