

# IoTを想定した ダム流入量予測精度向上に関する研究

## REAL-TIME DAM INFLOW FORECASTING ACCURACY IMPROVED BY INTERNET OF THINGS

天方匡純<sup>1</sup>・安野貴人<sup>2</sup>・藤井純一郎<sup>2</sup>・嶋本ゆり<sup>2</sup>・大久保順一<sup>2</sup>

1) 八千代エンジニアリング㈱ 技術創発研究所  
(〒111-8648 台東区浅草橋5-20-8 CSタワー3F)

2) 八千代エンジニアリング㈱ 技術創発研究所 AI解析研究室  
(〒111-8648 台東区浅草橋5-20-8 CSタワー3F)

私達はこれまで、河川の水位や流量を予測する際、主に予測雨量と流出・河道モデルのセットで対応してきた。しかし、前者も後者も不確実性が高く、実運用に役立つ汎用的な取り組みに至っていない。特に、ダム上流域では水位観測所やモデル構築情報も少なく、物理モデルによる水位・流量予測は、ダム下流の河川管理区間で実施される洪水予測よりも難しい条件にある。一方、センサーや通信ネットワークの技術進歩や安価化によりIoT (Internet of Things) が急速に進みつつある。そこで、分布型流出解析モデルから構成されるデジタル空間上に疑似的にセンサー設置を想定してDigital Twinを実現し、深層学習モデルの一種であるLSTM (Long Short-Term Memory) を活用してセンサーの数や位置の違いによるダム流入量予測精度の向上程度を検討した。この結果、センサー数を増やすことで基本的に予測精度は向上するが、データ特性の偏りは予測精度向上に貢献せず、それ故に流域内のセンサー設置数にも限界があることを示した。逆にデータ多様性に留意してセンサー位置を決めることで予測精度向上に繋がることも示した。

We have used a set of predictive rainfall and analytical models as main methods when predicting water levels or discharge levels. However, available flood forecasting technologies are inadequate and have not reached general utilization through practical operation. An important point is that fewer observation stations and less information exist for models in dam basins than for river management segments under dams. Therefore, predicting water levels or discharge levels based on physical models is difficult. As an alternative, the Internet of Things(IoT) is supported by the improvement and lessening costs of sensors and network technologies. Therefore, we set sensors artificially in the Digital Twin (A digital twin is a digital replica of a living or non-living physical entity. By bridging the physical and the virtual world, data is transmitted seamlessly allowing the virtual entity to exist simultaneously with the physical entity.(Saddik,2018)) basin produced using the distributed flow analysis model and using Long Short-Term Memory(LSTM), a sort of deep learning to confirm the accuracy of dam inflow forecasts obtained with different numbers and locations of sensors. Results demonstrate that the predictive precision fundamentally improves with increasing the number of sensors but that it doesn't improve with sensors which generate biased data. Furthermore, we can infer a sensor number limit. Results show that adequate sensor location improves the predictive precision.

**Key Words** : dam inflow forecasting (ダム流入量予測) , RNN (リカレントニューラルネットワーク) , LSTM (LSTM) , IoT (モノのインターネット) , Digital Twin (デジタルツイン)

### I. はじめに

大気中の温室効果ガス濃度の増加に伴って地球上のほぼ全ての地域で気温が上昇するのに対し、降水量は増加する地域と減少する地域があるとされる。そして、日本では、短時間強雨の発生回数が全ての

地域及び季節で増加すると予測されている(環境省ほか, 2018)。実際、これを証明するかのようになり、平成25年10月の大島町土石流災害、平成26年8月の広島土砂災害、平成27年9月の関東・東北豪雨、平成29年7月の九州北部豪雨、平成30年7月の西日本豪雨と、近年、従来の河川計画で想定していなかった

外力現象の多発により、多くの水災害・土砂災害が発生している。

また、短時間強雨の発生回数だけでなく、無降水日の増加も予測されており、私達の生活はこれまで以上の水リスクに晒されることになる。そこで、極端化する水リスクを平準化する仕組みが不可欠である。従来、高水流量と低水流量に大きな差がある我が国において、効率的な水運用を目指して建設されたのが巨大人工構造物であるダムである。ダムの貯水機能は一度建設されると柔軟に変えられるものではないが、運用の工夫により平滑化機能が拡張される。現在も貯水機能を最大限に活用するために、事前放流操作や異常洪水時防災操作といった運用規定が定められ、私達の生活に安全・安心をもたらすと共に潤いをもたらしている。しかし、前述の短時間強雨の発生頻度増大等の出水傾向の変化は、河川やダムの管理の困難さを増大させる一要因となると共に、ダム管理者への極度の精神的負担に繋がっていると考えられる。

洪水時のダム管理者の上記負担を軽減すると共に適切なダム運用を支援するためのツールとして、ダム流入量を予測するモデルが存在する。ダム流入量予測モデルはダム流域特有のものであり、下流域の河川管理区間ではダム流入量予測モデルと類似の洪水予測モデルで河川水位を予測している。洪水予測モデルは、国土交通省及び都道府県等で運用され、雨量を入力条件とし、流量を算定する流出解析モデルや流量を水位に変換する河道モデルから構成される(図-1上段)。ダム流入量予測モデル、あるいは、洪水予測モデル、いずれも、既往洪水に対してモデルを検証した上でモデル変数を最適化し、確定モデルとするが、各過程やモデルには入力条件、モデル

条件等の不確定性が数多く含まれ、リアルタイムでのダム流入量・水位再現も難しい現実があり、これらの予測精度向上が水防上の大きな課題である。昨今、国土交通省主体にベイズ理論に基づく粒子フィルター(工藤ら, 2009; 立川ら, 2011; 高崎ら, 2012)を活用した状態方程式のデータ同化が適用され始め、予測精度向上が期待される。ただし、流域内の水文状態を検出するデジタルデータが不足しており、偏ったフィルタリングにならざるを得ない場合が多い。今後、流域内の水位観測機器等が増えることで、分布型流出解析モデルや水位追跡計算モデルのパラメータの不確定性も低減すると考えられる。

