

深層学習の河川防災・環境分野への応用 —四万十川・鏡川水位、仁淀川大渡ダム流入量について—

中根 英昭¹ 若槻 祐貴²

山本 啓² 武田 拓巳² 端野 典平^{3*}

(受領日：2019年5月7日)

¹ 高知工科大学名誉教授

〒782-8502 高知県香美市土佐山田町宮ノ口185

² 高知工科大学大学院工学研究科環境数理コース

〒782-8502 高知県香美市土佐山田町宮ノ口185

³ 高知工科大学環境理工学群

〒782-8502 高知県香美市土佐山田町宮ノ口185

* E-mail: hashino.tempei@kochi-tech.ac.jp

要約：四万十川、鏡川、仁淀川を対象に、河川防災、河川環境保全にとって重要な、水位、ダム流入量を推定・予測するための深層学習を用いたモデリングを行った。使用したモデルは全結合層によって構成され、隠れ層が2~4個の多層パーセプトロンであり、使用した計算機は、GPUを搭載したWindows PCあるいはLinux PCである。全てのモデリングにおいて、河川水位あるいはダム流入量測定地点の上流流域圏の雨量時系列を深層学習の入力データとして用いた。そして、時系列の長さが数時間~数10時間の場合よりも、1~2年の方が観測データをより高精度で再現できた。特に、流域圏への降雨が少なく、低水位、低流量の時期に時系列の長いことによる精度の改善が著しかった。短い時系列の場合には表流水の情報のみが学習されているのに対し、長い時系列の場合には、表流水、土壌水分、地下水・湧水の情報も学習していることが理由として考えられる。中小河川である鏡川の場合には、上流の観測雨量を用いた水位予測では1時間前の予測は可能であるが、2時間以上の早期水位予測のためには気象庁の降水短時間予報の利用が必要であることが示唆された。

1. はじめに

深層学習（ディープラーニング）は、近年の人工知能の急速な発展にとって鍵となる技術である¹⁾。高知工科大学環境理工学群中根研究室では、2016年度に深層学習研究の環境構築を開始し、2017年度から本格的に深層学習を用いた研究に取り組んできた。2017年度までの取り組みについては、高知工科大学紀要15巻に、報告「環境分野への深層

学習応用研究の立ち上げについて」として収録・公開されている²⁾。また、中根は、深層学習が汎用的かつ柔軟で高い帰納的モデリング能力を持っていることに着目し、学生の教育、卒業後の仕事や社会貢献にとっての重要性、及び社会への正負のインパクトの大きさについての当時の考察をまとめた。これは、高知工科大学紀要15巻に、招待論文「環境分野の深層学習研究を通じた人材育成」として収録・公開されている³⁾。2018年度には前年度に引き続き、

修士課程1年生1名、4年生5名が深層学習を用いた研究を実施した。2018年11月からは、新たに着任した端野が研究に参加し、特に、当時修士課程1年生の若槻の指導を分担した。

若槻は、ダム・堰の影響が比較的少なく、地上観測による流域圏の雨量や水位データが十分に得られる一級河川として、四万十川を対象に、上流の雨量時系列データのみを入力データとして用いて、雨量時系列の長さが水位推定モデルの能力に及ぼす影響を中心に研究を行った。

4年生の山本は、高知市の防災にとって重要な二級河川（中小河川）である鏡川について、気象庁が作成している1kmメッシュの雨量データである「解析雨量」から得た時系列を入力データとして用いて水位推定・予測モデルを作成し、その能力を評価した。

4年生の武田は、透明度が高いことで有名な仁淀川において、洪水防止に大きな役割を果たしている大渡ダムの流入量に関して、上流の雨量観測所で観測した雨量の時系列を入力データとして用いて推定・予測モデルを作成し、その能力を評価した。

これら3つの研究に共通している特長は以下の2点である。

- (1) 入力データとして雨量のみを用いているため、6時間あるいは15時間までの予報が公開されている気象庁の降水短時間予報を用いて、大河川に対しても中小河川に対しても、防災準備に必要な時間的に余裕を持った水位やダム流入量の予測を可能にするモデリング手法であること。
- (2) 入力データとして、1年あるいはそれ以上遡る長期雨量時系列を用いることによって、表流水、土壌水分、地下水・湧水の情報を学習して深層学習モデルに取り込んでいること。

これらの特長によって、深層学習の中でも最も単純な手法である、全結合層を用いた多層パーセプトロンも、河川防災にとって実用的な水準のモデリングを可能にする手法であるとの結果が得られた。また、(2)の特長は、防災のみならず、生態系や水利にとって重要な、低水量、低水位の期間の推定・予測をも、同じモデリング手法を用いて行うことを可能にしている。このように、公開されているデータと単純な深層学習手法、及び安価なパーソナルコンピュータによって河川水位やダム流入量の推定・予測を可能にしたことによって、この三つの研究の成果は、市民や行政担当者自らが深層学習による河川防災、河川環境保護のためのモデル開発を可

能にすることにつながると考えている。2018年度に行われたこれらの研究の詳細を報告すると共に、今後の課題を示すことが、本報告の目的である。

本報告では共著者である若槻の修士課程1年目、山本、武田の卒業研究に焦点を絞って報告するが、2018年度の中根研究室の4年生の卒業研究では、深層学習を用いた下記の研究も実施された。

- 清水公貴：「変分オートエンコーダを用いたMNIST画像の2次元潜在変数空間へのマッピングとカテゴリ識別手法」
- 谷桃花：「CNN（Xception）と転移学習を用いたサクラの種類の判別」
- 宮崎美徳：「ディープラーニングによる早明浦ダムの流入量推定」

この3編の内、清水公貴君の卒業論文については、2019年7月31日に高知工科大学学術情報リポジトリにおいて全文が公開される予定である⁴⁾。

2. 長期雨量時系列を入力データとして用いた深層学習による四万十川水位の推定モデル

2.1 本研究に至る経緯

深層学習による河川水位やダム流入量の推定において、上流雨量の時系列を用いる試みは、中根研究室において2017年9月頃から吉野川上流の早明浦ダムの流入量推定において用いられた²⁾。この場合は、流入量観測時から96時間遡る1時間ごとの上流流域圏平均雨量データを用いた。データをランダムにシャッフルした後5分割し、5分割したそれぞれのデータを80%の訓練データと20%の検証データに更に分割して、流入量推定モデルの作成と検証を行った⁵⁾。すなわち、シャッフルに基づく5分割の「k分割交互検証法」⁶⁾を用いたことになる。5つの検証結果の平均値は、二乗平均平方根誤差（RMSE）によって表すと23.9m³/sであるが、入力データにダム貯水量を加えると19.6m³/sになった。このことは、96時間前以前の降雨の情報を間接的に入力データに加えたことになるが、観測流入量とモデル推定値の間の不一致の改善は、流入量の小さな場合ほど大きかった。

また、高知県の二級河川である安田川下流の観測水位に対して、96時間遡る流域圏平均雨量時系列データのみを入力とした場合と、96時間時系列データに1年遡る1か月平均雨量データを追加した場合では、後者がはるかに観測水位を再現すること、その改善の度合いは低水位のデータに対してはるかに大きいことが、文献²⁾の図6と図7、あるい

は文献³⁾の図5と図6の比較で示された。しかしながら、検証に用いた安田川東島観測所の降水の無い期間の水位が時間と共に頻繁に変化していることから、誤差評価の上で精度を欠く可能性があることが2018年度に入って明らかになった。

さらに、早明浦ダムの場合においても、安田川の場合においても検証における「情報の洩れ」の問題に対する対策が不十分であり、その影響は、5分割していない安田川の場合において大きいと考えられた。

そこで、2018年度には、上流の雨量時系列データを入力データとして、雨量データの時系列の長さが水位推定モデルの精度に及ぼす影響を検討することにした。まず、水位推定地点として四万十川中流域の津野川観測所を選択した。この観測所では、降水の無い期間に、水位の観測データが安定してほぼ同一の低い水位を示すからである。その検討結果を、「情報の洩れ」に関する考察と共に本節において示す。

2.2 先行研究と本研究の関係

ニューラルネットワークを用いた洪水時の河川水位予測は、2000年代から、Artificial Neural Network (ANN) と呼ばれる入力層1層、隠れ層1層、出力層1層の3層多層パーセプトロンを用いて行われてきた。海外では、例えば、2002年に発表されたBrath et al. によるイタリアのアルノ川における雨量からの流量予測⁷⁾、国内では、2003年に発表された八千代エンジニアリング等の研究グループによる阿武隈川洪水予測に関する研究⁸⁾などが典型的な例として挙げられる。これらのニューラルネットワークにおいては、隠れ層が1層しかない上に、隠れ層のノードの数がアルノ川の例では2個、阿武隈川の例では7個と少ないため、現在の2層以上の隠れ層と数100個以上のノード数を持つ多層パーセプトロン（深層学習のモデル）とは複雑な水文過程全体の表現力において大きな差がある。このため、阿武隈川の場合では、水位が平坦な部分と水位上昇～ピーク付近の二つの期間に分けて学習を行っている。

また、日本工営中央研究所の一言らは、2012年に発表した揖斐川を対象とした洪水予測の精度向上に関する論文において、数個の過去の観測雨量と観測水位を入力データとして用い、入力データの2倍の数の隠れ層のノード数の3層パーセプトロンによって、3時間までの洪水時の水位予測を数cmの誤差で行うことに成功した⁹⁾。さらに一言らは、2016年に、隠れ層が2層の多層パーセプトロン、す

