

降水量モデリングにおける CNN の特性

熊本大学工学部土木建築学科 学生会員 井ノ口 宗佑

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 横尾 和樹

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、深層学習は様々な分野で注目を集めており、深層学習を使用したダウンスケーリング手法の開発が進んでいる。深層学習の中で、畳み込みニューラルネットワーク(CNN)はしばしばダウンスケーリングに使用されている。例えば、板谷ら(2018)は深層学習を用いてダウンスケーリングを行い、入力データとして使用する変数および領域の選択により精度向上の可能性を示唆した[1]。また永里ら(2020)は入力変数の物理的な特性を考慮したうえで適切な入力変数を選択することの重要性を指摘した [2]。これらの研究では、入力データとして大気の再解析データのみ使用しているが、その解像度は地域レベルの現象を捉えるには十分とはいえない。そこで、高解像度な地形データを入力データに用いることにより、降水量推定制度の向上を図る。本研究においても、深層学習手法として CNN を使用する。そして、検証のために、地形データの有無における降水量の推定精度の比較を行う。

2. 手法

CNN は画像認識の分野で開発された深層学習手法の1つである。特徴として、畳み込み層やプーリング層を持った構造をしている(図-1)。畳み込み層は、画像の特徴の抽出を行い、プーリング層では、抽出されたデータの圧縮を行っている。

3. 対象地

対象地は、熊本県白川流域(図-2)を対象地に選定した。白川は、流域面積 480km²、流路延長 74km であり、阿蘇山根子岳を水源とする一級河川である。急流となって田園の広がる中流部を流下し、下流部においては、熊本市の密集市街地の中心を貫流し、有明海に流れ込

む特徴を持つ。

4. データセット

本研究では、入力データに大気の再解析データと地形データ、対象データに日単位の降水量データを用いた。大気の再解析データは、European Centre for Medium Range Weather Forecasts Interim Reanalysis(Era-interim)から取得した。Era-interim は、1979 年から現在までの全球レベルの大気情報の再解析データセットであり、時間解像度は 6 時間、水平解像度は 0.75 度×0.75 度である。地形データは、GTOPO30 から取得した。GTOPO30 は、水平グリッド間隔が 1km のグローバルデジタル標高モデル(DEM)である。降水量データは、Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources(APHRODITE)[3]から取得した。APHRODITE は、観測データを元に作成された日降水量データセットであり、0.05 度×0.05 度の解像度を持つ。

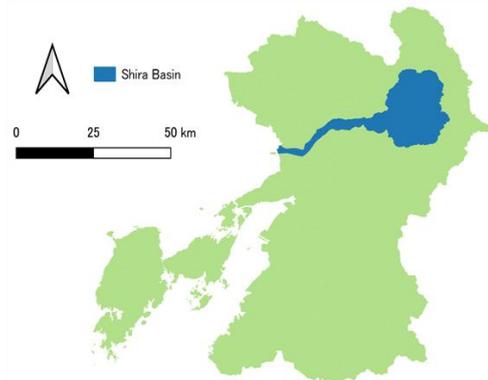
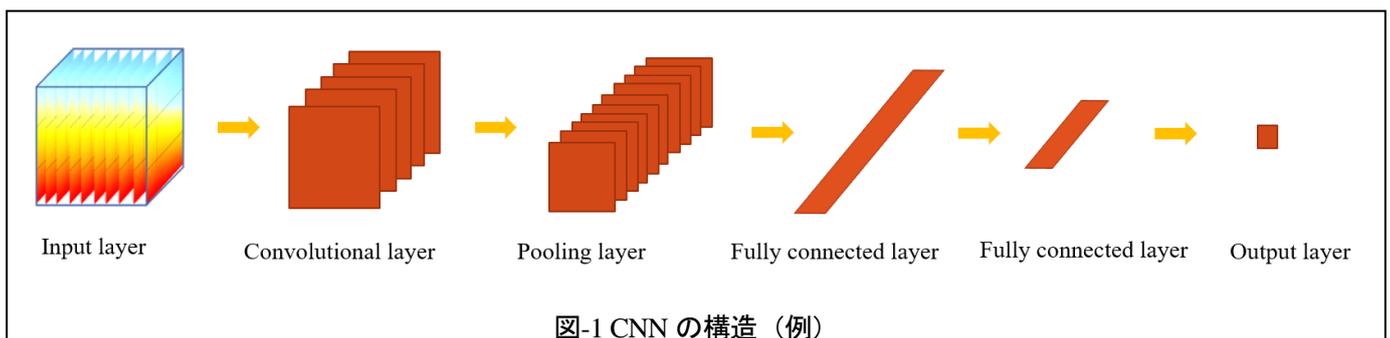


