

深層学習を用いたダムの貯水量予測

中央大学 学生会員 ○八木 将 中央大学 正会員 佐藤 尚次

1. はじめに

日本の河川は「狭い・急勾配・短い」といった特徴があるため、降った雨が一気に海へ流れてしまう。そのため洪水が多いことはよく知られているが、それらの日本の河川の特徴は同時に渇水が起こりやすい原因となっている。

日本の年間降水量は長期的にみると減少傾向にある。さらに、近年の異常気象の増加から水資源の安定した配給はより難しくなっている。洪水に比べて渇水は重要視されていないが、水資源問題は家庭用水・都市活動用水・工業用水・農業用水と広範囲の人々の生活に影響を及ぼすため軽視してよい問題ではない。

2. 研究の目的

渇水対策となるダムは洪水・渇水どちらの災害の対策ともなる施設である。しかし、それぞれの現象の特徴とダムの機能から分かるように、この先どちらの現象が起こるかを予測できないと災害に対する正しい貯水量がわからず、ダムがあるにもかかわらず災害を起こしてしまうことになる。実際、去年の7月初めに起こった愛媛県での洪水の際には降雨量が多くダムが耐え切れず放流し事故が発生した。さらに、2016年の6月から計5河川で行われた利根川水系での取水制限は、5月からの過剰放流が原因であると非難を浴びたこともある。

そこで、ダムのより良い利用のために NN(Neural Networks)を構築し水資源量予想を行う。

3. 対象とする地域

本研究で対象とする地域は、国内で渇水の被害を最も受けている四国地方の吉野川流域とした。

四国地方は、深刻な取水制限が長期間続くような渇水が高頻度で起きており、国内で数少ない慢性的に水不足である地域である。

長年渇水に苦しめられておりダムの数も多いため利用可能なデータは豊富である点と、積雪が少なく他地方に比べて水資源量の推定が容易である点が機械学習の対象として適している。

また、吉野川は水流が四国4県に及ぶ水系であり四国内で流域面積が最も広く流域内人口も多い。加えて、香川県民の水源とも呼べる香川用水の水源である早明浦ダム・池田ダムも流域内にあるため、四国で一番の重要河川といえる。その吉野川は渇水リスクを頻度と災害規模で考えると日本で最も渇水リスクが高い河川だといえる。

表-1 構築したプログラムの概要

Train:Test		5000:3000
ニューロン数	入力層	21変数
	隠れ層	1000
	出力層	1
LSTM層の活性化関数		tanh
再起計算用活性化関数		tanh
出力層の活性化関数		linear
最適化アルゴリズム		adadelta
予測一回分のデータ数		730(日分)
epoch(繰り返し計算数)		1000
ミニバッチサイズ		100
損失関数	Mean Abs Error	
metrics	loss, MeanAbsError, MeanSquareError	

4. 研究手法

本研究では日毎の吉野川の河口部・下流域・中流域にある3地点の水位と本流の早明浦ダム・池田ダム、支流である銅山川の宮川内ダム・柳瀬ダム・新宮ダム・富郷ダムの貯水量、雨量データ10地点、日付、取水制限レベル、早明浦ダムの放流量・流入量の計23変数を入力データとし7日後の各ダムの総貯水量の予測を出力する。構築する NN モデルは、時系列データを扱うことに長けている再起型ネットワーク (Recurrent Neural Networks), そのなかでも利用例が多く比較的新しいネットワークモデルである長短期記憶 (Long Short-Term Memory)を選んだ。NN モデルは Tensorflow をバックエンドとした Keras を利用し構築した。RNN の特性上、学習データは3次元データ(分割データ数, 予測1回分のデータ数, 変数の数)に変換してインプットする。

入力変数の考慮日数は最適化中であるが季節性を考慮したいので現在は2年分の730日とかなり長いスパンを現在はとっており、活性化関数は tanh, sigmoid, relu, linear から組み合わせを試し tanh, tanh, linear の組み合わせに決定した。損失関数は回帰分析であり外れ値を多くは含まないデータのため平均絶対誤差(Mean Absolute Error)を選び、最適化

アルゴリズムは確率的勾配降下法(SGD), Adagrad, RMSPROP, Adam, AdaDelta をそれぞれ試し AdaDelta を選んだ. 評価関数(metrics)は平均絶対誤差(Mean Absolute Error), 平均二乗誤差(Mean Squared Error), 誤差(Loss)を表示し学習曲線を表示させることで学習の推移と過学習の有無を確認する.

5. 結果の分析

モデルの評価として約 24 年分の毎日のデータ 8000 個のうち, Train データ 5000, Test and Validation データ 3000 から各評価関数を算出しグラフ化した. 各評価関数の推移を確認すると順調に学習は進み過学習も起きておらず, train_data と validation_data の誤差はどちらも近い値に収束している為, 両データの偏りも小さいかと考えられる.