なわち深層学習モデルを用いると、隠れ層が1層の場合に比べて洪水時の予測誤差が大幅に低減されることを示した¹⁰⁾。しかし、過去の観測水位を入力データとして用いていることが示しているように、ニューラルネットワークによって、降雨から水位上昇に至る水文過程全体をモデル化するという発想はなかったと考えられる。それは、一言らが2017年に発表した論文「深層ニューラルネットワークと分布型モデルを組み合わせたハイブリッド河川水位予測手法」¹¹⁾において、分布型モデルによって計算した水分貯留量を入力データに加えることによって、更なる精度の向上を図っていることから明らかである。

このように、2010年代に急速に発展した深層学習の特長である、深い層と大きなノード数による複雑なプロセスに対する優れた表現可能性とそれを可能にする技術—グラフィックプロセッシングユニット (GPU) による並列計算、活性化関数 ReLU、Adam 等の最適化手法、過学習を防ぐための正則化手法等—をフルに活用して地下への浸透と流出も含めた水文過程を、増水期も渇水期も含め年間を通して深層学習によってモデル化する試みは、著者らが知る限りわが国では存在しない。

海外においても、Chaopeng Shen によって2018年8月30日に“Water Resources Research”に出版されたレビュー、“A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists”¹²⁾の“3.1.3 In Hydrology”において、

“Compared to some other disciplines, hydrology has not witnessed the wide use of DL. Applications of DL have been few, especially in a big data setting, and the list of papers reviewed in this subsection is exhaustive to the author’s best compilation effort.”

と述べられている通り、2017年度から中根研究室で行っているような、入力ノードが数100から数万のビッグデータを扱って、深層学習のみによって水文過程をモデル化するような試みは、世界的に見ても、独立して行われた深層学習による水文過程モデリングの新しい研究の一つであると考えている。また、積雪の役割が重要でない場合に限るが、長期の雨量時系列データと単純な4～5層の多層パーセプトロンを用いて、年間を通した水文過程をモデル化できることを示した研究は、上記のレビューにおいても紹介されておらず、独自の特色を有する研究と考えている。

四万十川は、長くて緩やかな勾配の中流域が特徴の、本流にダムのない、幹川流路延長 196 km の一級河川である。入力データとして、芳生野、北川、梶原、影地、船戸、大野見、窪川、大正、日吉、近永、目黒、堂が森、江川崎の 13 地点の雨量の 1 時間値の時系列を用い、出力データ（教師データ）として、河口から約 40 km 上流に位置する津野川観測所の一時間毎の水位を用いた。13 の雨量観測地点と津野川観測所の位置を地図上に示す（図 1）。また、津野川観測所の水位の変化を図 2 に示す。冬に水位が低く、夏から秋にかけて毎年のように水防団待機水位に近い、あるいはそれを超える高水位が現れている。2014 年には、14.50 m の水位を記録している。

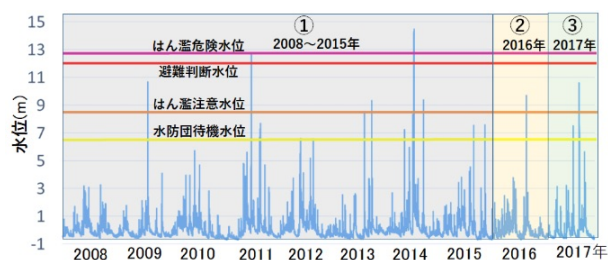
計算に用いた環境を表1に示す。使用した計算機は以前からデータ解析に使用していた Windows PC または OS を Linux(Ubuntu 17 等)に変更した PC にグラフィックボードを追加したものであり、地方大学の研究室でも数台用意することができ、市民が自費で1台所有できるような価格のものである。

学習モデルのプログラミングについては、“Keras Documentation”¹³⁾を参考にしてPythonを用いて行った。使用したモデルは全結合ネットワークの一種であるため sequential モデルを用いてプログラミングが可能であるが、今後の拡張性を考慮して Keras functional API を用いた。但し、後に述べる鏡川及び仁淀川大渡ダムに関する研究では Keras Sequential モデルを用いた。今後、プログラミングには functional API を用いる方向で教育・研究を進めている。

入力データについては、以下の考察に基づいて作成した。津野川観測所の水位には、上流約 150 km 圏の本流・支流からの雨水による表流水や地下水（湧水）が寄与している。河川に流出した後の雨水の多くは数時間～数 10 時間で観測所に到達すると考えられる。仮想的なパルス状の降雨に対して、降水が



図 1. 四万十川流域圏（薄黄色部分）と津野川観測所（水位観測地点）及びその上流の 13 の雨量観測地点。地図出典：国土交通省四国地方整備局中村河川事務所ホームページ，四万十川の雨量 <http://www.skr.mlit.go.jp/nakamura/ema22.html>



短時間で中下流に到達する場合にはパルス状に近い水位上昇を形成するが、地下に浸透して湧水として河川に流出する場合には、水位上昇に数か月から1年以上の時間遅れと共に、浸透する深さや地質により様々な広がりを持った水位上昇を形成すると考えられる。そこで、水位観測時間から遡る時間と共に徐々に時間分解能（雨量の平均時間）が大きくなるような雨量時系列データを入力とすることにした。具体的な時間分解能は下記の通りである。

- 230

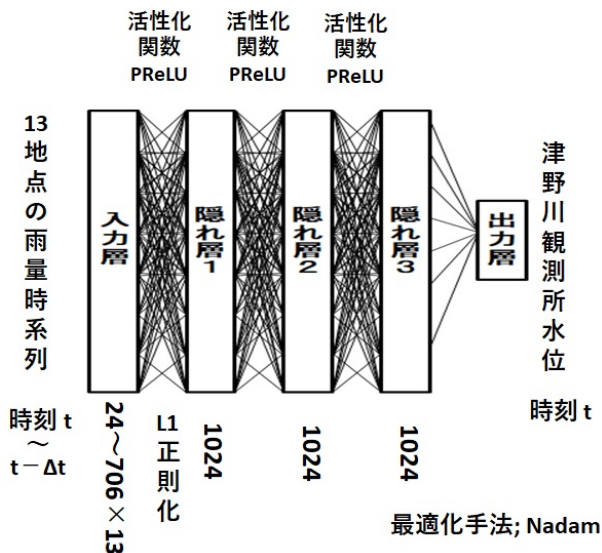


図3. 深層学習に用いた多層パーセプトロン。

「1024」等はノード数。2008年1月1日1時～2015年12月31日24時までの1時間ごと、70,128個の津野川観測所水位が出力層のデータ（教師データ）。水位観測の時刻 t から Δt 時間遡る上流13地点の雨量時系列が入力データ。例えば、 Δt が24時間の場合は時系列長=24。 Δt が2年の場合は時系列長=706。

表1. 深層学習に用いた計算機環境例

CPU	Core i7-7700
GPU	GeForce GTX1070 or GTX1080Ti
OS	Windows 10 / Linux (Ubuntu)
フレームワーク	Keras / Tensorflow
プログラム言語	Python 3.6 (Anaconda 4.4.0)

- 10日（153日+1時～720日）

図4は、遡り時間に対して分解能をプロットしたグラフである。例えば、水位観測の100日前の雨量は直前の132時間の平均雨量であり、642番目の時系列データとなる。

このようにして、17,280時間（720日）の時系列データを706要素の時系列に圧縮することにより、メモリや計算時間への負荷を軽減した。時系列の最長を約2年とした理由は、例えば、夏の渇水や洪水への前年の梅雨期や春雨の寄与等を表現するためである。水位を観測している津野川観測所と他の雨量観測所の間の距離の違いや支流ごとの傾斜や地質の違いの影響も想定されるため、上流観測点平

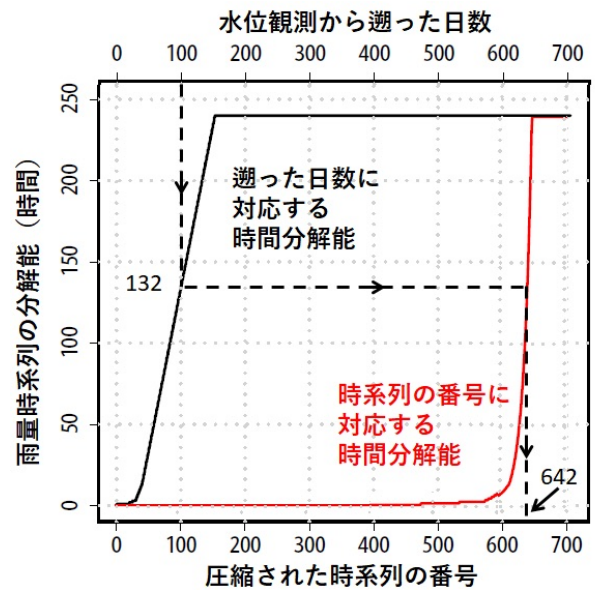


図4. 雨量時系列データの要素に対応する、水位観測時刻から遡った時間（日数）及び分解能

均雨量ではなく、全13地点の降雨の時系列を入力データとした。そのため、入力層のノード数は最大9,178（13×706）となった。

学習は、2008年1月1日1時～2015年12月31日24時までの1時間ごと、入力データと出力データ（教師データ）からなる70,128セットのデータをシャッフルして80%の教師データと20%の「検証データ」に分割して学習した。しかし、「検証誤差」が0.3%程度またはそれ以下という非現実的な値となったため、実際には上記の「20%の検証データ」は検証データとして扱わず、2016年1月1日1時～12月31日24時までのデータ（図2の②の期間のデータ）を検証データとして用いて、この検証誤差に基づいてハイパーパラメータ（隠れ層数、ノード数、正則化の係数、ドロップアウトの有無等）の選択を行った。つまり、「単純なホールドアウト法」⁶⁾で学習していることになる。しかし、学習データの20%は利用されておらず、現時点では、学習データを分割せずに100%使用して学習を行うべきであると考えている。