昨今、IoT(Internet of Things)の高まりと共に、日常空間内のデジタルデータ量は急激に増大している。これに伴い、センサー技術やネットワーク技術も身近なものとなり、国土交通省においても「革新的河川管理プロジェクト」の名の下で水位計の高度化・安価化等が目指されている。山間部のダム流域では、LTE(Long Term Evolution)電波が届かない、電源が無い等、私達の居住空間とは大きく異なるIoT条件にあるが、今後、成層圏通信が実現する等してIoT化は徐々に進んでいくと想定される。つまり、ダム流域でも水文量を把握するためのデジタルデータ量が増大することが想定される。本稿では、こういった流れを踏まえ、ダム流域内の水位・流量センサー数が増加することによるダム流入量予測精度の向上程度を検討するものである。

具体的には、対象流域にて分布型流出解析モデルを構築し、その計算結果を実流域での観測データに見立て、上流域水文量から下流域水文量(ダム流入量)を相関モデルにより予測する(図-1下段)。そして、水位・流量センサーの設置条件の違いによるダム流入量の予測精度の向上程度を検討する。なお、本稿では降雨量は入力条件とせず、あくまで、ダム流域内に複数の水位・流量センサーのみを設置することによるダム流入量予測精度への影響を確認する。降雨量を扱った検討は別途試みたい。

ダム流入量を推定する方法は、機械学習の一手法であるDNN(Deep Neural Network)を活用した手法とする。DNNは、現在、画像認識、音声認識、自然言語処理の分野で積極的に研究が進み、産業界への具体的な適用も始まっているところである。各分野で適応されるDNNのモデル構造は様々であり、洪水予測の分野でも中間層2層のDNNを活用した事例(一言ら, 2016)がある。また、自然言語処理のよ

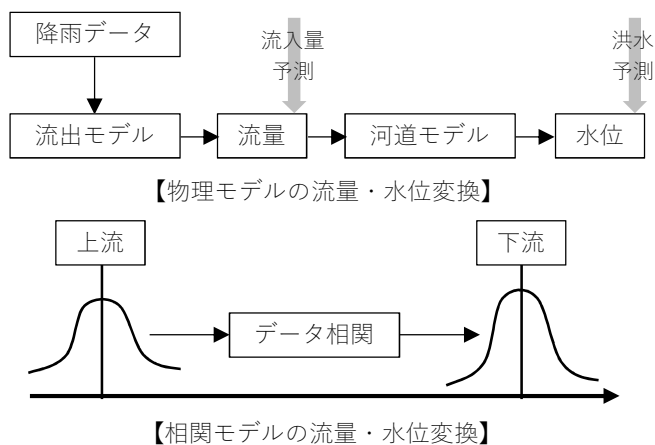


図-1 流量・水位算定方法

うに前後の言葉（データ）の繋がりを認識する必要がある分野では、RNN（Recurrent Neural Network）と呼ばれるデータ間の再帰関係を表現できるニューラルネットワークが用いられている。ただし、RNNでは保持できる前後関係は短く、長い文章を読み解くことができないため、長い前後関係を保持できるLSTM（Long Short-Term Memory）（Hochreiter *et al.*, 1995; Graves and Schmidhuber, 2005）が開発されている。本稿では、出水時のダム流入量の増加・減少過程を適切に予測するため、時系列データの前後関係を捉えることができるLSTMを選択した。

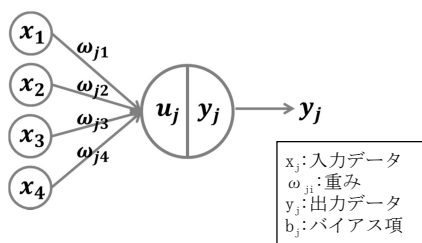
## II. LSTMの概要

ダム流入量予測に適用するLSTMは、DNNと呼ばれる人工知能の一種であり、「教師」とするデータを基にモデルパラメータを決定する教師あり学習手法の一種である。DNNのモデル構造は、これまでも洪水予測やダム流入量予測に利用されてきたニューラルネットワークの枠組みと類似している。ニューラルネットワークは、図-2のアルゴリズムを基本単位とし、これを図-3のように組み合わせることで表現力を高める。同様のアルゴリズムにパーセプトロンと呼ばれる基本単位が存在するが、これは活性化関

数（図-2の $f(u_j)$ ）にステップ関数を用いたものである。DNNは、中間層（隠れ層とも呼ぶ）の数を大幅に増やすことができ、非線形性に代表されるより複雑なデータ間の関係を読み解くことが可能になった。そして、特にLSTMは、前後関係が重要となる時系列データの特徴量抽出に威力を発揮する（Greff *et al.*, 2017）とされている。本稿では、対象とする時系列のハイドログラフの特徴量抽出にLSTMを適用する。

これまで、私達が洪水予測やダム流入量予測に利用していた3層ニューラルネットワーク（以下、3NNと記す）は図-3に示す通り、中間層を1層（1層当たりニューロン7個）とし、活性化関数にシグモイド関数を用いてきた（竹村ら, 2013；関ら, 2013）。中間層を更に多層化することで非線形性が高まり、データから複雑な特徴量を抽出できることは従来の研究で分かっていた。一方で、パラメータ設定に利用する誤差逆伝搬法（back propagation）の過程で、シグモイド関数が誤差関数の勾配消失を早め、中間層の多層化を困難としていた。このような中、活性化関数やニューラルネットワーク構造の工夫により中間層の多層化が実現されるようになってきた。