しかし, 予測値と正解値を見比べてみるとまだ学習が不十分で予測がうまくできていないことがわかる. これは損失関数の視点では誤差が十分小さいと判断し計算が収束に向かってしまったせいか, あるいは現状の条件ではこの結果が計算限界であることが考えられる. よって, 今後もネットワークのさらなる最適化が必要である.

予測結果は渇水イベントに対しておおむね予測できているものと、貯水量の低下を全く終えていないものがある。これらの違いは入力データである雨量等を確認してもわからなかったため、予測に必要なデータが入力できていないと考え新たなデータを入力する必要がある。ちなみに、渇水を長期的に予測するシステムとして学習させているので、台風などによる急激な貯水量の上昇は全く予測できておらず、渇水後での貯水量の立ち上がりも入力データとの時間差である 7 日分遅れている。

6. おわりに

過学習の経過を観察するために学習曲線を出力していたが過学習の傾向がみられるまで計算を進めたほうが細かい点まで予測できている面があり評価指標も MAE だけでなく評価する必要がある。

急激な変化でない長期的な貯水量の低下を終えていないのが現状では学習不足か入力データ不足なの

かわかっていないため、その点を解明し改善する必要がある。現状で考えている入力データの候補は、気温やそれを利用した蒸発散量やさらに多くの治水ダムや各種用水用の貯水池の貯水量などを入力として利用する予定である。しかし、増やすだけでは計算時間や計算結果に悪影響を与えるので、入力データの取捨選択をする必要もある。

また、渇水イベントに焦点を絞った予測システムを構築しているので学習データを全データとせず渇水イベントに絞った。データで学習・予測をすることも考えている。

謝辞

本研究では、国土交通省四国地方整備局 吉野川ダム統合管理事務所、水資源機構池田総合管理所 第一管理課、徳島県水管理政策課、より貴重なデータを使用させていただきました。ここに記して謝意を表します。

参考文献・出典

- 1) 土屋 大 藤村 和正 :早明浦ダム流域における長期流出解析と渇水時の流入量予測, 第 35 回土木学会関東支部技術研究発表会, 2007
- 2) 国土交通省四国地方整備局
<http://www.skr.mlit.go.jp/yoshino/> 2021/01/14
- 3) 国土交通省水門水質データベース
<http://www1.river.go.jp/> 2021/01/14
- 4) 国土交通省ダム諸量データベース
<http://mudam.nilim.go.jp/home> 2021/01/14
- 5) 我妻 幸長, はじめてのディープラーニング, SBクリエイティブ株式会社, 2018
- 6) 巢籠悠輔, 詳解ディープラーニング TensorFlow/Keras・Pytorch による時系列データ処理, 株式会社マイナビ出版, 2019
- 7) Tensorflow
<https://www.tensorflow.org/> 2021/01/14
- 8) 国土交通省四国地方整備局 吉野川ダム統合管理事務所
- 9) 水資源機構池田総合管理所 第一管理課
- 10) 徳島県水管理政策課

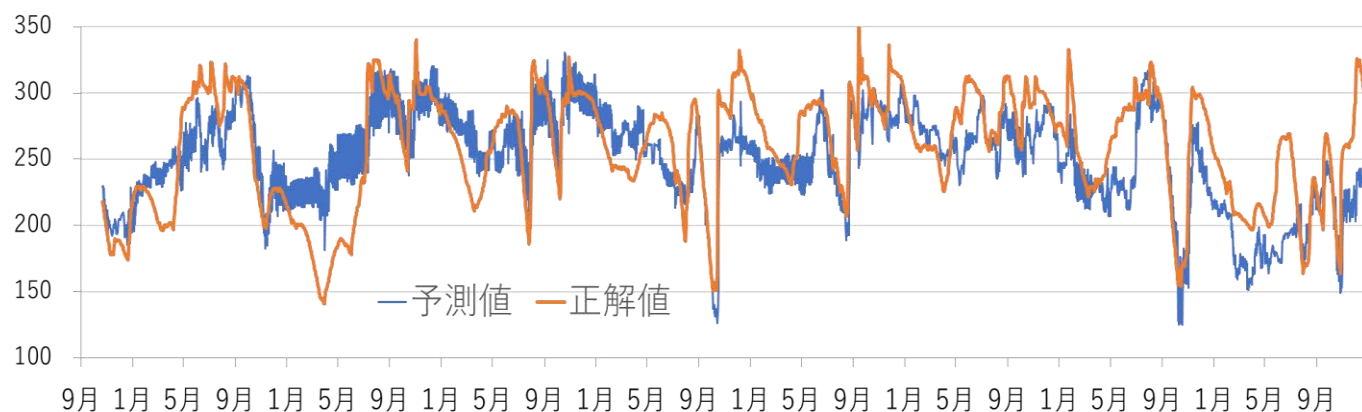


図-1 予測値と正解値の比較