学習に際しては、バッチ数を100とした。また、学習回数（epoch数）は100としたが、学習回数ごとに異なった重みの組（モデル）が生成されることから、151回から200回の学習で得られる50個の重みの組を保存し、50個のモデル推定結果を平均して検証及びテストを行った。後に述べる正則化に関する検討に当たっては、50個の検証結果の平均値

だけではなく、そのばらつきも考慮した。

2.6 時系列データ作成に伴う情報の洩れの問題

前節で、20%の「検証データ」による検証誤差が非現実的に小さかったことを述べたが、その原因は時間的に隣接した時系列データ（ベクトル）の要素の大部分が共通であることによると考えている。すなわち、時刻 t の雨量時系列データに対して、時刻 $t+1$ の雨量時系列データは先頭の雨量データのみが異なり、2番目から最後から一つ手前のデータまでは共通である。最後のデータは10日間の平均値であるため、ほとんど同じ値である。この場合、ランダムに20%のデータを選んで検証データを作成すると、検証データの全てが学習データの一部と極めて似ていることになる。従って、時間的に十分離れたデータを検証データとして用いる必要がある。情報の洩れが問題になるのはモデル推定値と観測値の間の誤差評価においてであるが、誤差評価に大きく寄与するのは水位の大きな期間のデータ、特に水位の立ち上がりからピーク及び立下りの初期のデータである。これは水位の値がそもそも大きい上に、立ち上がりの部分においては、立ち上がりのタイミングの違いが大きな値の差を引き起こすからである。このような水位のピークをもたらす降雨は、日本においては数時間から10数時間程度過去の上流の降雨であるため、数10日以上離れたデータは検証データとして用いても実際上問題は生じないと考えられる。従って、別の年のデータを検証データとして用いる場合には、事実上学習データから検証データへの情報の洩れはないと考えて良い。

別の年のデータを検証データや、最終的な精度確認のためのテストデータとして用いる際に問題になるのは、年が違うために、降雨や水位の値や時間変化が異なることである。そのために、誤差が過大に評価される。従って、単純なホールドアウト法を用いる場合、想定される多くの降雨パターンや水位上昇パターンを含む学習データを用いることが望ましく、検証データとしては、学習期間における典型的な降雨、水位パターンを含むものが望ましい。もちろん、 k 分割交差検証法を用いる方がより望ましいが、ハイパーパラメータの選定のための多数回の計算が必要であることを考慮して、単純なホールドアウト法を採用した。

2.7 深層学習モデルの設計及び過学習対策

深層学習モデルの設計にあたって、隠れ層の数及びノード数を変えて予備的な検討を行い、隠れ層の

数として3を、ノード数として1024を選んだ。本来は、ハイパーパラメータの選定は、グリッドサーチ等の方法によって系統的に行うべきであり、現在実施しているところである。過学習を抑制するための正則化及びドロップアウトについて検討したところ、ドロップアウトは50個のモデル推定値と観測値の間の誤差の平均値は大きくなかったものの、ばらつきが大きかったため採用しなかった。最終的に、検証データである2016年のデータを用いて、隠れ層第1層に正則化なし、L1正則化（正則化係数； 1×10^{-9} 、 1×10^{-8} 、 1×10^{-7} ）及びL2正則化（ 1×10^{-7} ）を適用して観測値との平方根自乗誤差（RMSE）を比較した。年間平均RMSEはどの場合も大差はなかったものの、L1正則化（正則化係数； 1×10^{-8} ）の場合が1年間の平均的な誤差、高水位時の水位ピーク付近の誤差の両者において相対的に優れていると判断した。

2.8 入力データの雨量時系列長の変化がモデル推定値に及ぼす影響

2.8.1 通年水位再現に及ぼす雨量時系列長の影響

入力データである上流13地点の雨量時系列の長さを24時間、96時間、240時間（10日）、17,280時間（720日）に設定し、学習にもハイパーパラメータの検討にも用いていない2017年1年間のデータについて水位再現実験を行った。結果を図5に示す。時系列長が24時間の場合には、水位のピーク的位置の一致は良いが低水位の部分の一致が悪い。水位のベースラインが観測値と無関係に0付近で一定になってしまっているだけでなく、パルス状のノイズも大きい。低水位時の河川水位を維持している雨量についての情報が入力にほとんど含まれていないことによって生じる誤差と考えられる。96時間、10日と全般的にかなり観測値との一致が良くなるが、10日においても低水位の部分ではノイズが見られる。降雨時系列長が720日の場合には、低水位から増水時まで観測値との一致が良く、ノイズもほとんど見られない。低水位期間のモデル化に必要な降雨のデータの情報が入力データに十分含まれていることによると考えられる。

図6に観測値と深層学習モデル推定値の散布図を示す。時系列長の増加と共に原点を通る傾き1の直線の近くに分布する点が多くなるが、低水位部分における改善が顕著である。平方根二乗平均誤差（RMSE）は、図に示していない時系列長を含め、24時間、96時間、10日、180日、360日、540日、720日に対し、0.518 m、0.303 m、0.223 m、0.179 m、

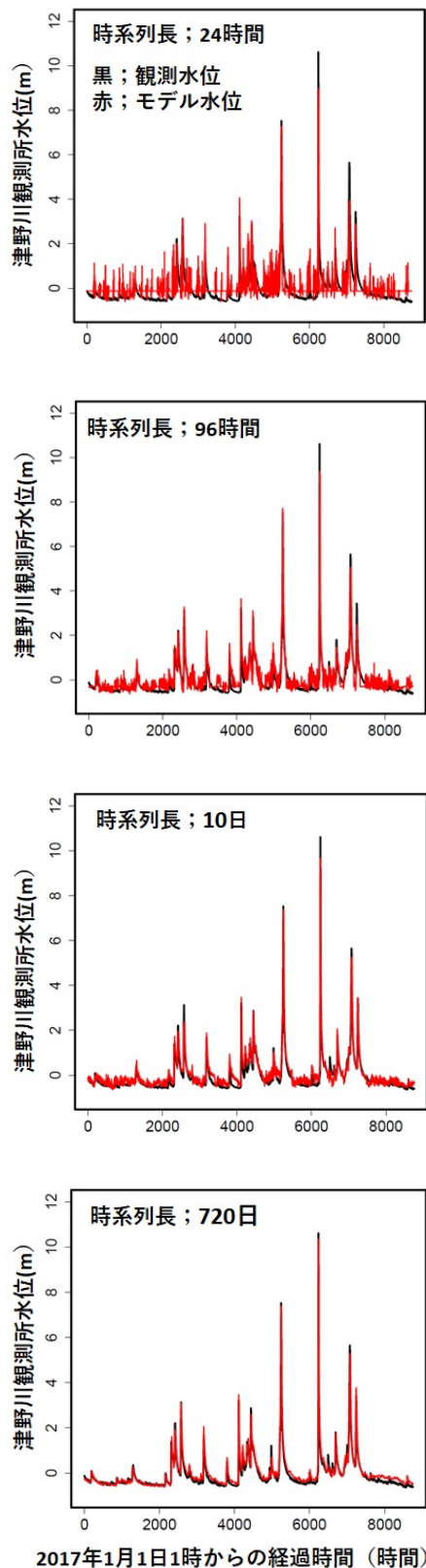


図 5. 2017 年の四万十川津野川観測所水位のモデル推定値に対する、入力データ（上流雨量時系列）の長さの影響。黒実線は観測値、赤実線はモデル推定値。

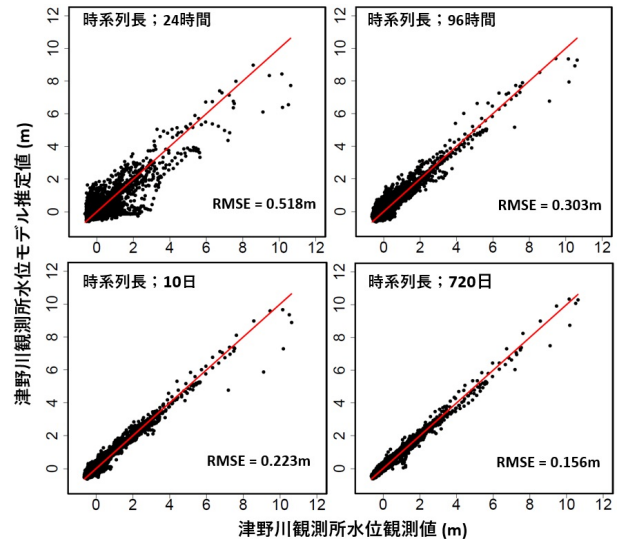


図 6. 図 5 に対応する散布図。赤実線は原点を通る傾き 1 の直線。平方根二乗平均誤差 (RMSE) をグラフ右下に示す。

0.169 m、0.162 m、0.156 m と時系列長の増加と共に漸減した。このように、約 2 年遡る雨量時系列データを入力とすることによって、低水位や洪水時を含めた 2017 年 1 年間の水位を良好に再現した。