また、時系列データを扱う場合、時刻 $t$ のデータ



$$u_j = (\omega_{j1}x_1 + \omega_{j2}x_2 + \omega_{j3}x_3 + \omega_{j4}x_4 + b_j)$$

$$y_j = f(u_j)$$

図-2 ニューラルネットワークの基本構造

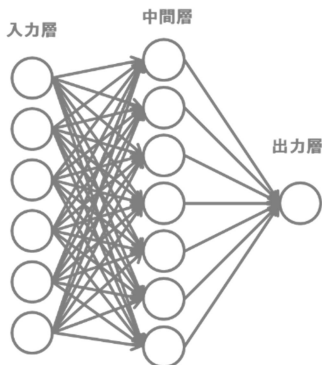


図-3 3層ニューラルネットワーク

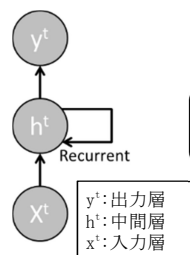


図-4 RNNの基本構造

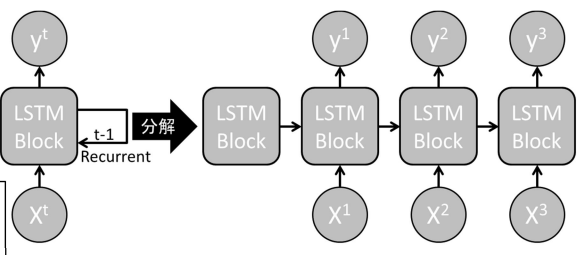
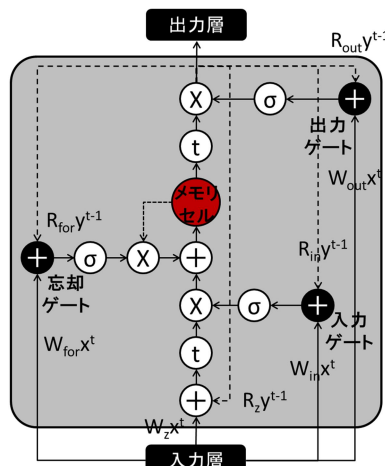


図-5 LSTMの基本構造



$\sigma$ :シグモイド層  
 $t$ :tanh層  
 $W_{for}x^t$ :時刻 $t$ の忘却ゲートへの入力値  
 $R_{for}y^{t-1}$ :時刻 $t-1$ の忘却ゲートへの出力値  
 $W_{in}x^t$ :時刻 $t$ の入力ゲートへの入力値  
 $R_{in}y^{t-1}$ :時刻 $t-1$ の入力ゲートへの出力値  
 $W_{out}x^t$ :時刻 $t$ の出力ゲートへの入力値  
 $R_{out}y^{t-1}$ :時刻 $t-1$ の出力ゲートへの出力値  
 (注)ただし、入力値は全て同じ、出力値も全て同じ

図-6 LSTM Block構造の概要

と時刻 $t + \Delta t$ のデータの関係进行学习することにより、時刻 $t$ のデータから時刻 $t + \Delta t$ のデータを予測することが可能となる。しかし、従来のニューラルネットワークは、図-3に示す通り、原則として入力データの前後関係を保持する機構を持たない。例えば、入力層に入力箇所が3個あり、「時刻 $t$ で始まるデータ」、「時刻 $t + 1$ で始まるデータ」、「時刻 $t + 2$ で始まるデータ」のように1ステップずつ時刻をずらして入力する工夫をすると、前後関係を保持して目的データとの関係を構築出来る。しかし、数日先のように予測スパンを長くしたり、XRAINのようなメッシュ雨量を利用すると、冗長な類似データセットが大量に出来るため好ましくない。

一方、RNN、LSTMは入力データの前後関係を保持する機能が整備されており、入力データを必要最小限にしてデータの前後関係の特徴量を捉え、出力データとの関係を構築することができる。RNNは図-4のように時刻 $t - 1$ 以前の間層を時刻 $t$ の間層に紐づけることができ、これにより入力データの前後関係を保持している。LSTMは、図-5のように基本的な構造はRNNと同様であるが、中間層をニューロンではなくLSTMブロックと呼ばれるメモリセルと3つのゲートを持つブロックに置き換え、データの前後関係の長期依存性能 (long-term dependencies) を高めている。つまり、DNNの勾配消失問題をネットワーク構造で緩和している。具体的には、図-6のメモリセルのアイデアを中心とした3つのゲートによるデータの制御である。3つのゲートは、忘却ゲート、入力ゲート、出力ゲートに分けられ、それぞれシグモイド層によりデータを通す・通さないを判定する。

LSTMが長期依存性能を発揮できるのはメモリセルの効果であり、過去のデータをシグモイド層やtanh層を通さず「線形和」として保持し、誤差逆伝搬しても勾配が極端に大きくなったり小さくなったりしない (Hochreiter *et al.*, 1995) ためである。忘却ゲートは、メモリセルの情報を廃棄する/しないを判定する役割を担う。例えば、出水期間から非出水期間に移行した場合、出水期間のデータ特性は忘れても良い。その上で、非出水期間のデータ特性をメモリセルに保存するための新たな情報追記を判定する必要がある。入力ゲートは、入力層からのデータを新たに加えてメモリセルの情報を更新する/しないを判定し、入力層からのデータはtanh層を通してメモリセルに新たに加えられる候補データと

して生成される。この結果、忘却ゲートと入力ゲートを介してメモリセルの情報が更新されることになる。そして、更新されたメモリセルからのデータがtanh層を通過し、それを出力する/しないを判定する出力ゲートを通じて、出力層へと吐き出される。学習時に長期記憶を必要とするAdding Problemにおいて、RNNは長さ100の配列でさえも上手く記憶保持できないが、LSTMは500-1000の配列が記憶保持できる (Li *et al.*, 2018) とされている。

