

人工知能技術の河川ハイドログラフ予測への適用に関する研究

長岡技術科学大学院 環境社会基盤工学専攻 非会員 ○久々江耀平
長岡技術科学大学院 環境社会基盤工学専攻 正会員 細山田得三

1. はじめに

令和元年 10 月に発生した台風 19 号をはじめとして近年の日本では台風や大雨による洪水や土砂災害が多発している。こうした洪水対策施設の能力を超過した災害からの避難には、水文流出解析等による河川水位の予測情報が重要な指標となる。既知の予測手法は、流域ごとに細かいパラメータの調整や、地形データを与える必要がある。正確な予測には、地形データの更新等が必要となるが、その工程には多くの時間と労力を要する。

こうした水文流出解析の性質を受け、細かいパラメータ調整が不要ながら種々の予測において良好な結果を残している AI 技術¹⁾に着目した。

本研究では、種々の AI 技術のうち、長岡を対象に過去の降雨量、河川流量といった長期間の時系列データから河川流量予測が可能である LSTM (Long Short-Term Memory) モデルを用いた。LSTM は近年になり注目され、他分野では既に広く用いられ良好な結果を残している。この LSTM モデルの適用性を検討に際し、予測精度向上のため元データ群から一つの時系列データとして切り出す期間と学習期間の最適な程度を検討した。

また、既に河川水位予測に用いられている手法として粒子フィルタリングが挙げられる。この粒子フィルタモデルは、LSTM モデルと同様に過去のデータから回帰的に未来のデータを予測する。このことから既知の粒子フィルタモデルを実装し、2018 年の河川流量に対して逐次近似による予測を行い、予測結果と観測値の比較を 2 乗平均平方根誤差により示し、両モデルの精度の比較検討を行った。

以上より本稿は LSTM モデルの河川ハイドログラフ予測への適用性を検討するものである。

2. LSTM (Long Short-Term Memory) モデル

本章では、LSTM モデルのアルゴリズムを解説する。

なお、1987 年 1 月 1 日から 2019 年 12 月 31 日の観測地点長岡における河川流量データ (国土交通省北陸地方整備局信濃川河川事務所によるものを借用) と 1 時間降雨量データ (気象庁より借用) のうち、2017 年から過去 30 年間のデータから LSTM モデルを実装した。そのモデルに 2018 年 1 月 1 日から 2018 年 12 月 31 日までの降雨量を入力し、各時系列データの末尾の時刻から 1 時間後・12 時間後の流量を予測した。

(1) アルゴリズム

AI 技術と総称されるもののうち、NN (Neural Network) はヒトの脳の学習過程に着想を得て構築された計算アルゴリズムであり、いくつかの問いと答えの組み合わせから学習し、新たな問いに対して学習に基づいた答えを導き出す。NN は大きく入力層・隠れ層・出力層に分かれたニューロンから成る。

RNN (Recurrent Neural Network) は NN に時間軸の概念を導入したアルゴリズムであり、入力値・出力値を時系列データとして扱う。時系列データとは、データ群において時間軸方向の依存があるデータである。しかし、RNN は長期間予測が不可能なアルゴリズムであった。

LSTM は長期間予測のためのアルゴリズムとして開発され、RNN からの改良点として図-1 に示すように隠れ層を LSTM ブロックに置き換えている。RNN を構成する通常のニューロンは一様な機能を持つが、LSTM ブロックのニューロンはそれぞれ固有の役割を持つ CEC (Constant Error Carousel)・入力ゲート・出力ゲート・忘却ゲートから構成されている。入力ゲート・出力ゲートは時間依存性のあるデータを受け取った場合のみ CEC へと誤差を伝播する。CEC は誤差を記憶するニューロンである。忘却ゲートは時系列データ内でパターンが変わった場合、CEC から過去の情報を捨てるためのニューロンである。これらの LSTM ブロックにより長期間の予測が可能となった。

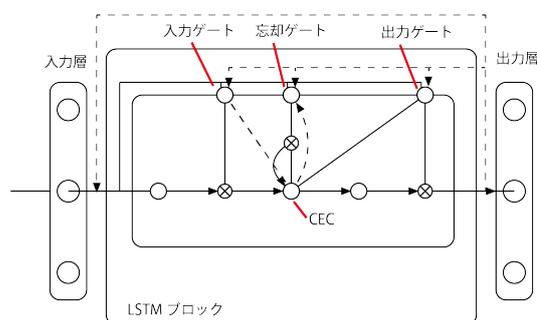


図-1 LSTM の展開図

3. 結果と考察

本章では対象地点での予測結果を示し、考察する。

ここで全期間とは 2018 年 1 月 1 日から 2018 年 12 月 31 日、融雪期間は同年 4 月 2 日から 4 月 9 日である。

表-1 各切り出し期間における2乗平均平方根誤差

切り出し期間	全期間		融雪期間	
	1時間後	12時間後	1時間後	12時間後
2日間	57.66	150.4	80.23	189.7
3日間	54.78	151.5	73.34	172.6
1週間	56.05	160.2	83.13	182.6
2週間	62.49	173.1	79.28	198.8
1ヶ月	70.17	175.2	85.82	221.7