2.8.2 2017 年最高水位時の水位再現に及ぼす雨量時系列長の影響

2017 年の最高水位ピークは、9 月 17 日～18 日の台風 18 号接近時に現れた。この期間について、24 時間、96 時間、10 日、720 日の雨量時系列長の入力データに対するモデル推定値と水位観測値の比較を図 7 に示す。時系列長 24 時間の場合も水位ピークの形はある程度再現しているが、裾の部分の再現性は低い。96 時間では、ピークの形の再現性は相当に良くなる。10 日、720 日と時系列長が更に長くなるにつれ、ピークの立ち上がりから先端部、立ち下りまで、観測水位を相当に良く再現するようになる。

2.9 学習に用いた水位を上回る水位の推定

ニューラルネットワークによる水位予測に関しては、「想定外つまり未経験の洪水規模における予測精度が著しく低下する可能性を有している点がある。」⁸⁾、と危惧されてきた。これは、深層学習以前のニューラルネットワークが非線形な現象を十分に表現する能力を持っていない、あるいは、洪水時のみの限られたデータのための「スモールデータ」によって学習していたことによると考えられる。深層学習とビッグデータを用いて作成したモデルであれ

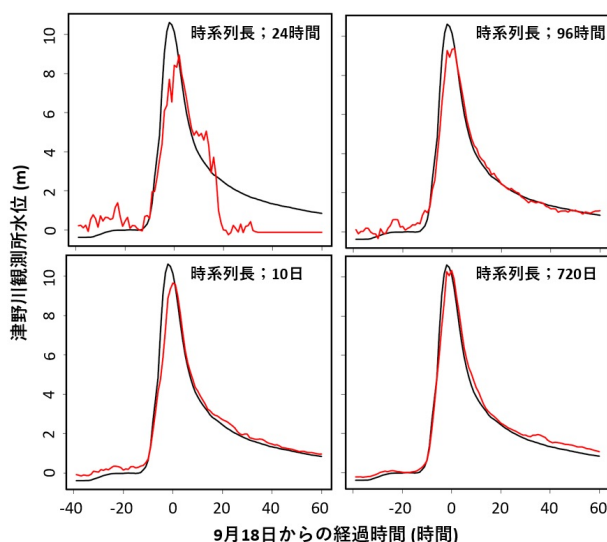


図 7. 2017 年台風 18 号に伴う高水位のモデルによる再現と入力データ（上流雨量時系列）長の影響。黒実線は観測値、赤実線はモデル推定値。

ば、学習した水位の上限付近を適切にモデル化している可能性が高く、学習に用いた水位の観測期間には経験したことのない洪水時の水位を再現する能力も優れていることが期待される。そこで、最大水位が 10.68 m 以下である 8 年間（2008、2009、2010、2012、2013、2015、2016、2017 年）のデータによる学習済みモデルを用いて、2014 年の台風 11 号豪雨時の 8 月 7 日～13 日の水位（最高水位 14.50 m）の再現を試みた（図 8）。概ね水位の時系列を精度良く再現しており、推定ピーク水位と観測値の差は 0.47 m に過ぎない。津野川観測所は中流に位置しており、水位上昇によって川幅が極端に増加することがないため、学習された過去の雨量と水位の関係がそのまま適用できると考えられる。下流部の水位推定においては、川幅が水位上昇の際に大きく変化することが考えられ、同様の結果が得られるか否かについては今後確認が必要である。しかし、過去の記録を大きく上回る水位を深層学習によって推定し、さらに予測することが不可能ではないことの一端を示すことができたと考えている。

2.10 2 節のまとめと今後の課題

上流の、2 年程度の長期の雨量時系列を入力データとし、多層パーセプトロンを用いて、洪水時を含めて通年の河川水位を高い精度でモデル化できることを示した。また、この手法により、学習事例より 30% 以上高い水位の推定が可能である例を示すことができた。

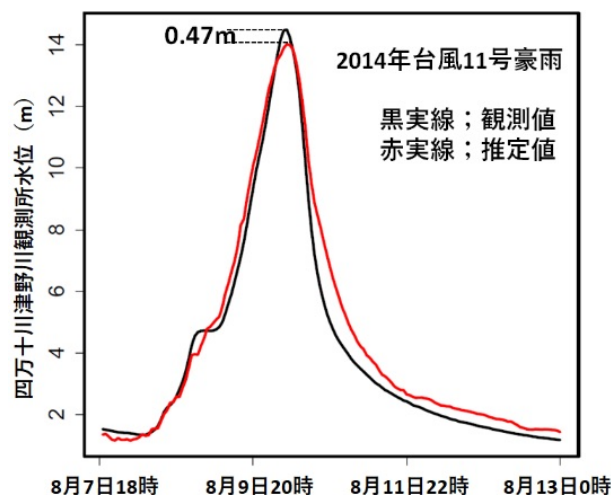


図 8. 学習に用いたデータの最大水位（10.68 m）を 30% 以上超える水位の再現。水位ピークは 14.50 m。

本研究で開発したモデルは、降雨量のみを入力データとして用いていることから、気象庁降水短時間予報による降雨の予報値を用いて水位予測を行うならば、6 時間あるいは 15 時間前までに記録的な増水への対応を行うことが可能になることを示唆することができたと考えている。気象庁の降水短時間予報を河川水位予報に用いることが容易になれば、豪雨の発生と河川水位増加の時間差の小さな中小河川の水位予測が可能になる。現在、低価格で、洪水時のみに水位測定を行う「危機管理型水位計」の設置が進められている。この水位計のデータを、「死活監視」として 1 日 1 回行われる水位観測も含めて利用し、本研究の成果を活用することが可能と考えている。そして、このような中小河川の水位予測を担保するための気象庁等による降雨予測精度の向上、データ利用の利便性の向上の気運がさらに高まることを期待したい。

なお、本報告では触れなかったが、1～6 時間程度先の水位を学習することによって、4 時間程度までの水位予測が実用的な精度で可能であることを確認している。また、雨量時系列の特徴抽出を予測に利用することが可能な、一次元畳み込みニューラルネットワーク（一次元 CNN）を用いると、予測精度が向上することも確認している。時系列データの解析に良く使われる LSTM（Long Short Term Memory）が予測精度の向上に資する可能性の確認を含め、一級河川である四万十川の場合には、地上観測雨量のみを入力とした場合にも、予測精度の更なる向上の余地は残されていると考えている。

本研究では学習データを 80% しか活用していな

いこと、時系列データの分解能の設定の最適化やハイパーパラメータの探索をより系統的に行う余地があることから、現在、計算手法のバージョンアップを行って再計算しており、若干精度が改善された結果が得られると期待しているところである。

3. 気象庁解析雨量時系列を入力データとして用いた深層学習による鏡川宗安寺観測所の水位推定・予測モデル

3.1 背景と目的

高知市内中心部を流れる鏡川のような中小河川は、流域面積が小さく、上流の降雨から下流における水位上昇までの時間が短いため、上流の雨量観測値や水位観測値から下流の水位の予測を行うことは困難である。高知工科大学環境理工学群中根研究室では、鏡川上流2地点の雨量データとダム諸量（流入量、全放流量、貯水量等）を入力データとして、高知平野の入り口に位置する宗安寺観測所の水位を、深層学習を用いて高い精度で推定できおり、1時間までの予測も可能であることを示した²⁾。しかし、予報値のないダム諸量を含んだ水位推定・予測モデルであるため、2時間以上の予測は困難である。ダムを有する河川の水位を上流の雨量データのみから推定・予測するモデルであれば、気象庁が提供している降水短時間予報を用いて6時間程度の早期の洪水予測が可能になる。本研究では、降水短時間予報と同じ1kmメッシュの書式の過去データである解析雨量の時系列を入力データとした深層学習によって宗安寺観測所の水位の推定・予測を行い、降水短時間予報活用のノウハウを蓄積することを狙いの一つとしている。しかし、気象レーダーを基にした解析雨量と、観測所の雨量計を用いた観測値（以後、地上雨量と記す）は必ずしも一致しないことから、地上雨量を入力データとして用いた場合との比較を行う。この二つが本研究の目的である。

3.2 深層学習モデル

深層学習には、2節の四万十川の場合と同様の多層パーセプトロンを用いた。多層パーセプトロンのハイパーパラメータ（隠れ層の数、隠れ層のノードの数、バッチサイズ）については様々な組み合わせを検討した結果、隠れ層の数4、隠れ層のノード数1024、バッチサイズ128を採用した（図9）。正則化についてはL1正則化を用い、正則化係数として 1×10^{-5} 、 1×10^{-6} 、 1×10^{-7} の3種類を検討した。また、正則化を適用する層に関して1層目のみの場合

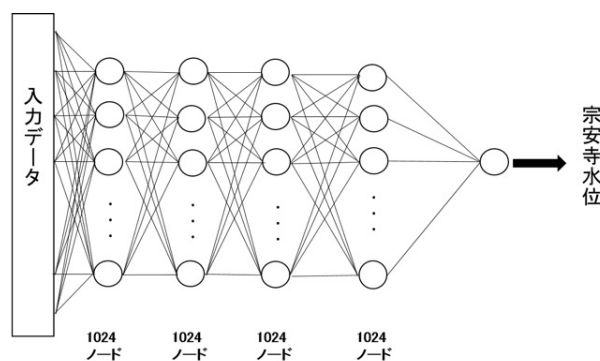


図9. 使用した多層パーセプトロンの構造

と2層目のみの場合を検討した。その結果、正則化係数 1×10^{-6} を1層目のみに適応したものを採用した。各層の活性化関数にはPReLU関数、最適化関数にはNadam、誤差関数には平均自乗誤差（Mean Squared Error）を使用した。