### III. 対象流域の概要

検討対象とする道志ダム上流域は、図-7に示す通り、神奈川県相模川水系道志川に位置し、流域面積が112.5km<sup>2</sup>であり、上流域に降った雨量が流入量として道志ダムに到達するまでの時間は概ね60~90分である。

道志ダムの目的は発電のみであり、操作規程に従い常時満水位を超えないよう貯水位を維持する操作が求められており、水利権上の最大取水量である発電使用水量11.4m<sup>3</sup>/sを上回る流入が継続すると予測される場合、ゲート放流を行う必要がある。そして、そのために道志ダムへの移動時間や放流連絡に要する時間、休祝祭日における職員の参集時間等の

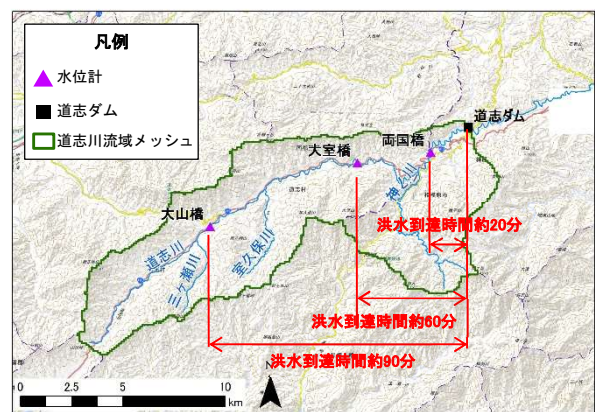


図-7 本研究の対象流域概要

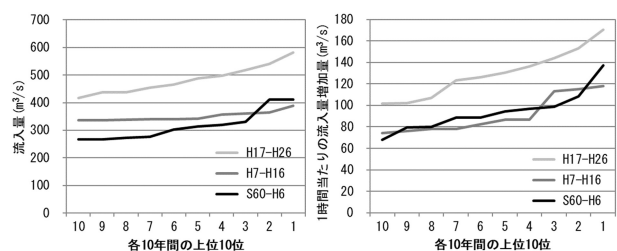


図-8 近年の出水傾向

(左：流入量，右：流入量増加量)

放流までのリードタイムの確保が必要である。

一方、過去30年間のダム流入量を整理すると図-8のとおり、近年の気候変化等により最大流入量及び1時間当たり流入量増加量（各年の1時間当たりのダム流入最大増加量）が増加傾向を示しており、堆砂対策等の必要な措置を講じると共に安全なダム管理と効率的な水運用を進めていくための精度の高いダム流入量予測システム構築が不可欠である。

#### IV. モデルの構築

本章では、道志ダム上流域において、LSTMを適用してダム流入量予測モデルを作成するための利用データ及び手順について示す。

##### 1. 学習・予測過程の概要

図-9に示す通り、学習出水（52出水）に対してダム流入量を予測するLSTMモデルを作成したのち、未学習出水（3出水）に対して予測結果・精度を確認する。

##### 2. 入力データ

LSTMを活用してダム流入量予測モデルを作成するため対象出水を定め、その入力データを整理した。具体的には、平成17年から平成27年までのゲート放流を伴う出水から、①ダム流入量が $90\text{m}^3/\text{s}$ 以上となった出水（道志ダム操作規程ではダム流入量 $90\text{m}^3/\text{s}$ 以上を洪水という）、②最高水位が常時満水位-1.5m以上の出水、③流域平均雨量が $20\text{mm}/3\text{hr}$ 又は総雨量 $100\text{mm}$ 以上の出水、の条件にて55出水を抜

き出した。そして、平成17年から平成26年までの52出水を学習対象とし、平成27年の3出水で検証した。55出水のうち最大規模出水のダム流入量は $637\text{m}^3/\text{s}$ であり、平成27年の7月3日出水は既往36位の規模、7月16日と9月9日出水はそれぞれ既往8位、9位の規模である。また、7月3日出水は上記③の条件に該当し、7月16日と9月9日出水は上記①の条件に該当する。センサーデータと見立てる入力データは、水系筋の任意地点データを活用できるように、道志ダム流域の分布型流出解析モデルの計算結果を仮想的に用いた。分布型流出解析モデルは土木研究所モデル（吉野ら、1990）を用いた。このモデルは、3種類のタンクで構成される流出解析モデル（図-10）であり、表層モデル、地下水モデル、河道モデルの各要素モデルからなる。それぞれの変数は、土値利用データ（出典：土地利用細分メッシュ図、国土地理院）、土壌データ（出典：土壌分類図、国土数値情報）、地質データ（出典：表層地質区分図、国土数値情報）を基に設定し、落水方向は5mDEM（出典：国土地理院）を $100\text{m}$ メッシュに加工（図-11）して設定した。そして、平成17年から平成26年度までのゲート放流を伴うダム流入量上位10個の代表出水についてパラメータを検証しモデルを確定した。その後、上記55出水（学習出水52+未学習出水3）に対応するレーダー・アメダス解析雨量（ $1\text{km}^2$ メッシュ）を分布型流出解析モデルに入力し、その計算結果となる各地点（メッシュ）の30分毎の流量をLSTMの入力データとした。通常、河川の観測対象は水位であるが、流出解析モデルの計算結果である流量を水位に変換せず、そのままLSTMの入力データとして用いた。また、実観測に伴うノイズも考慮していない。

