

# Transformer を用いた洪水予報の改善

熊本大学工学部土木建築学科 学生会員 浦越 淳  
 熊本大学 くまもと水環境・減災研究教研究センター 正会員 石田 桂  
 熊本大学自然科学教育部 学生会員 永里 起義  
 熊本大学自然科学教育部 学生会員 坂口 大珠

## 1. 序論

近年、2018年の西日本豪雨、九州内においては2020年の熊本豪雨など、日本各地において豪雨災害が多発している。それに伴い大規模な洪水等が発生し、多くの被害をもたらしている。そのため、今後の被害軽減や防止のために河川流量予報の高精度化が重要である。

河川流量予報には事例ベースモデルや分布型流出モデル等、様々な手法が存在している。また近年、様々な分野で深層学習 (deep learning) が用いられ高い評価を受けている。河川流量予報にも深層学習は用いられており、例えば、福丸ら (2022) は深層学習手法の一つである Long Short-Term Memory (LSTM) を用いて河川流量予報を行っている<sup>1)</sup>。一方で、深層学習は発展が著しく、深層学習の新しい手法を用いることで河川流量予報にも改善の可能性があると考えられる。例えば、2017年に発表された Transformer という深層学習モデルは、河川流量予報ではないが時系列予測に用いられ、高い精度を示している<sup>2)</sup>。

そこで、本研究では Transformer を用いた河川流量予報モデルの構築を目的とする。Transformer を用いることにより予報の高精度化が期待される。

## 2. 手法

Transformer は自然言語処理分野において高い精度が得られることで注目を集めた深層学習モデルであり、その

他の分野でも高い精度が得られることが確認されている。Transformer は Encoder と Decoder の2つの部分に大きく分けられる (図-1)。従来の RNN や CNN といった再帰計算を用いた深層学習モデルとは異なり、Attention のみを用いた構造になっている。そのことにより、並列計算が可能になり計算の高速化につながっている。Attention は以下の式で表される。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

ここで  $Q$  は Query,  $K$  は Key,  $V$  は Value,  $\sqrt{d_k}$  はスケーリング因子である。まず Query と Key の類似度  $e$  を、類似度を計算する関数  $f$  を用いて計算する。その後、正規化関数  $g$  を用いて  $e$  を正規化し重み  $a$  を算出する。そして  $a$  と Value の要素積を取り、出力を計算する。

Transformer は再帰計算を用いないことが特徴であるが、それにより時系列データの順序や位置関係を得られない問題が発生する。そのため、Positional Encoding という処理を用いて、データの位置関係を取得する。Positional Encoding は以下の式で表される。

$$PE_{(\text{pos}, 2i)} = \sin(\text{pos}/10000^{2i/d_{\text{model}}})$$

$$PE_{(\text{pos}, 2i+1)} = \cos(\text{pos}/10000^{2i/d_{\text{model}}})$$

ここで  $\text{pos}$  は配列の要素の位置、 $i (i=1, 2, \dots, d_{\text{model}})$  は入力次元である。つまり  $\text{pos}$  は行列の何列目かを、 $i$  は行列の何行目かを表している。