3.3 鏡川の概要と雨量・水位観測データ

鏡川の概要を図10に示す。鏡川は流域面積が170万 km^2 、流路延長が31kmの二級河川（中小河川）である。鏡川では2014年8月に台風12、11号が相次いで接近した際、高知市全域33万人に避難勧告が出された。その際、宗安寺水位観測所では4.97mの過去最高の水位を記録した。

宗安寺水位観測所の上流には、鏡ダムを通して宗安寺に至る支流と、鏡ダムの下流に流れ込む支流がある。本研究では、この二つの合流点の下流で高知平野の入り口にある宗安寺水位観測所の水位推定・予測モデルの開発を行った。入力データとして鏡川上流に位置する平石雨量観測所、柿の又雨量観測所の2地点の雨量データを使用した。平石雨量観測所は図10の青色で示した観測所であり、鏡ダムを経由する河川の上流に位置する雨量観測所である。一方、柿の又雨量観測所は図10の緑色で示した観測所であり、鏡ダムを経由せずダム下流で合流する支流の上流部に位置する雨量観測所である。

3.4 入力データの作成

本研究では雨量データとして、解析雨量データ（気象庁提供）¹⁴⁾と観測所の地上雨量データ（高知県土木部河川課提供）の2種類の雨量データを用いた。解析雨量は1kmメッシュの2次元データであるが、平石及び柿ノ又観測所に最も距離に近いデータを選んで用いた。この2種類の雨量データのうち、解析雨量データには欠測値がないが、地上雨量データには複数の欠測値が存在するため、そのま



図 10. 鏡川の水系流域図と雨量及び水位観測地点。高知県「鏡川水系河川整備計画（平成 29 年 9 月）」を基に作成。

ま用いることはできない。そこで、ニューラルネットワークを用いて解析雨量から地上雨量を推定して欠測値を置き換えた。ニューラルネットワークとしては、256 ノードの隠れ層が 1 層の多層パーセプトロンを用いた。この「ニューラルネットワーク補間」とも言うべき方法によって、実測値を良く再現できることを確認した。雨量データはすべて 30 分毎の 1 時間積算雨量である。

解析雨量データ、地上雨量データのそれぞれを入力データとして水位推定モデルを作成した。その際、2 節の四万十川の水位推定の場合と同様に、遡る時間と共に時間分解能が低下する雨量時系列データを作成した。すなわち、水位観測時からの遡り時間が 20 日までは 1 時間の分解能（平均時間）とし、それ以降は分解能を最大 30 日まで低下させ、1 年間遡る雨量時系列データを 189 要素で構成した。この入力データを平石観測所及び柿ノ又観測所について作成し、378 ノードの入力層を構成し、出力層（教師データ）である宗安寺の水位と対応させた。作成したデータを、2008 年～2012 年を訓練データ（学習モデル作成用）、2013 年 2 月～11 月を検証データ（ハイパーパラメータ調整用）、2014 年 1 月～9 月をテストデータ（最終的なモデルの評価）の 3 種類に分割した。

3.5 深層学習による宗安寺観測所水位の推定

解析雨量データを入力データとして作成した水位推定モデルと、地上雨量データを入力データとして作成した水位推定モデルを用いて、テストデータの水位を推定した。その結果を散布図として図 11 に示す。ここでは、横軸が水位の観測値、縦軸は水

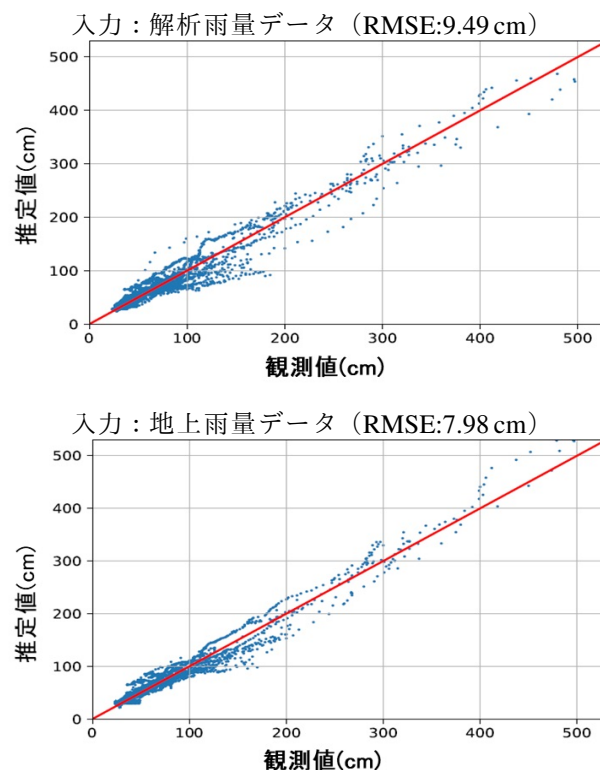


図 11. 解析雨量データ（上段）、地上データ（下段）を入力データとして作成した、宗安寺観測所水位の深層学習モデルの推定値と観測値の散布図。

位のモデル推定値である。散布図は、両モデルにおいて原点を通る傾き 1 の直線付近に点が分布していることを示している。RMSE 値には 20 % 程度の差があり、低水位の場合に違いが大きい、高水位では大きな差はない。このことから、防災用には解析雨量データも地上データに劣らない精度があるが、環境保全を検討する際には、地上雨量データが入力データとして優れていると言える。このような違いの理由として、地上雨量データは実測値であり、解析雨量データは、気象レーダーとアメダスデータを基に作成した解析値であることが考えられる。しかし、洪水時は 4 m 以上の水位を問題にしていることを考慮すると、解析雨量データを入力とした推定値も十分実用的な精度を持っていると言える。

3.6 2014 年 8 月 1 日から 8 月 10 日の豪雨の際の宗安寺観測所水位の推定

解析雨量データを入力データとして作成した水位推定モデルと地上雨量データを用いて、2014 年 8 月 1 日～8 月 10 日の高知県に、台風 12 号、11 号が相次いで接近した期間の水位推定を行った。この間

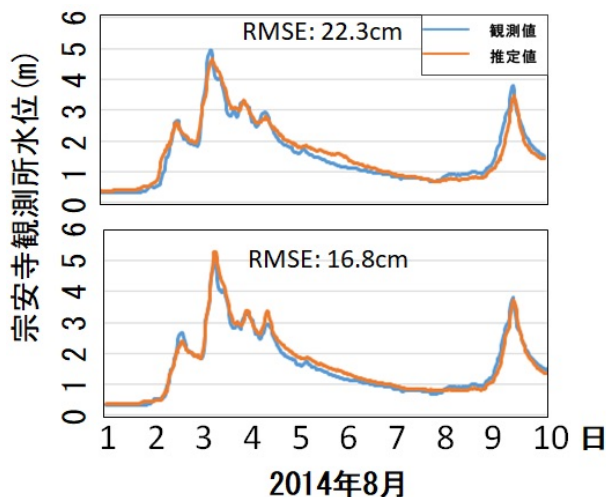


図 12. 2014 年 8 月 1 日から 8 月 10 日の解析雨量データ（上段）、地上雨量データ（下段）を入力データとして推定した宗安寺観測所水位と観測水位の比較。

の水位の最大値は 4.97 m であり、2009 年から 2012 年の期間の水位は全て 4 m 以下である。水位観測値とモデル推定値の時系列を RMSE 値と共に図 12 に示す。解析雨量のデータで作成したモデル、地上雨量データで作成したモデルの推定結果と観測値を図 12 に示す。上段が解析雨量データで作成したモデル推定結果との比較、下段が地上雨量データで作成したモデルの推定結果との比較である。

図中の時系列プロットの橙色がモデル推定値、青色が観測値を示している。どちらのモデルも豪雨時の水位の変化を良く再現している。また、推定が困難である水位の立ち上がり部分やピーク水位についても再現性が良い。入力データによって RMSE 値に若干の差があるのは、主に、立ち上がり部分のわずかな違いとピーク後の裾の部分の違いによるものと考えられる。

3.7 宗安寺観測所水位の予測

水位予測は、1 時間後、2 時間後、3 時間後等の観測水位を学習することによって可能である。水位推定モデルに用いたハイパーパラメータを用い、深層学習によって水位予測モデルを作成した。入力データには 2009 年から 2014 年 7 月 31 日までの解析雨量の雨量時系列データ、教師データには同じ期間の宗安寺の水位データを用いた。この水位予測モデルに、2014 年 8 月 3 日 1 時から 24 時の間の 30 分ごとの解析雨量データ（1 年さかのぼる雨量時系列）を入力データとして与え、0～3 時間予測を行った。