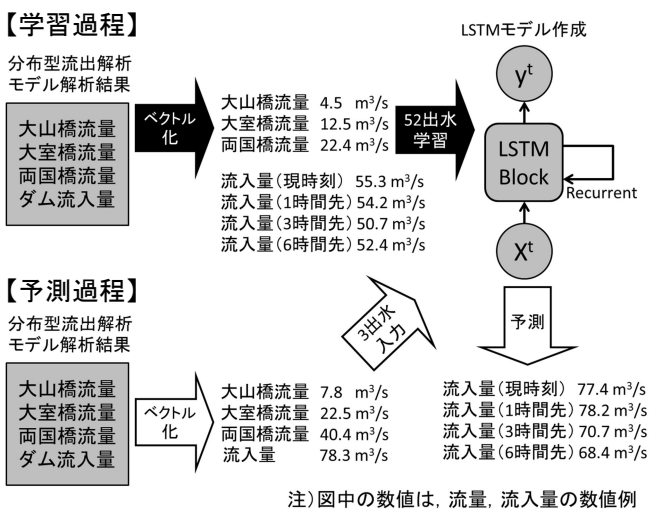


図-9 LSTMの学習・予測過程の概要

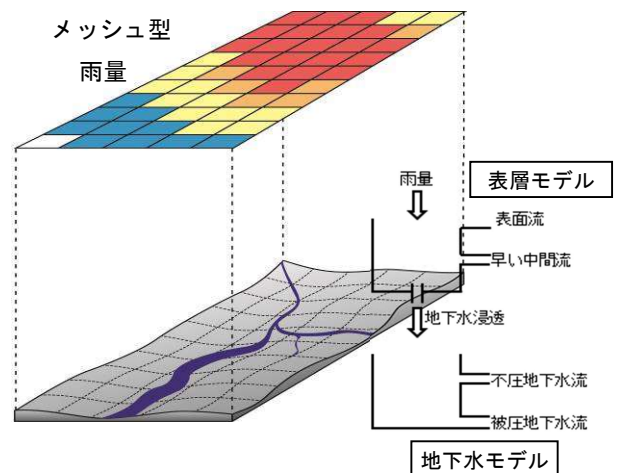


図-10 分布型流出解析モデルの構造

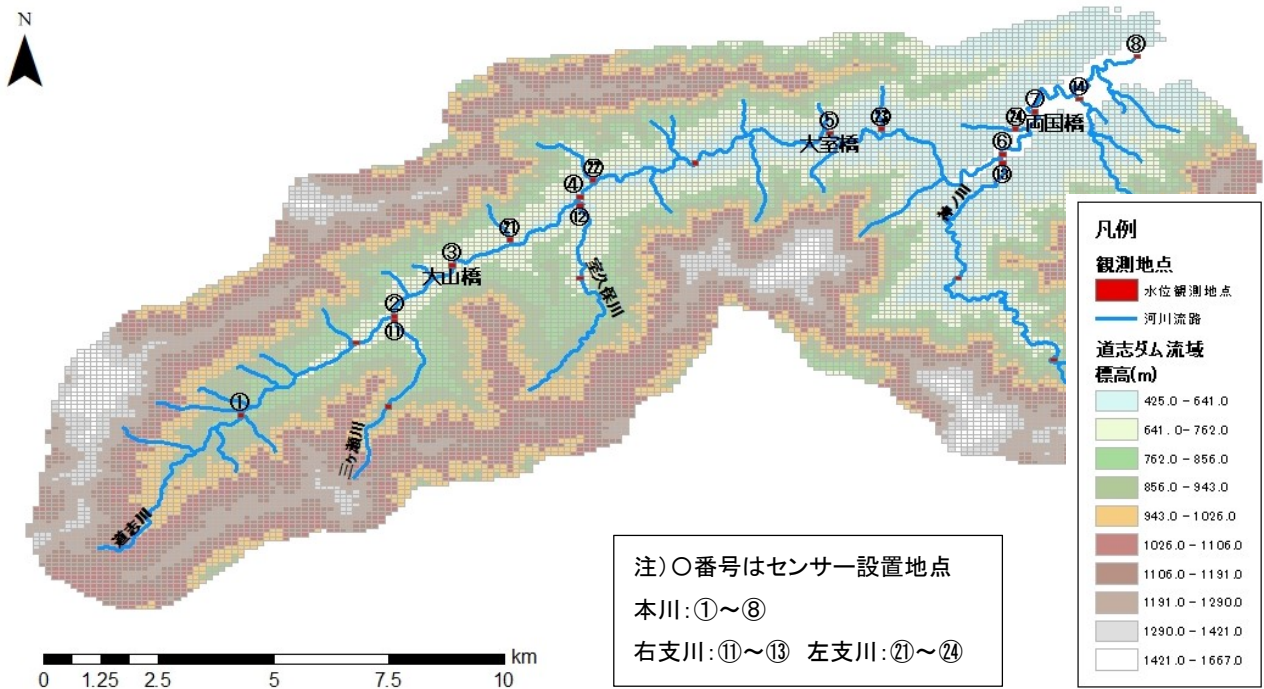


図-11 分布型流出解析モデルのメッシュ図

表-1 入力・出力フォーマットの概要

洪水 イベント No	時刻	入力データ				出力データ			
		大山橋 計算流量	大室橋 計算流量	両国橋 計算流量	ダム 計算流入量	ダム 計算流入量			
		(m <sup>3</sup> /s)	(m <sup>3</sup> /s)	(m <sup>3</sup> /s)	(m <sup>3</sup> /s)	0時間 (m <sup>3</sup> /s)	1時間先 (m <sup>3</sup> /s)	...	6時間先 (m <sup>3</sup> /s)
1	900	0.5	1.2	20.3	2.0	2.0	2.0	...	2.8
1	930	0.5	1.4	20.8	2.1	2.1	2.2	...	3.5
1	1000	0.6	1.6	21.5	2.0	2.0	2.5	...	4.5
...	1030	0.6	1.8	22.7	2.2	2.2	3.0	...	5.2
2	1400	...	...	...	...	...	...	...	...
2	1430	...	...	...	...	...	...	...	...
...	1500	...	...	...	...	...	...	...	...