表-2 各学習期間における2乗平均平方根誤差

学習期間	全期間		融雪期間	
	1時間後	12時間後	1時間後	12時間後
1年間	54.78	151.5	73.34	172.6
10年間	54.29	146.5	76.45	171.5
20年間	53.10	143.7	71.38	175.6
30年間	63.59	150.4	95.27	180.6

表-3 両モデルの実測値・予測値間の2乗平均平方根誤差

Model	全期間		融雪期間	
	1時間後	12時間後	1時間後	12時間後
LSTM	54.78	151.5	73.34	172.6
粒子フィルタ	60.15	165.8	93.31	210.6

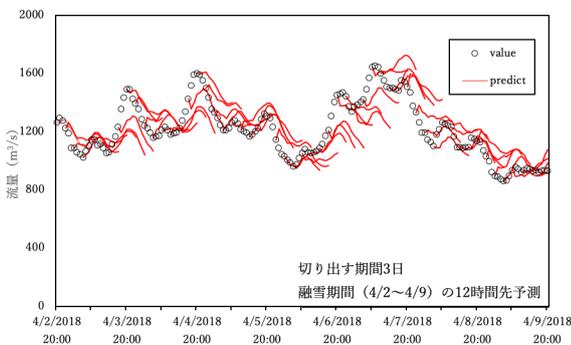


図-2 切り出し期間3日間の融雪期間の予測結果

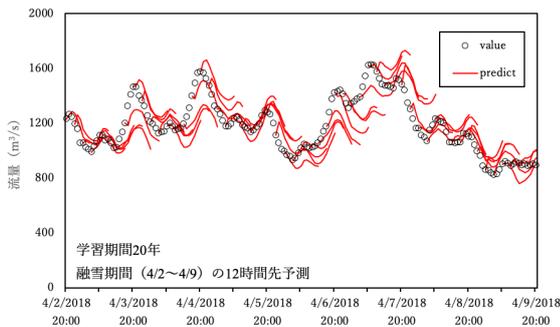


図-3 学習期間20年間の融雪期間の予測結果

(1) LSTMモデルによる予測結果

まず、学習期間を1年間とし、最適な切り出し期間を検討する。代表例として図-2に時系列データの切り出し期間を3日間とした際の融雪期間の予測値と実測値を示す。表-1に各切り出し期間の時系列データから予測した流量と実測値との2乗平均平方根誤差 (m^3/s) を示す。表-1より最も精度の高い予測は切り出し期間を3日間とした時系列データによるものである。また、融雪期間の精度は通年の予測精度より劣る。これは降雨量以外の要因による河川流量の増加をモデルに反映できていないためである。

次に時系列データの切り出し期間を3日間とし、モデルの学習期間の検討を行った。代表例として図-3に学習期間を20年間とした際の融雪期間の予測値と実測値を示す。表-2に各切り出し期間の時系列データから予測した流量と実測値との2乗平均平方根誤差 (m^3/s) を示す。表-2より20年間に設定した時最も精度が高くなるが、30年間では精度が低くなる。このことから30年以上の学習期間では過学習であり、20年間程度が最も適した学習期間といえる。

また、図-2、図-3より直近の観測値(流量・降雨量)が大きいと過大に予測し、観測値が小さい場合は過小に予測する特徴がある。

(2) 粒子フィルタモデルとの予測結果比較

表-3に各モデルにより(LSTMモデルは切り出し期間3日間、学習期間1年間)予測した流量と実測値との2乗平均平方根誤差 (m^3/s) を示す。表-3よりLSTMモデルの2乗平均平方根誤差は粒子フィルタモデルよりも小さい。LSTMモデルは既知の手法よりも同等以上の精度を持つといえる。

4. まとめ

本研究では、過去の降雨量データと河川流量データの関係を学習した後、入力された降雨量から河川流量を予測するLSTMモデルを実装した。また、適切な実装条件の検討を行いモデルの精度向上を図った結論として次のことが得られた。

- LSTMモデルによる河川流量予測は切り出し期間3日間、学習期間20年間程度が適している。
- LSTMモデルは粒子フィルタより予測精度が高い。
- 降雨以外の要因による変化を予測できていない。

以上から改良の必要はあるもののLSTMの河川ハイドログラフ予測への適用性が示された。また、今後の展望として次のことが挙げられる。

- LSTMモデルを双方向化による精度向上、self attention法によりモデルの予測理由を可視化する。
- 台風発生時の降雨量・河川流量にクローズアップし、d4PDF等から台風来襲時の河川流量の将来実験を行う。

参考文献

- 1) 吉田均, 野村泰稔, 広兼道幸, 一言正之, 小田和広, 秋山孝正, 宇津木真司: AIのインフラ分野への応用, 電気書院, pp63-81, 2019.
- 2) 巢籠 悠輔: 詳解ディープラーニング 第2版, マイナビ出版, 2019.
- 3) 新保聖志, 志野直紀: 大流域における粒子フィルタ法を用いた洪水予測システムの実用化について, 2017