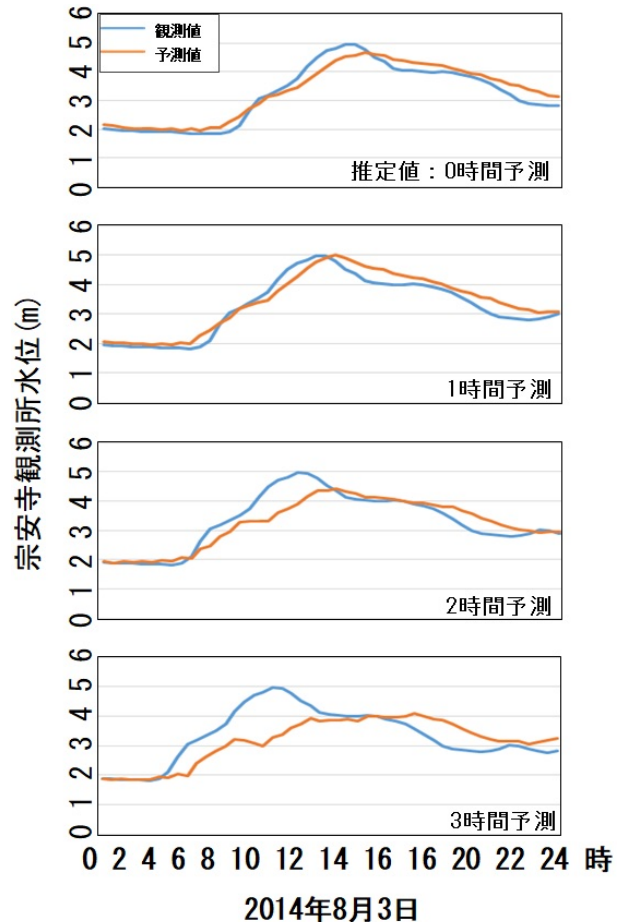


図 13. 2014 年 8 月 3 日の、解析雨量データを入力データとして用い、深層学習によって作成した宗安寺観測所のモデル推定水位、1 時間～3 時間予測水位及び観測水位の時間変化。

結果を図 13 に示す。

推定値（0 時間予測値）と 1 時間予測値には大きな差はないが、2 時間以上の予測では、ピーク水位が過小評価となり、ピークに達する時刻が後ろにずれている。2 時間以上の予測に用いた入力データ（雨量時系列）には、河川上流の観測所付近に降った豪雨がピークを過ぎた情報が含まれず、豪雨の立ち上がり時の情報しか含まれていないからである。したがって、鏡川における 2 時間以上の水位予測には雨量の予報値が必要であり、気象庁の降水短時間予報を用いることができれば、より早期の水位予測が可能になると考えられる。

3.8 3 節のまとめと今後の課題

本研究は、ダムを有する中小河川である鏡川においても、ダムの諸量を含まず、上流の雨量時系列データのみを用いて作成した水位推定モデルによっ

て、ダム下流の宗安寺水位観測所の水位を、防災情報として十分な精度で推定できること明らかにした。

入力データとして解析雨量データと地上雨量データを用いた。地上雨量データには、欠測値の補間の際して、解析雨量と地上観測雨量の関係を、ニューラルネットワークを用いてモデル化する「ニューラルネットワーク補間」を試みたところ、単なる直線回帰よりも良好な補間を行うことができた。

結果として得られた水位推定モデルについては、地上雨量データを入力データとして用いる方が観測水位の再現性に優れていたが、高水位領域では大きな差がなかった。従って、防災目的には解析雨量も実用的な精度を持っているが、環境保全について検討する場合は地上観測による雨量を用いる方が良いと言える。雨量の予報値を用いない限り、1時間予測までは水位上昇とピークに達する時刻、ピーク水位を良く予測できるが、2時間以上の予測では、ピークに達する時刻やピーク水位が十分な精度では予測できないことが確認できた。

今後の課題として、下記が挙げられる。

- 本研究では上流2地点の雨量データを入力データとして作成したモデルを用いた検討を行ったが、鏡川流域圏全域、あるいはより広範囲の解析雨量データを用いた水位推定・予測モデルを用いた場合にどの程度推定・予測が改善されるかを確認すること、
- 気象庁の降水短時間予報の6時間先までの雨量予報値を入力データとして予測を行い、早期予測がどの程度の実用的な精度を有するかを確認すること、
- 四万十川のような一級河川の場合にも共通の課題であるが、入力データを1次元あるいは2次元の画像データと見なして畳み込みニューラルネットワーク(CNN)を適用することによって予測精度が向上するか、あるいは、時系列データの解析に良く使われるLSTM(Long Short Term Memory)が予測精度の向上に資するかを確認することが挙げられる。

4. 深層学習を用いた仁淀川大渡ダム流入量の推定・予測

4.1 背景と目的

大渡ダムが位置する仁淀川の流域圏は、年平均降水量が約2,800mmであり、全国平均の1,700mmと比較すると約1.7倍という全国有数の多雨地帯である。また、台風の常襲地帯であり、昭和38年の台

風9号では、床上浸水1,569戸、床下浸水289戸、昭和50年8月の台風5号では、床上浸水5,272戸、床下浸水1,792戸に及ぶ洪水被害が発生した¹⁵⁾。これは大渡ダムの竣工以前の被害であり、大渡ダムの竣工以降洪水被害は減少した。しかし、平成26年8月の台風12号による豪雨では、いの町や日高村等で床上浸水271戸、床下浸水237戸に及ぶ洪水被害が発生したことから¹⁵⁾、洪水の問題が完全に解決したわけではない。このような洪水被害を最小限に抑えるためには、ダムによる適切な洪水調節が必要不可欠である。洪水調節とは、豪雨の影響で増加した流入量をそのまま放流せず、一定量を貯留することによって下流の水位上昇を防ぐことである。適切な洪水調節を行うためには、事前に放流を行い、貯水量を減らしておくことや、適切な流入量を予測し、それに合わせた放流を行うことが重要である。流入量は上流域の降雨や、地形、地質等によって決まるため、非線形性が強い。一般的には、分布型流出モデルや貯留関数法を用いたモデル、タンクモデル等の流出モデルを用いて流出解析を行い、流入量を予測する。しかし、このような物理モデルにおいては、流域の標高データから斜面を特定し、流域を分割し、流域内の水循環を流出システムとして数式で表現するため、複雑なモデリングが必要である。また、その流域に合わせた近似が必要であるため、ある河川のモデルを他の河川に使用することは困難である。

上記のような流出モデルによる定量化が困難である雨量と流入量の関係を、非線形な関係のモデル化に優れた深層学習を用いてモデル化し、流入量の予測可能性を検討することが本研究の目的である。すなわち、大渡ダムの上流に存在する14か所の雨量観測所の実測データのみを用いて、多層パーセプトロンを用いた深層学習による回帰モデルを作成し、大渡ダムに流れ込む水量の予測可能性を検討する。その際、入力に雨量データのみを用いることによって、他のダムに対して応用できるような汎用性の高いモデルの作成が可能になる。本研究ではこの深層学習モデルを用いて流入量の予測を行い、ダムの洪水調節に資する予測値となり得るかについて検討した。

4.2 仁淀川及び大渡ダムの概要

仁淀川は、源流を愛媛県の石鎚山(標高1,982m)に発する一級河川であり、図14のように中流から高知県に入り太平洋に流れ込んでいる。幹川流路延長は124km、流域面積は1,560km²であり、背景で



図 14. 仁淀川流域圏概要と大渡ダムの位置。(国土交通省河川局平成 20 年 3 月「仁淀川水系河川整備基本方針」参考図に基づき作成。)

も述べた通り全国有数の多雨地帯である。流域面積のうち 86%が森林であり、上流域は石鎚山をはじめとする急峻な山地である。中流域もほとんどが山地で構成されており、下流に広がる平野には土佐市やいの町等の主要な市街地が位置する。この平野は地盤高が仁淀川の計画高水位より低い事に加えて、仁淀川から離れるほど地盤が低くなるという特徴があるため、仁淀川からの背水の影響を受けやすい。また、堤防決壊時の被害ポテンシャルが高いため、想定氾濫区域は流域外にまで及び、想定氾濫危険区域内人口は 4.7 万人にのぼる¹⁶⁾。これらの条件のため、古くから氾濫による水害に悩まされてきた地域である。

このような洪水被害を軽減する治水対策として、昭和 61 年に大渡ダムが建設された。大渡ダムは「背景と目的」に述べたように、仁淀川の中流に位置する多目的ダムである。ダムの高さは 96m、堤頂長 325 m、有効貯水容量 5,200 万 m^3 、集水面積は 688.9 km^2 であり、これは仁淀川の流域面積の 44%である¹⁷⁾。ダム竣工以降何度も洪水調節を実施し、下流の水位低減、破堤氾濫に対する被害ポテンシャルの軽減等の実績を上げている。これらのことから、仁淀川における大渡ダムの洪水調節は、下流の洪水被害の軽減に対して重要な役割を持つと言える。

4.3 使用した計算機とソフトウェア

CPU として Core i7 を搭載した Windows PC に GPU (GeForce GTX 1070 あるいは GeForce GTX 1080ti) を装着し、Tensorflow をバックエンドとしたフレームワーク Keras 及び Python3.6.1 を用いてプログラムを作成して並列計算を行った。

4.4 深層学習モデルの構造とハイパーパラメータのチューニング

深層学習には、3 節の図 9 と同様の、全結合層からなる多層パーセプトロンを用いた。ハイパーパラメータのチューニングとは、モデルの構造を最適化し、推定精度を向上させることである。具体的にはモデルの層数やノード数、バッチサイズ、正則化等を調整し、検証データの推定精度を比較することによって最適化することである。しかし、このハイパーパラメータチューニングには非常に時間がかかるため、本研究では精度が向上した条件を確定して次の条件を変更している。例えば、層数は 4 層が最善と判断した後は、層数を 4 層に固定して他の条件を変更するというものである。条件の変化による推定結果の比較には、検証データ全体の RMSE と洪水時の RMSE を使用した。本研究では、PC のメモリ容量やディスク容量の都合上、モデルの重みの容量が少なくて済む層数、ノード数を検討した。また、学習回数が少ない状態、つまりは短時間の学習で推定精度が高いものをより良いモデルとした。