### 3. 出力データ

教師データとなる出力データは、分布型流出解析モデルの計算結果の一つであるダム流入量を用いる。ただし、ダム流入量は、運用時の最新時刻（以下、現時刻という）のデータを最大限活用する目的でLSTMの入力データとしても活用する。また、洪水到達時間が最大でも1.5時間程度と想定されるため、実用上は2時間先予測が利用限界と考えられるが、表-1に示す通り、模擬的に6時間先までのダム流入量計算結果も学習対象に加えた。なお、表-1は大山橋、大室橋、両国橋、ダム流入量のデータを入力値として、現時刻から6時間先までのダム流入量を予測する際のデータフォーマットサンプルである。

### 4. モデルの構築方法

入力と出力（教師）のデータを組み合わせ、LSTMを介して、その関係性を学習することでダム流入量

予測モデルを構築する。ここでは、データ間の関係に基づくデータ活用の方法を示し、モデル構築方法について述べる。大山橋（③：番号は図-11、図-12に示すモデル上の地点番号を示す）、大室橋（⑤）、両国橋（⑦）の3観測所の位置に該当する分布型流出解析モデルのメッシュ位置の計算結果を活用して、ダム流入量（⑧）を予測するモデルを構築する。リアルタイムで入手できるデータを想定し、予測精度向上のため入力データにはダム流入量も加える。また、6時間先まで予測するため、表-1の通り、6時間先までのダム流入量を学習対象とする。なお、雨量は一切用いない。未学習出水の検証時は、現時刻までの入力値（大山橋、大室橋、両国橋の計算流量及びダム計算流入量）から6時間先までのダム流入量を予測する。構築したLSTMモデル構造は表-2のとおりである。モデルアーキテクチャーの感度分析のため、中間層1層の56個のLSTMブロックを連ねた場合（表-3）、中間層2層として各層に7個のLSTMブロックを連ねた場合（表-4）の計算も実施したが、予測精度が高かった表-2及び図-13の中間層1層のLSTM7ブロックのモデルを採用した。その予測結果を図-14に示す。

本計算では、機械学習フレームワークとしてTensorFlowとKerasを用いた。計算と教師の誤差は、平均二乗誤差で評価し、SGD（確率的勾配降下法）により最適化計算を実施した。マシンスペック i7-

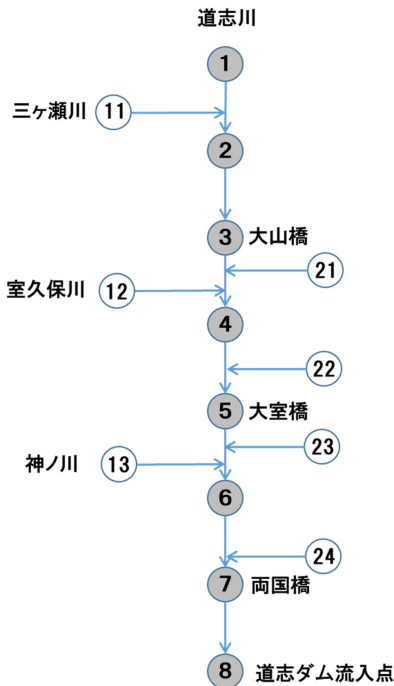


図-12 センサー配置図 (図-11と連動)

表-2 1層LSTM7ブロックモデル

層名	レイヤータイプ	出カデータ数	パラメータ数
lstm_1	LSTM層	7	364
dense_1	全結合層	1	8
パラメータ総数			372

注) LSTM層はLSTMBlockからなる層を示し、全結合層は図-2に示すニューロンから成る層を示す

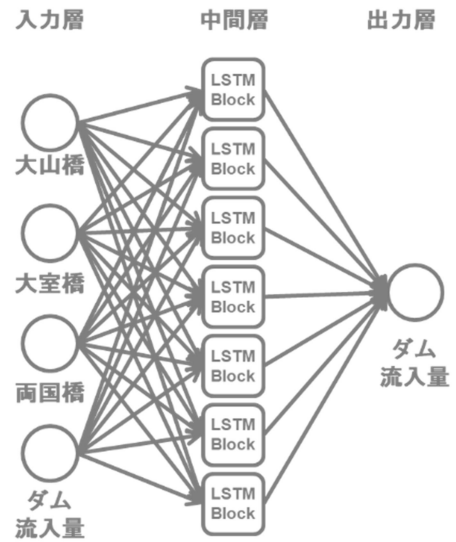


図-13 1層LSTM7ブロックモデル構造

表-3 1層LSTM56ブロックモデル

層名	レイヤータイプ	出カデータ数	パラメータ数
lstm_1	LSTM層	56	13888
dense_1	全結合層	1	57
パラメータ総数			13945

表-4 2層LSTM (7×7) モデル

層名	レイヤータイプ	出カデータ数	パラメータ数
lstm_1	LSTM層	7	364
lstm_2	LSTM層	7	420
dense_1	全結合層	1	8
パラメータ総数			792

