

深層学習による流量予測精度と流域面積との関係

RELATION BETWEEN FLOW RATE PREDICTION USING DEEP LEARNING AND BASIN AREA

北海道大学工学部環境社会工学科 非会員 齋藤 翔太 (Shota Saito)
北海道大学大学院工学研究院 ○正 員 田中 岳 (Gaku Tanaka)

1. はじめに

昨今の頻発する洪水災害に対して、その被害は河川整備が行き届かない地方で拡大する傾向にある。このように洪水災害に対して脆弱な地方では、財政、人材の課題から、利便性を第一に洪水予測システムを選択せざるを得ない状況にある。

一般的な洪水予測システムは、一つの集中型モデル、集中型モデルの分布型、あるいは分布型モデルなど、様々なタイプの流出モデルが基盤となるが、これらとは異なり、画像認識などに応用される深層学習をこの予測計算に用いる試み^{1,2)}が注目されている。然しながら、深層学習は多量の学習データを必要とするため、その適用例は水文資料が十分に整った一級河川に限られ、地方自治体が管理する二級河川への例は極めて少ない。

本研究では、模擬された降雨流出現象（仮想的な学習・検証データセット）に対して、深層学習による流量予測の精度と流域面積との関係を検討し、この手法を中小河川に適用するための条件等の基礎的な知見を得ることを目的としている。

2. 深層学習モデルと模擬された降雨流出現象

(1) 深層学習モデル

ここ数年の間に、深層学習のための多種多様なオープンソースが容易に入手可能となってきた。本論文では、Python のオープンソース機械学習ライブラリの scikit-learn を用いる。そこで、入力値には、流量を予測する目的地点（流域末端）の一時間前と現在の流量、それよりも上流側の二時間前から現在までの流量と雨量を与えて三時間後の流量を予測する。以下に、モデル内で行われる計算の過程を記す。図-1 は、深層学習モデルの概念図である。この図に示す中間層、出力層の各素子の中では次のような計算を行う。

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

ここで、 u ：各素子の入力値； x ：入力値； θ ：バイアス、 K ：各層の素子数； $f(u)$ ：活性化関数（中間層：ReLU 関数、出力層：恒等関数）； z ：素子の出力。出力層の出力値と目標出力との誤差が小さくなるように学習を行い、その結果を次式で評価する。

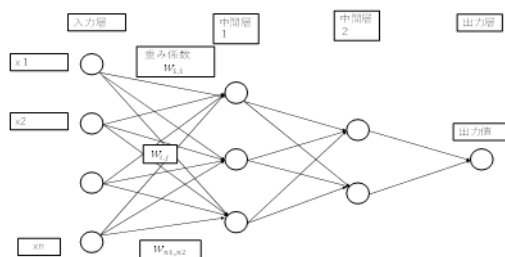


図-1 深層学習モデルの概念図

$$E(w) = \frac{1}{2} \sum_{n=1}^N (d_n - y(x_n; w))^2$$

E ：二乗誤差； N ：サンプル数； d ：目標出力； y ：ネットワークの出力値。重み w の更新は Adam で行った。これは Kingma らが提唱する確率的勾配降下法の最適化手法³⁾で、以下の式により計算される。

$$m_{t+1} = \beta_1 m_t + (1 - \beta_1) \nabla E(w^t),$$

$$v_{t+1} = \beta_2 v_t + (1 - \beta_2) \nabla E(w^t)^2,$$

$$\hat{m} = \frac{m_{t+1}}{1 - \beta_1^t}, \quad \hat{v} = \frac{v_{t+1}}{1 - \beta_2^t}, \quad w_{t+1} = w_t - \alpha \frac{\hat{m}}{\sqrt{\hat{v}} + \epsilon}$$

m, v はそれぞれ一次、二次のモーメントを表す。なお、これらの初期値は 0 、 $\alpha = 0.001$ 、 $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$ 、 $\epsilon = 10^{-8}$ である。

(2) 降雨流出現象のモデル化

実際の河道網の構造が自己相似性⁴⁾を示すことから、本論文では、三本の河道が合流する Peano network⁵⁾を用いる（図-2）。なお、図の実線は河道を、破線は各河道の両岸に連結された三角形斜面を、●印が流域末端を、 $i (\leq 2^n)$ は河道（リンク）の位置を表している。また、斜面、河道の各要素モデルには、

$$s_h = k_h q_h^{p_h}, \quad \frac{ds_h}{dt} + q_h = r$$

なお、 s_h ：貯留高； q_h ：流出高； r ：降雨強度； t ：時間； k_h ：貯留係数； p_h ：貯留指数。

河道要素モデル：

$$s_i = k_i q_i^{p_i}, \quad \begin{cases} \frac{ds_{i+1}}{dt} + q_{i+1} = A q_i + q_i + 2q_m, & i = (2l-1)m \\ \frac{ds_1}{dt} + q_1 = A q_h \end{cases}$$

なお、 s_i ：貯留量； q_i ：流出量； k_i ：貯留係数； p_i ：貯留指数； $m = 2^{n-k}$ ； $k (\leq n)$ ：自然数； $l (\leq 2^{k-1})$ ：自然数； A ：一河道二斜面からなる最小要素の面積。添え字 i は、河道の位置（図-2）と対応する。また q_m は、 $i+1$ 番目の河道の上流端に連結された左右の流域からの流入量を意味する。

(3) データセット作成条件

ここでは、降雨継続時間が $72(\text{hr})$ 、平均降雨強度は $5, 30, 50(\text{mm/hr})$ になる矩形降雨 R と三角形降雨 T の計 6 パターン、 $\Delta t = 1(\text{hr})$ として、それぞれ 101 ケースの模