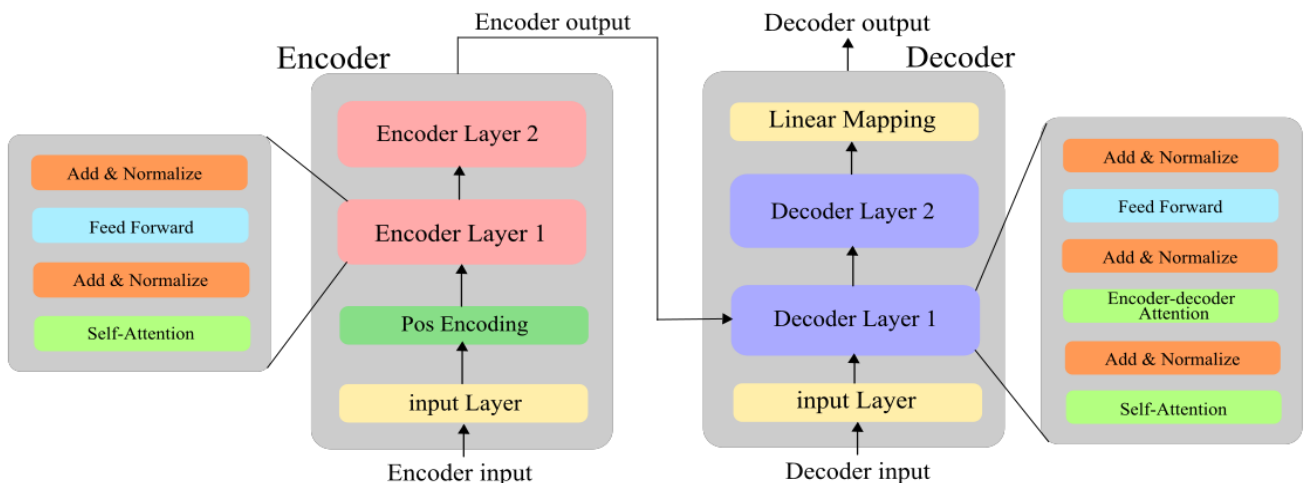


図-1 Transformer の構造

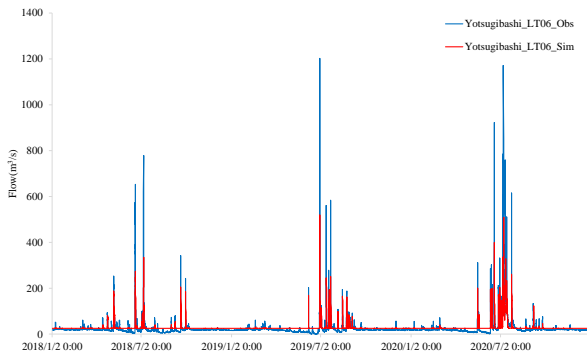


図-2 6時間先の流量予報

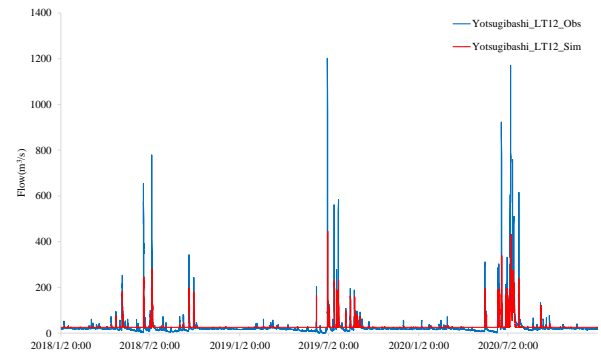


図-3 12時間先の流量予報

表-1 ハイパーパラメータ

Dropout の確率	0.2
Multi-Head Attention の Head 数	2
Encoder Layer の数	4
Decoder Layer の数	4
$d_{\text{model}}$	128
損失関数	MSE
最適化関数	Adam
バッチサイズ	512*4
入力データの長さ	24

### 3. 対象地・データセット

熊本県白川流域を対象地に選定した。白川は流域面積が480平方キロメートル、流路延長が74キロメートルである、阿蘇山根子岳を水源とする一級河川である。中流部は急流で水の流れが速く、熊本市街部が広がる下流部や低平地の広がる河口部は緩やかな地形となっているため、洪水を引き起こしやすく被害も拡大しやすいという特徴を持つ。

白川流域の1時間ごとの流量は国土交通省水文水質データベースより観測値を得た。対象とする流量の観測点は代継橋である。1時間ごとの降水量は解析雨量から取得し、小流域毎の面積平均値を算出し入力として用いた。

### 4. 適用

Transformerの実装には、Pythonのオープンソース機械学習ライブラリであるPytorchを用いた。Transformerにおける本研究で設定したハイパーパラメータを示す(表-1)。本研究では、24時間分のデータを入力し12時間先までの流量予報を行った。

本研究ではデータセットを、モデルの学習を行う訓練期間(2006-2015)、過学習が生じていないか確認する検証期間(2016-2017)モデルの推定精度の最終的な評価を行うテスト期間(2018-2020)の3つに分けた。学習においてPatienceを30としたEarly Stoppingを用いた。Transformerの学習過程では初期条件やバッチ選択にランダム性があ

り、同じモデルの場合でも学習ごとの結果にばらつきが生じるため、各モデル設定につき100回ずつ学習を行った。また、モデル精度の評価基準としてNash-Sutcliffe Efficiency(NSE)を用いた。NSEは、負の無限大から1で表され、値が1に近づくほどモデルの精度が良いとされる。NSEの式は以下に示す。

$$NSE = 1 - \frac{\sum_{t=1}^T (Q_o^t - Q_m^t)^2}{\sum_{t=1}^T (Q_o^t - \bar{Q}_o)^2}$$

$Q_o$ : 実測流量  $Q_m$ : 計算流量  $\bar{Q}_o$ : 実測流量の平均値

### 5. 結果・考察

本研究における6時間先および12時間先の河川流量予報の結果を示す(図-2, 3)。NSEの値はそれぞれ0.478, 0.286を示した。どちらも0.5を下回っており、高い予報精度は得られなかった。特にピーク時の流量に関しては、6時間先、12時間先共に予測値が実測値の10分の1以下の値をとっていた。今後、他のモデルとの比較やハイパーパラメータの調節を行う必要があると考えられる。

### 6. 結論

本研究では白川流域を対象に、深層学習手法の一つであるTransformerを用いた河川流量予報モデルの構築を行った。結果が示すように、Transformerを用いることにより、河川流量予報において高い予報精度を得ることはできなかった。今後、LSTM等の他のモデルとの比較を行い、また、ハイパーパラメータの変更等でさらなる精度向上を目指す。

### 参考文献

- 1) 福丸大智, 赤松良久, 新谷哲也, 藤井晴香: 深層学習を用いた6時間先の河川水位予測実現に向けた基礎的検討, 自然災害研究協議会中国地区部会研究論文集, 第8号 17-20, 2022
- 2) 本田良司, 小笠原範光, 小玉亮, 日本気象協会: Transformerによる時系列予測, 人工知能学会全国大会論文集, 第34回, 2020