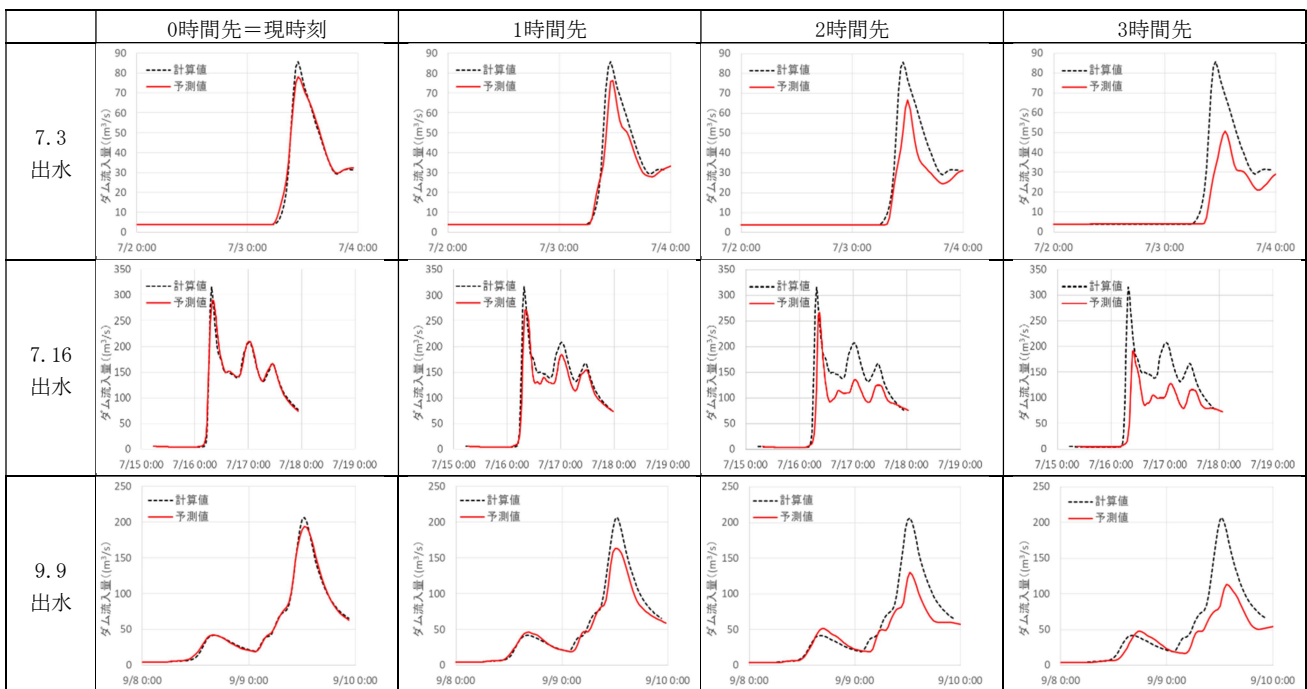


図-14 平成27年出水の予測結果

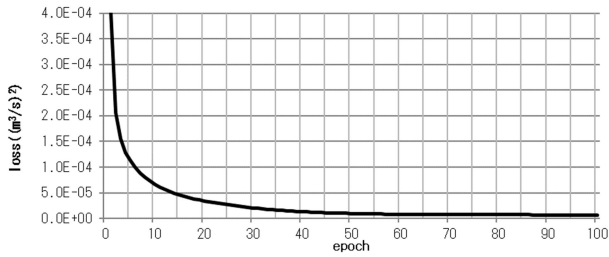


図-15 1層LSTM7ブロックのパターン3における学習誤差プロセス

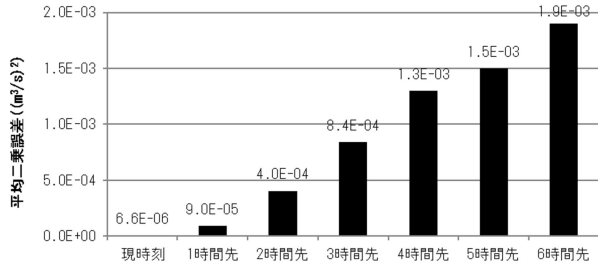


図-16 各予測時間のモデルの最終学習誤差

6567U CPU@3.3GHzのRAM16GBにて、学習時間10分、epoch数100回で計算を切り上げている。出力層となる全結合層には活性化関数を繋げていない。一例として、現時刻における最適化計算（後述するパターン3の計算結果）の学習誤差収束状況を図-15に示すと共に、図-16に各予測時間の100epoch目の学習誤差を示す。

## V. センサー数とダム流入量予測精度の関係

本章では、流域内に複数センサーを設置した結果、どの程度ダム流入量予測精度が向上するかを確認するため、センサー数と予測精度の関係、そして、センサー位置と予測精度の関係を示す。具体的には流出系統や流域面積を参考にセンサー位置を決め、センサーを増やしていく場合のダム流入量予測精度との関係を確認する。想定するセンサー設置パターンを表-5に示す。既存の3つの水位観測所（大山橋、大室橋、両国橋）はセンサー位置として考慮する。表-5のパターンに従い、52出水学習した後に3出水のダム流入量予測精度をパターン毎に比較する。LSTMアーキテクチャー、計算条件等は同一とし、比較指標は以下のRMSEを用いる。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (h_c(i) - h_p(i))^2}$$

$N$  : 総計算ステップ数、 $h_c(i)$  :  $i$ 時の計算値（観測値扱い）、 $h_p(i)$  :  $i$ 時の予測値

表-5 計算パターン

センサー地点番号	①	②	③	④	⑤	⑥	⑦	⑧	備考
本川	●	●	●	●	●	●	●	●	
右支川		●		●		●			
左支川									
備考	三ヶ瀬川		大山橋	室久保川	大室橋	神ノ川	両国橋	予測対象	
パターン1							●	●	
パターン2			●					●	
パターン3			●		●		●	●	現状（基本）
パターン4	●		●		●		●	●	基本+本川増
パターン5		●	●	●	●	●	●	●	基本+支川増
パターン6	●	●	●	●	●	●	●	●	最大限
パターン7		●		●		●		●	支川のみ