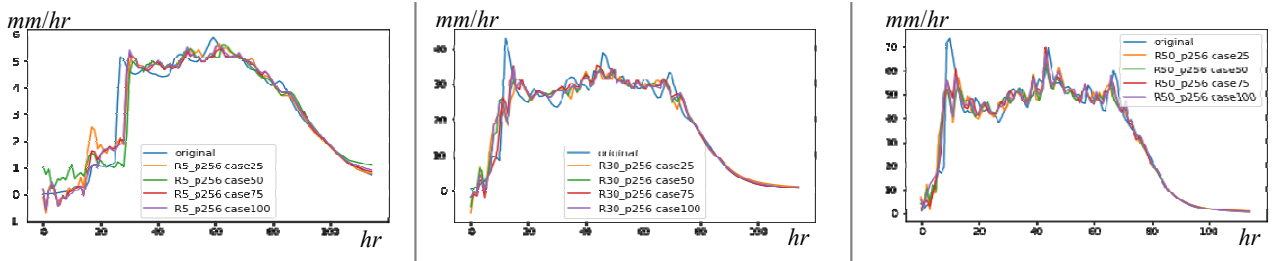
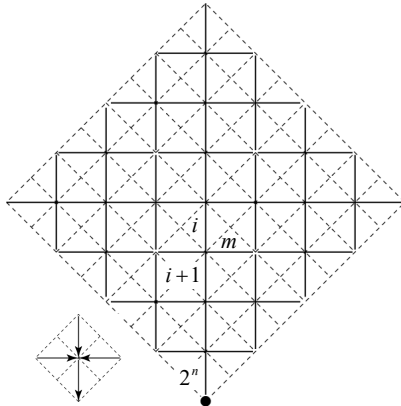


図-3 p256の降水規模（矩形降雨）毎の計算結果（流出高）

図-2 流域と模擬河道網⁵⁾表-1 パラメータ n と流域面積 (km²)

n	1	2	3	4
地点 p	2	4	8	16
面積	8.000	32.00	128.0	512.0
n	5	6	7	8
地点 p	32	64	128	256
面積	2048	8192	32768	131072

擬降雨時系列を作成し、先に述べた流出モデルを介して学習・検証データセットを構築した。なお、各降雨時系列は、その変動係数を 0.5 に固定し、他の条件は以下のように設定し、

$$n_h = 0.10, i_h = 0.2, l_0 = 2000 (m), i_1 = 0.01, \\ w_1 = 0.8 (m), C = 32.7$$

$n = 8$ に対応した流域規模まで流出計算を実施した。 n と流域面積との関係を表-1 にまとめる。なお、例えば、流量を予測する目的地点が地点 $p8$ の場合の学習には、それよりも上流側の流量として地点 $p2, p4$ と最上流地点 $p1$ の計算結果と降雨のみを用いている。

3. 計算結果と考察

図-3 に計算結果（検証データと深層学習による流出予測結果の比較）の一例を示す。ニューラルネットワークの性能評価は、以下の Root Mean Squared Error (RMSE) で行われることが多い。

$$RMSE = \sqrt{\frac{\sum (P_i - Q_i)^2}{n}}$$

なお、 P_i : 目標出力; Q_i : 予測結果。この値が 0 に近いほど高精度であることを示す。計算結果の RMSE（降雨時系列の標準偏差で基準化）のうち降水規模が 5mm/hr のものを表-2 に示す。この表からデータ数の少ない地

表-2 R5（矩形降雨 5mm/hr）での各地点の RMSE

	25case	50case	75case	100case
$p256$	0.237	0.286	0.211	0.214
$p128$	0.250	0.202	0.179	0.169
$p64$	0.096	0.085	0.078	0.070
$p32$	0.106	0.082	0.072	0.073
$p16$	0.148	0.145	0.147	0.149
$p8$	0.482	0.468	0.464	0.465
$p4$	0.569	0.568	0.572	0.573
$p2$	0.741	0.733	0.736	0.728

点 $p2 \sim p8$, データ数が多く流域面積も大きい地点 $p128$ と $p256$ については、予測結果の精度が低いことが分かる。また、学習したケース数に着目すると、地点 $p2 \sim p8$ までは、ケース数に変わりなく予測精度が低く、他の地点ではケース数が大きくなるにつれて予測精度の向上がみられる。しかし、ケース数 75 以上になるとその変化は小さくなる。

4. おわりに

本論文では、模擬された降雨流出現象（仮想的な学習・検証データセット）において、深層学習による流量予測精度と流域面積との関係を検討した。その結果、ここで実施した数値実験の範囲ではあるが、深層学習による流出予測精度は、流域面積と観測点数に依存することが認められた。然しながら、その適用条件の特定までには至っていない。今後は、この点を明らかにする予定である。

参考文献

- 1) 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, I_187-I_192, 2016.
- 2) 一言正之, 櫻庭雅明: 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), Vol. 73, No. 1, pp. 22-33, 2017.
- 3) Kingma, D. P. and Ba, J. L.: Adam: A Method for Stochastic Optimization, ICLR, 2015.
- 4) Rodríguez-Iturbe, I. and Rinaldo, A.: Fractal River Basins: Chance and Self-Organization, Cambridge University Press, pp. 1-547, 2001.
- 5) Gupta, V. K., Castro, S. L. and Over, T. M.: On scaling exponents of spation peak flows from rainfall river network geometry, *J. Hydrol.*, Vol. 187, pp. 81-104, 1996.