

Transformer を用いた降水量予報

熊本大学工学部土木建築学科 学生会員 大内田華音
 熊本大学 くまもと水循環・減災研究教育センター 正会員 石田桂
 熊本大学自然科学教育部 学生会員 永里赳義
 熊本大学自然科学教育部 学生会員 坂口大珠

1. 序論

近年、異常気象が多発しており、集中豪雨による河川の氾濫は全国に大きな被害をもたらしている。集中豪雨の発生回数の増加に伴い、洪水による被害は年々深刻化している。そのような中、河川の氾濫による洪水の被害を最小限に抑えるために、河川流域範囲の降水量予報の高精度化が求められている。

降水量予報の手法として、昨今深層学習を用いた事例が多く存在する。例えば、Shi et al (2015 年) は、深層学習による降水量予報に、高い適用可能性を示している¹⁾。

一方で、深層学習モデルの一つである Transformer は、従来のモデルと比較すると、学習効率が良く、学習時間の割に高精度な結果が得られることから、近年注目されている。例えば、本田ら (2020 年) は、気象データの時系列予測において、Long-and short-term temporal patterns with deep neural networks (LSTNet) と比較し、Transformer により良い精度が得られることを示している²⁾。

そこで、本研究では、高精度化を目指し、Transformer を降水量予報に用いる。対象地は熊本県に位置する白川流域とし、3次元の大気データから降水量を予報するモデルを実装する。

2. 手法

Transformer は、2017年に"Attention is all you need"³⁾で提案された、自然言語処理を目指したモデルである。その構造は、図-1で示すように、Encoder と Decoder で構成されている。Long Short Term Memory (LSTM) や Gated Recurrent Unit (GRU) などの従来の Recurrent Neural Networks (RNN) 系モデルと比較すると、の再帰計算が行われる層は用いられておらず、その代わりとして Attention という機構のみが活かされているのが大きな特徴である。Attention は時系列データのどこに注目すべきか、動的に特定し、時系列データに重みづけをすることができる。そのため、並列計算が可能になり、学習時間が短縮される。ただし、Attention のみでは、時系列データの順序・位置関係の情報を取得することができない。そのため、Positional Encoding という機構を用いて、時系列データの位置情報の埋め込みを行う。このように、Transformer は、時系列データの位置関係を保持したまま、直接的に学習することができるため、学習効率が良く、学

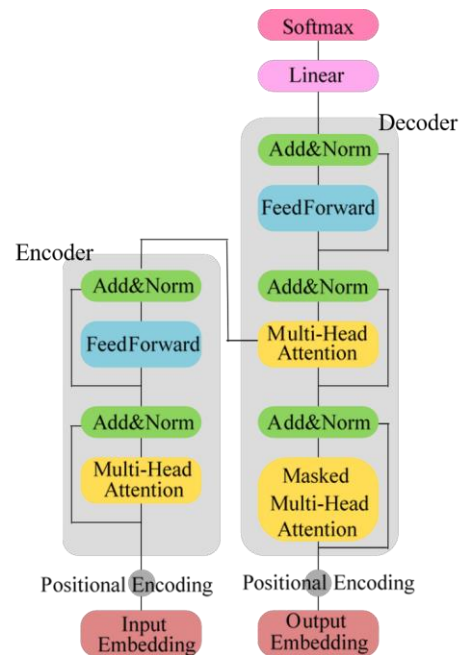


図-1 Transformer の概略図

習時間の割に、高精度な結果が得られるのが特徴とされている。

3. 対象地・データセット

本研究では白川流域を対象地とした。白川は熊本県高森町阿蘇根子岳を水源とし、熊本市密集市街地を貫流し有明海に注ぐ一級河川である。白川の流域面積は480km²、幹川流路延長は74kmである。

入力データである気象データは、ERA5より取得した。ERA5は、European Centre For Medium-Range Wether Forecasts (ECMWF)によって作成された全球レベルの大気情報の再解析データセットである。また、1時間毎の降水量データを、気象庁の解析雨量より取得し、白川流域の流域平均値を算出し対象データとした。使用する入力変数は、2次元変数であるSkin Tempreture, Total cumulns water, 3次元変数である比湿, 気温, 経度方向の風速, 緯度方向の風速, ジオポテンシャルの計7つとした。また、以上の入力変数の範囲は、白川流域の範囲である東経130.5度~131.25度, 北

