

ニューラルネットワークを用いた土器川流域における 河川流量予測の入力データ項目の検討

香川大学工学部	賛助会員	石橋日菜子
香川大学工学部	賛助会員	柴 尾 葵
香川大学工学部	賛助会員	宮本えりか
香川大学工学部	賛助会員	吉 原 航 平
香川大学創造工学部	正 会 員	石 塚 正 秀

1. はじめに

近年、人工知能（AI）が注目されており、あらゆる分野で利用されている。AI 技術は水工学分野でも活用されており（一言ら（2016）、中谷ら（2017））、例えば、滝口ら（2018）は、ニューラルネットワーク（Artificial Neural Network：ANN）による河川流量予測モデルを積雪地域の奈良俣ダム流域（集水面積：95.4 km²）に適用し、ANN のパラメータ設定と入力データの選択について考察した。その結果、日平均気温、日降水量、最深積雪深、日平均風速の 4 つの因子が積雪地域の河川流量予測に重要である結果を示した。

本研究は、同様のモデルを用いて、河川流量の少ない土器川流域を対象として、河川流量予測に対して、どのような入力データが影響するのかについて考察を行った。

2. 研究の手法

(1) 計算方法

ANN の計算は、滝口ら（2018）にもとづいて実施した。パラメータの設置および計算手順の詳細についても同様に滝口ら（2018）を参考とした。

(2) 計算対象流域

香川県の一級河川である土器川（流域面積：127 km²）を対象とした。土器川は讃岐山脈を水源として、仲多度郡まんのう町および丸亀市を流れ、瀬戸内海に注ぐ。気象データには気象庁アメダスの滝宮観測所、河川流量には国土交通省水文水質データベースの祓川観測所のデータを使用した。

(3) 計算条件

全計算期間は、1993-2007 年の計 15 年間とした。用いる入力データは、日平均風速、日照時間、日降水量、日平均気温の 4 項目である。なお、計算の都合上、2 項目を dummy データとして、計 6 項目の入力データを用意した。また、教師信号は土器川（祓川橋地点）の日平均流量とした。

(4) 計算期間の設定

1993 年から 2003 年を学習期間とし、入力データと教師信号を用いて ANN の重みを決定する。つぎに、2004 年から 2005 年を試験期間として過学習を判断し、学習済みの ANN に入力データを与え、出力結果

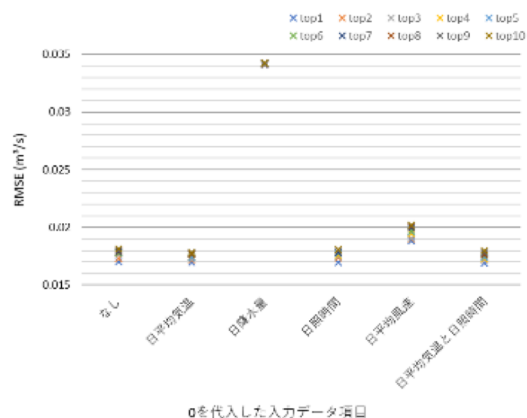


図-1 入力データの組合せ毎の RMSE の上位 10 位の比較 (左から、case 1～case 6)

と教師信号との RMSE（平均二乗誤差平方根）を算定する。そして、2006 年から 2007 年を検証期間とし、最適な ANN の入力パラメータを用いて河川流量予測を行い、その予測精度を RMSE によって評価する。

3. 研究の結果

(1) 入力データ項目と ANN パラメータ値の選定

試験期間における RMSE の値が小さい上位 10 個の ANN を入力データの個数ごとに選択し、パラメータの分布状況を調べた（図-1）。なお、入力データの種類は全部で 6 通りあるため、計 60 個の ANN が選択されたことになる。試験期間（2004-2005 年）に行った最適な入力データ項目を判定するために、図の横軸には、0 を代入したデータ項目を示し、case 毎に RMSE の値が小さい上位 10 個の値をプロットした。RMSE の値が小さいほどそのパラメータの寄与度が低く、大きいほど高いことから、日降水量は寄与度がかなり高い結果が得られた。また、日平均風速が 0 のときも RMSE がわずかに大きくなっていることから日平均気温や日照時間よりも寄与度が高いといえる。

また、ANN のパラメータ値について、Learning Rate は 0.5、Number of Epochs は 300、500、Batch Size は 7、15 のようにより小さい値、Neuron Size は 54、36 のようにより大きく設定した条件において、試験期間における予測精度が高いことが分かった。

表-1 入力データ項目ごとの RMSE の比較

	tem p	ra in	sun	w ind	RM SE	Rank
case1	○	○	○	○	0.0071	4
case2	×	○	○	○	0.0067	2
case3	○	×	○	○	0.0109	6
case4	○	○	×	○	0.0069	3
case5	○	○	○	×	0.0075	5
case6	×	○	×	○	0.0067	1