図 15 はハイパーパラメータチューニングを行っていない状態のモデルによる推定結果であり、図 16 はチューニングを終えたモデルの推定結果である。様々な条件で比較を行った結果、層数は 4 層で十分と判断した。また、L1 正則化をモデルに追加した。これは過学習を抑制する手法の一つであり、ネットワークの重みを制限することによって複雑化を防ぐ効果がある。図 15 と図 16 の散布図を比較すると、500~2000 m^3/s の推定値が大幅に改善されていることが分かる。これは流入量増加のタイミングと減少のタイミングをより正確に捉えることができたことによる。このチューニングにより、検証データ全体の RMSE の値も 20.5 m^3/s から 16.5 m^3/s に改善した。

4.5 使用した観測データの前処理、及び入力データの作成

4.5.1 使用した観測データ

本研究では、国土交通省水管理・国土保全局が公開している水文水質データベース¹⁸⁾から取得したデータを使用した。入力データとなる雨量データとして、大渡ダムより上流に位置する梅ヶ市、面河、大味川、上直瀬、久万、永久、御三戸、黒藤川、美川、西谷、地芳、獅子越、休場、大渡の 14 地点の雨量観測所の雨量データの内、2002 年 1 月から 2017 年 12 月の 16 年間分を使用した。水文水質データベースの雨量データには欠測値が存在するため、高

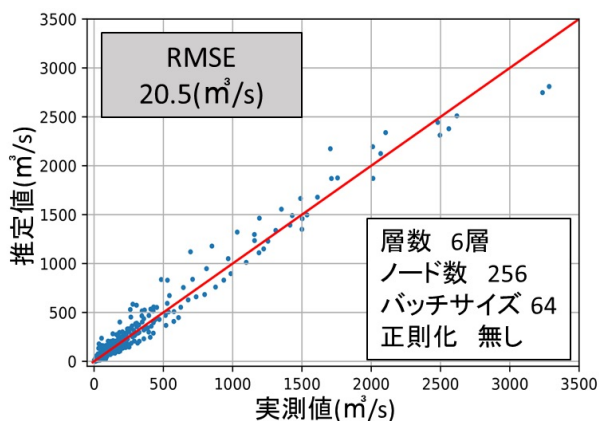


図 15. ハイパーパラメータチューニング前のモデルによる推定結果

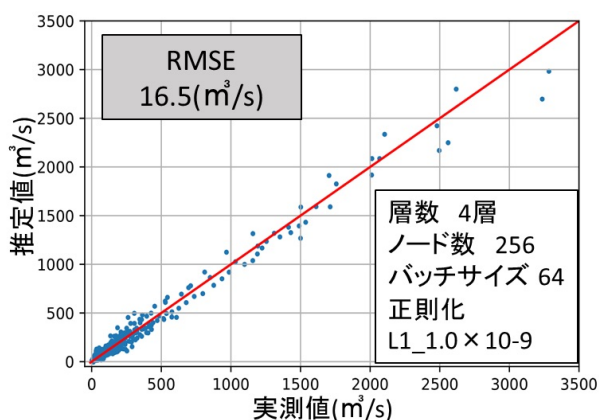


図 16. ハイパーパラメータチューニング後のモデルによる推定結果

知県土木部河川課の雨量データを2008年から2017年まで使用し、対応するデータの欠測を補間した。それでも補間できない欠測は、欠測のある観測所付近の欠測のない観測所データをそのまま使用して補間した。

また、教師データとして2003年1月から2017年12月の大渡ダムの流入量データを使用した。流入量データにみられる欠測については、高知県土木部河川課の大渡ダム流入量データを併用して補間した。最終的に補間できない流入量の欠測のある時刻については、対応する雨量データと共に学習データから取り除いた。

4.5.2 流入量データの異常値処理

教師データとして使用したダム流入量は、上流からダム貯水池に流れ込んでくる水量のことであり、1秒あたりの流れ込む水の体積 [m^3/s] で表される。このダム流入量は直接観測することが困難であり、

観測地点における貯水位の時間変化とダム放流量の観測値から計算によって求められる。流入量の算出には、貯水位を観測し、ダムに固有の貯水位－ダム容量曲線を用いて1時間毎の貯水量を求め、この間の放流量を加えて計算する方法が用いられる。しかし、貯水位は風やゲート操作などが原因となる波浪により不規則に変化する場合があるため、この方法を用いて流入量を計算した場合、算出結果が通常では考えられない負の値を示す場合や、実際より大きく見積もった結果を示す場合がある¹⁹⁾。このような異常値を除去するために、R言語の平滑化スプラインライブラリを利用したプログラムを作成した。このプログラムを用いて流入量データに対して平滑化スプライン関数を当てはめ、元データとスプラインフィッティング結果の残差を求めた。その際、データを流入量の変動が大きい期間と変動が小さい期間に分割し、それぞれに自由度を設定した。これは流入量の大きいデータの残差と小さいデータの残差に差が生じることを防ぐためである。全データの残差の標準偏差 (σ) を求め、 2σ 以上に分布するデータを異常値としてスプラインフィッティングした値で置き換えた。本研究では、このように異常値処理を施した大渡ダム流入量を教師データとして使用した。

4.5.3 雨量データの時系列データ化

本研究では、四万十川や鏡川の水位推定・予測と同様に、観測所の雨量データをそのまま利用するわけではなく、過去の雨量の情報をモデルに与えるために、時系列として入力に使用している。これは元々1列である1つの観測所の雨量データに、図17のように列方向に過去の情報を与え、時系列化を施すものである。この時系列データには、時系列を過去に1時間ずつ遡る1時間データ区間と、1時間区間より過去のデータを2時間平均、3時間平均のように複数のデータを平均し、1つのデータとして利用する時間平均データ区間が存在する。この時間平均データ区間にはデータを圧縮する役割があり、少ないデータで長期間の時系列の情報を取り入れることが出来る。このような時系列を上流14ヶ所の雨量観測所データそれぞれに施し、入力データとして使用した。

4.5.4 仁淀川大渡ダム流入量推定モデル

四万十川水位推定、鏡川水位推定と同様に、学習データ、検証データ及びテストデータについては、「単純なホールドアウト法」⁶⁾を用いた。その際、例

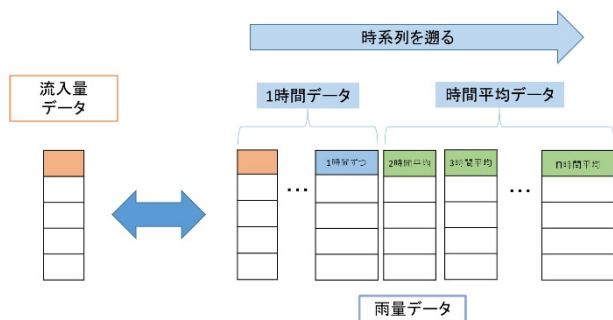


図 17. 雨量観測データの時系列化の方法

例えば、平年より流入量と雨量が少ない年を検証データとした場合、流入量の大きいデータに対して検証が不十分になる。このことを考慮し、雨量、流入量が平年通りである2013年を検証データとして使用した。そして、学習データとしては、2003年～2012年及び2014年のデータを用いた。また、テストデータとして、2017年のデータを用いた。なお、検証の際の検討に基づき、入力データである雨量時系列の遡り期間は1年とし、「1時間データ」の期間を10日間に設定した。

このような学習によって作成した大渡ダム流入量推定モデルを用いて、2017年の入力データ（上流雨量時系列）を用いて推定した流入量を観測値と比較すると、平方根自乗誤差（RMSE）は $17.8 \text{ m}^3/\text{s}$ であった。

4.5.5 仁淀川大渡ダム流入量の予測

時刻 t の流入量データが n 時間予測流入量である場合、学習に用いる入力データは、 $t - n$ が先頭の上流雨量時系列データになる。そのため、予測の際には、時刻 t の直近の n 時間の上流雨量の情報は学習には使われない。その結果、 n 時間予測値が推定値（0時間予測値）より精度が低くなると考えられる。

2017年の全観測データに対する予測結果の誤差は下記のとおりである。2時間予測までは推定モデルの場合と同程度の誤差であるが、3時間予測以上では急速に誤差が増大する。流入量が急激に増加する場合の予測結果を図18に示す。2時間予測までは、流量が急激に増大するが、タイミングの遅れもなく、ピーク流量もそれ程低下していない。

図19に示すように、長期的な降雨による流入量の増大の場合には、急激な流量の変化が無い場合、2時間以降の予測もそれ程精度が悪くならない。

表 2. テストデータ（2017年）全体に対する誤差

予測時間	RMSE (m^3/s)
0 時間（推定モデル）	17.8
1 時間	17.5
2 時間	15.8
3 時間	28.9
4 時間	46.6
5 時間	62.6