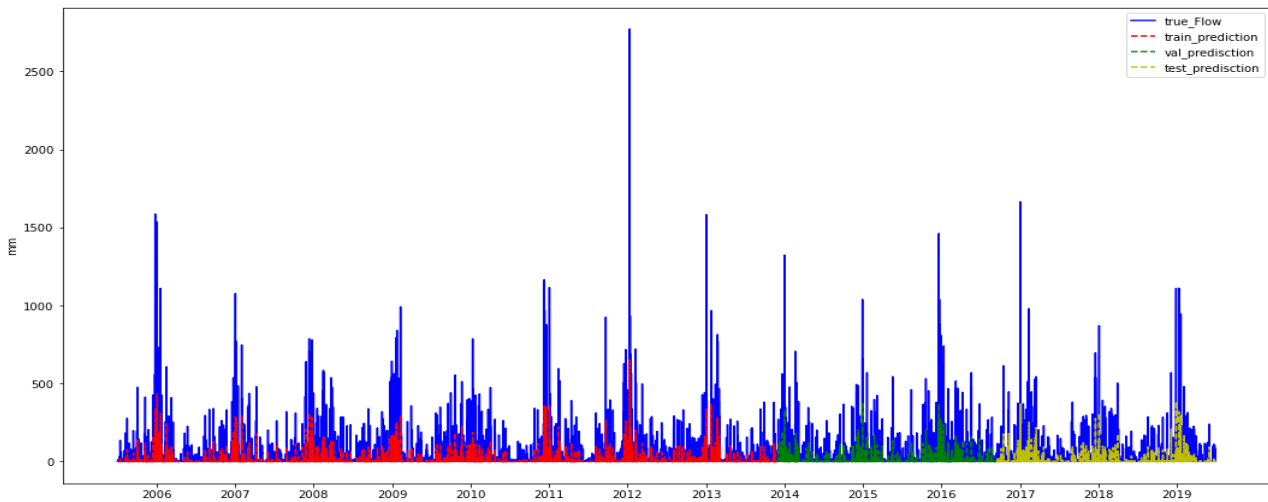


図-2 6時間先の降水量予報

表-1 ハイパーパラメータの設定

Encoderへの入力長	6
Decoderの出力長	6
バッチサイズ	8
EarlyStoppingのPatience数	30
Multi-Head AttentionのHead数	8
最適化関数	Adam
誤差関数	平均二乗誤差

緯 32.75 度～33 度とした。3 次元変数は 1000hPa の圧力面のもののみを用いた。対象期間は 2006 年～2019 年とした。

4. 適用

実装には、深層学習ライブラリである Keras を用いた。また、Encoder の入力長は 6 時間、Decoder の入力長も 6 時間とし、6 時間先までの降水量予報を行った。データセットは、ハイパーパラメータの調整を行う訓練期間（2006～2013 年）、過学習を起こしていないかを確認するための検証期間（2014～2016 年）、モデルの推定精度の最終的な評価を行うテスト期間（2017～2019 年）の 3 つに分けた。また、モデルを評価する指標として、平均平方二乗誤差 Root Mean Square Error (RMSE) (1) を用いた。 y_i は i 番目の実測値、 \hat{y}_i は i 番目の予測値、 n はデータ総数を示す。RMSE の取り得る範囲は 0 から無限大であり、0 に近いほど、そのモデルの精度は良いとされる。なお、ハイパーパラメータの設定は表-1 に示す。

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (1)$$

5. 結果

6 時間後の降水量予報の結果を図-2 に示す。青色の数値が実測値、赤色の数値が訓練期間のモデルの予測値、緑色の数値が検証期間のモデルの予測値、黄色の数値がテスト期間のモデルの予測値を示している。図-2 より、実測値の変動の特徴は捉えられてはいるが、ピーク時の精度が低く、全体的に実測値が予測値の 2 倍程の値になっていることがわかる。なお、RMSE は、訓練期間が 0.429mm、検証期間が 0.416mm、テスト期間が 0.408mm という値になった。

6. まとめ

本研究では、Transformer を用いて 3 次元大気データを入力とし 6 時間先までの降水量予報を行った。Transformer を降水量予報に適用することができたが、現状で誤差が小さいとは言えない。今後、Transformer による予報精度向上のために、入力に用いる変数の種類・時空間範囲の選定、Transformer の構造及びハイパーパラメータの最適化などを行う必要がある。

参考文献

- 1) Shi X, Chen Z, Wang H, Yeung D, W & K Wong, W & C Woo : Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, 29th Annual Conference on Neural Information Processing Systems, 2015.
- 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, 34, 2020.
- 3) Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin : Attention is all you need, In Proceedings of the 31 st conference on Neural Information Processing Systems, 2017.