図-2 白川流域



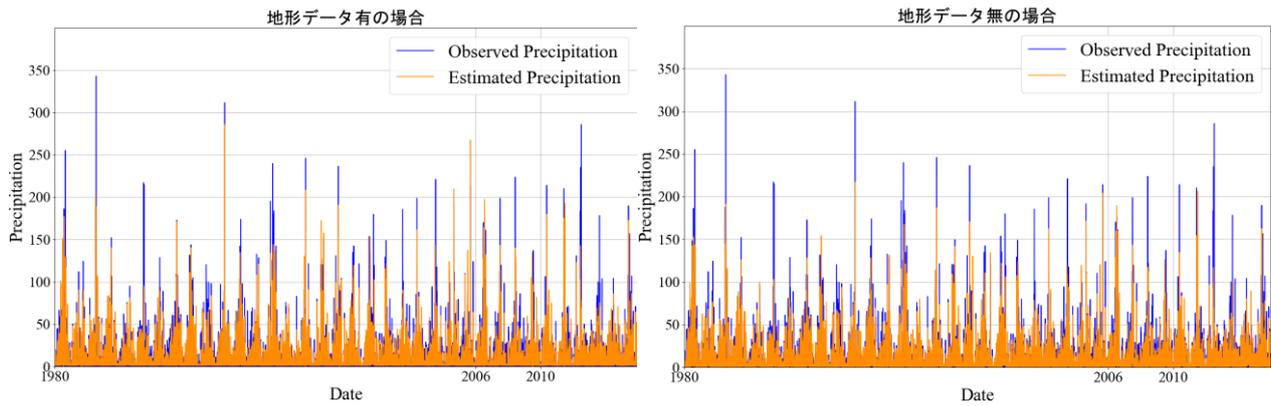


図-3 地形データを用いた場合と用いない場合における推定量-観測量の比較

表-1 HYPER PARAMETERS

ハイパーパラメーター	
バッチサイズ	512
バッチ選択	Shuffle 法
誤差関数	平均二乗差誤差
最適化関数	RMSprop
カーネルサイズ	3
ストライド	1
活性化関数	Softplus 関数
patience	40

5. 適用

まず初めに、GTOPO30 と ERA-interim のデータを緯度経度を指定することで同じ範囲に設定した。また、解像度が異なるため、それぞれの配列を線形に等間隔なものに生成し直し、データの補間を行ったものを入力データとした。CNN には、様々なパラメータがあり、用いるハイパーパラメーターは、表-1 に示す。モデルの学習・評価期間は、トレーニング期間(モデルの学習に用いる期間)を 1980 年～2005 年、検証期間(学習データに対し過学習を起こしていないか検証する期間)を 2006 年～2010 年、テスト期間(最終的なモデルの検証に用いる期間)を 2011 年～2015 年のそれぞれ 3 つに分割した。CNN 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 200 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。

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

Q_o : 観測流量 Q_s : 推定流量

6. 結果と考察

地形データである GTOPO30 を入力データとして追加の有無における推定量-観測量の比較した結果を図-3 に示す。図-3 をみると、地形データを用いる場合の方がピーク時において、実際の観測量により近い推定量を表している。しかし、両方に大きな差はなく、テスト期間における NSE の値も地形データを用いない場合は 0.660、地形データを用いる場合は 0.687 とこちらも差があるとは言えない。そのため、数値地形モデルを加えるだけでは降水量精度における影響は少ないと考えられる。また、本研究では ERA-interim から取得する気温や風速など標高に関する変数に地形データを用いた補正を行っていない。そのため、精度向上に影響が少ないと考える。

7. 結論

本研究では、深層学習手法の 1 つである CNN を用いた降水量ダウンスケーリングにおける地形データを用いた場合の降水量精度の影響について検証した。結果より、地形データを用いる場合でも、降水量精度に与える影響は少ないと言える。しかし、標高と関係のある入力変数に補正をかけることによって、より詳しい入力データを取得することができ、降水量精度の向上に繋がる可能性があると考えられる。

8. 参考文献

[1]板知明,芳村圭,深層学習を用いた水文気象場のダウンスケーリング手法の開発,土木学会論文集 B1(水工学),74(4),I_151-I_156,2018
 [2]永起義,石田桂,上田誠,横尾和樹,木山真人,尼崎太樹,深層学習を用いた水文気象場降水量ダウンスケーリングの特性,土木学会論文集 B1(水工学),Vol.76,No2,I_373-I_378,2020
 [3] Kamguchi, K . Arakawa, O . Kitoh, A . Yatagai, A . Hamada, A . Yasutomi, N. Development of APHRO_JP, the first Japanese high-resolution daily precipitation product for more than 100 years. Hydrological Research Letters, 4: 60-64 2010