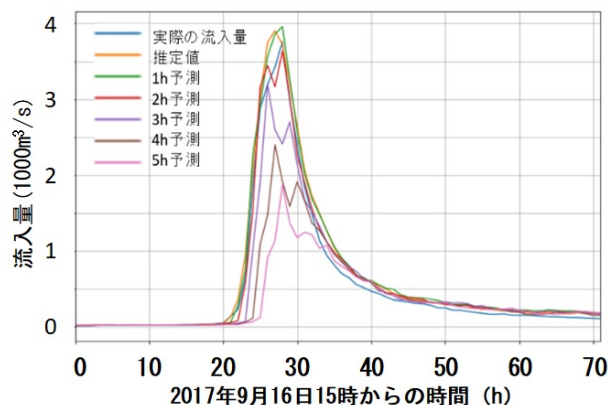


図 18. 流入量の急激な増加の例とその予測

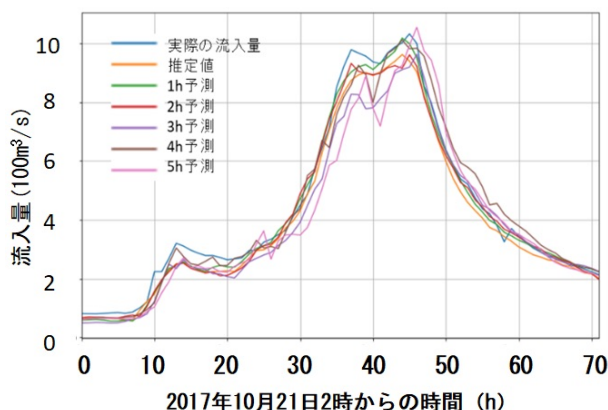


図 19. 長期的な降雨による流入量の増加の例とその予測

4.6 4節のまとめと今後の課題

本研究では、愛媛県から高知県にかけて流れる仁淀川大渡ダムの流入量を、上流14観測所の雨量観測時系列を入力データとして学習する手法を開発した。その際、下記の課題について検討した。深層学習モデルについては、四万十川水位、鏡川水位の場合と同様、全結合層からなる多層パーセプトロンを用いたが、下記の点について特に工夫・検討を行った。

- ハイパーパラメータの選定、
- ダム流入量の異常値処理、

- 雨量データの時系列化、
- 検証データの選定の工夫、

そして、流入量推定モデルだけでなく、5時間までの流入量予測モデルの精度について評価を行い、2時間予測までは推定モデルと大差ない精度が得られること、3時間以上の予測では、誤差が大きくなることを確認した。

今後の課題として、気象庁の降水短時間予報の6時間先までの雨量予報値を入力データとして用いること、14雨量観測所の時系列データを画像データとして扱い、1次元あるいは2次元畳み込みニューラルネットワーク（CNN）を用いて、より早期の予測が可能かについて検討することが挙げられる。

5. おわりに

一級河川で降雨から中下流の水位上昇まで数時間かかる四万十川の水位、二級河川で降雨から2時間以内に水位が上昇する鏡川の水位、一級河川仁淀川上流の大渡ダムの流入量を深層学習によって推定・予測した。1年あるいはそれ以上まで遡る上流の雨量時系列を入力とした深層学習によって、水位の高低、流量の多寡にかかわらず推定・予測が可能であることを示すことができた。入力データと、教師データである水位、流量の関係の学習の過程で、表流水、土壌水分、地下水・湧水の情報のみならず、ダムのオペレーションの情報さえモデル化したことが成功の鍵であったと考えている。

本研究で使用した深層学習のモデルは、隠れ層が2~4個の多層パーセプトロンであり、学生、市民、行政担当者にとっても自らプログラミングすることが困難ではない。観測データに基づいて自らモデリングと推定・予測することによって、河川に親しみ、防災、環境について認識を深め、対応力を強化するツールになり得ると考えている。

地球温暖化の進行、森林の成長や伐採、土地利用変化等による将来の流域圏の変化を、帰納法である深層学習は予測できない。温暖化影響とそれに対する適応のための、自然と社会・経済シナリオに基づくシミュレーションは増々重要になる。しかしながら、現在と将来のそれぞれの時期において、河川防災・環境の観点からの気候変動適応のツールとして、深層学習は大いに活用されると期待している。

謝辞

河川、データサイエンス、気象、シミュレーションに関してディスカッションして下さった、高知工科大学古沢浩教授、柴田清孝教授、全卓樹教授、那

須清吾教授、吉村耕平助教に感謝いたします。また、河川データを提供して下さった高知県土木部河川課にお礼申し上げます。なお、本研究の一部は、「SI-CAT 気候変動適応技術社会実装プログラム」の一環として実施されております。

文献

- 1) 情報処理推進機構, “第1章 技術動向”, AI白書 2017, pp. 15–166, 2017.
- 2) 中根英昭, 若槻祐貴, “環境分野への深層学習応用研究の立ち上げについて”, 高知工科大学紀要, Vol. 15, pp. 111–120, 2018.
- 3) 中根英昭, “環境分野への深層学習応用研究の立ち上げについて”, 高知工科大学紀要, Vol. 15, pp. 69–78, 2018.
- 4) 清水公貴, “変分オートエンコーダを用いた MNIST 画像の2次元潜在変数空間へのマッピングとカテゴリー識別”, 高知工科大学2018年度学位論文(学士), 高知工科大学学術情報リポジトリ, pp. 1–25, 2019.
- 5) 若槻祐貴, “ディープラーニングによる高知市の降水現象の有無の判別と早明浦ダムの流入量推定”, 高知工科大学環境理工学群2017年度卒業論文, 2018.
- 6) Chollet, F., “Python と Keras によるディープラーニング”, 4.2 節 “機械学習モデルの評価”, マイナビ出版, pp. 99–103, 2018.
- 7) Brath, A., A. Montanari, E. Toth, “Neural networks and non-parametric methods for improving realtime flood forecasting through conceptual hydrological models”, Hydrology and Earth System Sciences, Vol. 6, No. 4, pp. 627–640, 2002.
- 8) 槻山敏昭, 外山久典, 笹原和彦, 眞間修一, 関基, 竹村仁志, “ニューラルネットワークによる阿武隈川洪水予測の基礎的検討”, 河川技術論文集, pp. 73–78, 2003.
- 9) 一言正之, 服部洋佑, 梶正範, 小野寺勝, 桜庭正明, “揖斐川を対象としたニューラルネットワーク洪水予測の精度向上に関する研究”, 河川技術論文集, Vol. 18, pp. 441–446, 2012.
- 10) 一言正之, 桜庭雅明, 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), Vol. 72, No. 4, pp. I.187–I.192, 2016.
- 11) 一言正之, 桜庭雅明, “深層ニューラルネットワークと分布型モデルを組み合わせたハイブリッド河川水位予測手法”, 土木学会論文集 B1 (水工学), Vol. 73, No. 1, pp. 22–33, 2017.

- 12) Shen, C., “A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists”, Water Resources Research, 2018.
- 13) Google, “Keras Documentation”, <https://keras.io/ja/>
- 14) 気象庁, “解析雨量”, <https://www.jma.go.jp/jma/kishou/known/kurashi/kaiseki.html>
- 15) 国土交通省高知河川国道事務所, “過去の出水状況”, http://www.skr.mlit.go.jp/kochi/work/niyodo/old_flood/old_flood.html.
- 16) 国土交通省四国地方整備局, 高知県, “仁淀川水系河川整備計画 点検結果について”, 2014. <http://www.skr.mlit.go.jp/kochi/niyodoseibikeikaku/conference/learn03/07.pdf>
- 17) 平成29年度四国地方ダム等フォローアップ委員会, “大渡ダム定期報告書 概要版”, 2018. <http://www.skr.mlit.go.jp/kasen/damu/oodo.pdf>
- 18) 中村興一, 藤間聡, 中村篤史, “ダム貯水池振動特性と貯水位平滑化について”, 水工学論文集, 第48巻, pp. 1387–1392, 2004.
- 19) 国土交通省, 水文水質データベース <http://www1.river.go.jp/>

Application of Deep Learning to River Disaster Prevention and Environmental Conservation

**— on the Shimanto River and Kagami River Water Levels, and
the Ohdo Dam Inflow of the Niyodo River —**

Hideaki Nakane¹ Yuki Wakatsuki²

Kei Yamamoto² Takuya Takeda² Tempei Hashino^{3*}

(Received: May 7th, 2019)

¹ Emeritus Professor, Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami City, Kochi 782–8502, JAPAN

² Graduate School of Engineering, Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami City, Kochi 782–8502, JAPAN

³ School of Environmental Science and Engineering, Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami City, Kochi 782–8502, JAPAN

* E-mail: hashino.tempei@kochi-tech.ac.jp

Abstract: Targeting the Shimanto River, Kagami River, and Niyodo River, we carried out Deep-Learning modeling to estimate and predict the water levels and the inflow into the dam for river disaster prevention and river environment conservation. We utilized models with multilayer perceptrons composed of full connection layers with 2 to 4 hidden layers, and for computers we used Windows PCs or Linux PCs equipped with GPUs. In all modeling, the rainfall time series data of the basin areas upstream of the measurement points of the river water levels or the dam inflow were used as input data for deep learning. The measurement data can be reproduced with higher accuracy in the case of a 1 to 2-year rainfall time series than in the case of several hours to several tens of hours. In particular, the improvement on accuracy was remarkable during the periods with low rainfalls, low river water levels and low dam inflow. These results can be explained as follows: in the case of short time series, only information on surface water is learned, while in the case of long time series, information on surface water, soil moisture, ground water and spring water can also be learned. In the case of the Kagami River, a smaller-sized river, a one-hour prediction was possible with high accuracy by deep learning modeling using the observed rainfall data upstream. However, it is recommended that the use of the Meteorological Agency's rainfall short-term forecast is necessary for longer-term prediction.