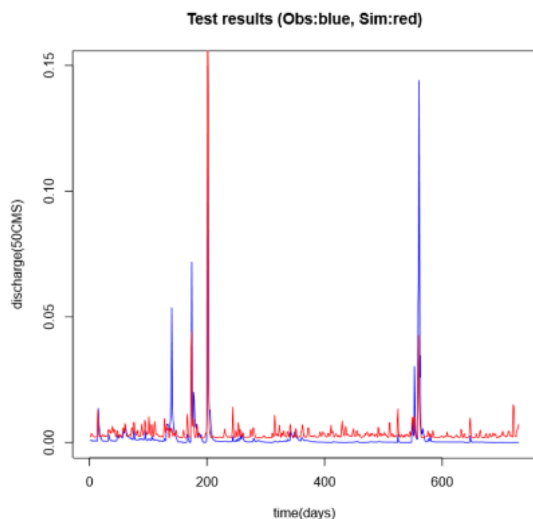


図-2 日降水量と日平均風速を考慮した場合 (case 6) の観測値と予測値の日平均流量の比較 (青線が観測値、赤線が予測値) (横軸は日付)

(2) 河川流量予測の検証

入力データ項目については 6 通りの組合せで検証を行った。ANN のパラメータ値については、先ほどの結果から Learning Rate を 0.5、Number of Epochs を 300、Batch Size を 7、Neuron Size を 54 に設定した。入力データ項目ごとの RMSE の値を表-1 に示す。まず、先ほどの試験期間で算出された RMSE が 0.016 から 0.035 だったのに対して、今回行った検証期間で算出された RMSE は 0.006 から 0.010 と小さくなっていることから、全体的に予測精度がより高くなっていることが分かる。また、試験期間と同様に、日降水量を 0 にした case 3 の RMSE が最も大きくなり、ついで日平均風速を 0 にした case 5 の場合が大きくなる結果が得られた。

また、日平均気温、日照時間をそれぞれ 0 にした case 2 と case 4、日平均気温と日照時間をともに 0 にした case 6 の RMSE がより小さくなっていることから、土器川流域において河川流量予測に用いる入力データ項目は、日降水量と日平均風速が重要であるといえる。

(3) 観測値と予測値の比較

図 2、3 は、流量ハイドログラフについて、それぞれ case 6、case 3 の観測値と予測値の比較を示す。降水量と風速を考慮した case 6 では、洪水時における青と赤の線が重なっていることから、精度が高いといえる。一方、降水量を考慮しなかった case 3 では、青と赤の線が重なっていないため、精度が悪いといえる。

Test results (Obs:blue, Sim:red)

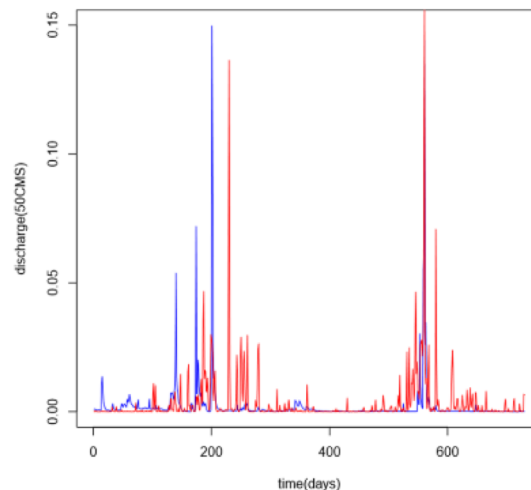


図-3 日降水量を考慮しない場合 (case 3) の観測値と予測値の日平均流量の比較 (青線が観測値、赤線が予測値) (横軸は日付)

4. まとめと今後の課題

以上の結果から、土器川流域において最適な入力データ項目は日降水量と日平均風速であることが分かった。日平均気温と日照時間については、考慮しないとき (case 6) の方が予測精度は高いことから、河川流量予測においては入力データとしての重要性が低いことが分かった。ただし、case 6 の RMSE の値は 0.0067 であり、0 ではなかったことから、これらの他に入力データとして重要な項目があるのではないかと考える。他の入力データとして考えられるものの例として、相対湿度、水温、気圧などが挙げられる。

ANN のパラメータの設定においては、今回は各 3 パターンの値で検討したが、パターン数を増やすことでより最適なパラメータ値を抽出することができれば、さらにより小さい RMSE を算出することができると考えられる。また、学習期間を長くすることで、様々な出水現象を学習することができることから、より精度の高い予測が行えるのではないかと考える。

参考文献：

- 1) 一言正之、櫻庭雅明、清雄一：深層学習を用いた河川水位予測手法の開発、土木学会論文集 B1 (水工学)、72、No.4、I_187-I_192、2016.
- 2) 中谷祐介、石崎裕大、西田修三：深層学習を用いた感潮河川の水質変動予測、土木学会論文集 B1 (水工学)、73、No.4、I_1141-I_1146、2017.
- 3) 滝口修司、キムスンミン、立川康人、市川温、萬和明：ニューラルネットワークを用いた積雪地域の河川流量予測における重要入力因子の抽出、土木学会論文集 (水工学)、74、No.4、I_877-I_882、2018.
- 4) 気象庁：アメダスデータ、www.jma.go.jp/jp/amedas/
- 5) 国土交通省：水文水質データベース <http://www1.river.go.jp/>