RMSEは時系列の計算値と予測値の平均偏差を示し、正規化されておらず異なる出水間での精度比較には適さないが、算定式が簡便であり理解しやすいため用いた。

### 1. 到達時間と予測精度の関係

河川は上流から下流に流れ、流路延長が長くなるほど上流から下流への到達時間は長くなる。つまり、上流側のセンサー情報を活用して下流側の水位・流量予測をする場合、上流側センサー位置によって下流側への到達時間が異なる。その結果、予測のリードタイムが異なる。

具体的にB地点からA地点に流下する河川区間を考える。この区間には流入河川が無いと仮定し、B地点からA地点への洪水時の到達時間は1時間とする。A地点でリアルタイムの水位観測をする際には、A地点のセンサーを用いることが最も合理的である。また、A地点で1時間先の水位予測をする際には、B地点のセンサーを用いることが最も合理的である。この観点で、図-17のパターン1とパターン2を確認する。パターン1とパターン2は、ダム上流域に水位センサーを1つだけ設置することを想定し、下流（両国橋）、上流（大山橋）にセンサーを設置した場合の計算結果を示している。なお、前記の通り、センサー位置は分布型流出解析モデル上で設定し、当該地点の流量計算結果をセンサー値に見立てる。現時刻のダム流入量予測の精度が高いのは最もダム流入地点に近い両国橋地点データを活用したパターン1の予測である。また、1時間先、2時間先のダム流入量予測精度が高いのは上流にある大山橋地点データを活用したパターン2の予測である。

以上の整理より、単純に上流から下流に流れる河川条件を想定すると、予測のリードタイムに適した位置の上流側センサーを利用することがダム流入量予測の精度向上に繋がると考えられる。例えば、洪



水到達時間を最大限に活用して予測のリードタイムを確保するためには、より上流側にセンサーを設置すれば良い。しかし、一般的に上流側のセンサーほど流域負担面積が小さく下流側の水位・流入量予測への情報提供力が小さい。また、到達時間（＝リードタイム）は流量規模により異なり、雨の降り方により上記関係は複雑になる。従って、予測のリードタイムを想定して特定位置にセンサーを置くとその効果は限定的となる。

以上のことより、流域内の様々な事象に対応してダム流入量予測精度を高めるためには、多様なデータを取得するため複数のセンサーを設置し、このデータを適切に分類できることが重要である。

## 2. センサー数と予測精度の関係

次に複数センサー設置の効用について示す。センサー数と予測精度の関係を示すため、図-17のパターン1～4までの計算結果を確認する。

パターン3は、既往の3つの水位観測所（大山橋、大室橋、両国橋）を全て活用してダム流入量を予測する。パターン3は、現時刻予測ではパターン1よりも精度が悪く、1時間先と2時間先予測ではパターン2よりも精度が悪い。これより、パターン3は、パターン1とパターン2の中間の特性を有することが分かる。このように複数データを重ねることによって一方のデータ特性に偏らない汎化性の高いモデルを構築することができる。一方で、どっちつかずの分類精度

が低いモデルが出来ることにも繋がる。しかし、多様な学習データを増やすことで分類精度を高めることができる。パターン4は、パターン3から更に本川筋のセンサー数を増やしたものであり、流域面積の大きい支川の合流直後にセンサーを設置しているところに特徴がある。パターン4は、現時刻の予測精度はパターン3と大きく変わらないが、1時間先、2時間先の予測精度はパターン3よりも向上している。これは本川とは異なる支川の情報を加味できているためと考えられる。同様の効果が図-18のパターン5でも見られる。パターン5は、パターン3から支川筋のセンサー数を増やしたものであり、パターン4と同様に本川とは異なる支川の情報を加味できていると考えられる。言ってみれば、パターン4とパターン5は同様のデータ特性を有すると想定される。図-11で確認すると、パターン3から追加されたセンサー位置もパターン4とパターン5で近い所にある。

一方で、図-18のパターン6のようにセンサーを増やしてもパターン5と比較して予測精度が向上していない例もある。これは、特性の類似したデータを重ねても予測精度を向上させることが難しいことを示唆している。更に取得データの類似性の観点から予測精度向上のためのセンサー設置数には上限があることも想定される。

以上より、センサー数を増やし流域情報を増やすことは予測精度向上に重要である。しかし、類似した情報を重ねた場合、偏った分類となる恐れがあり、

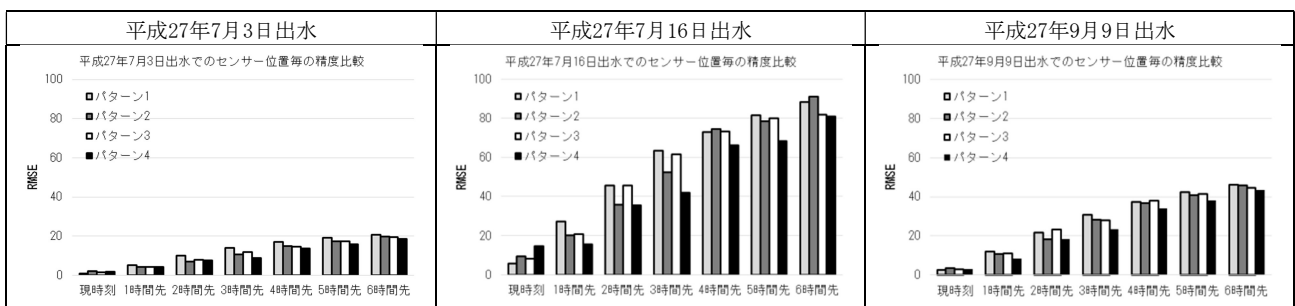


図-17 パターン毎の精度検証

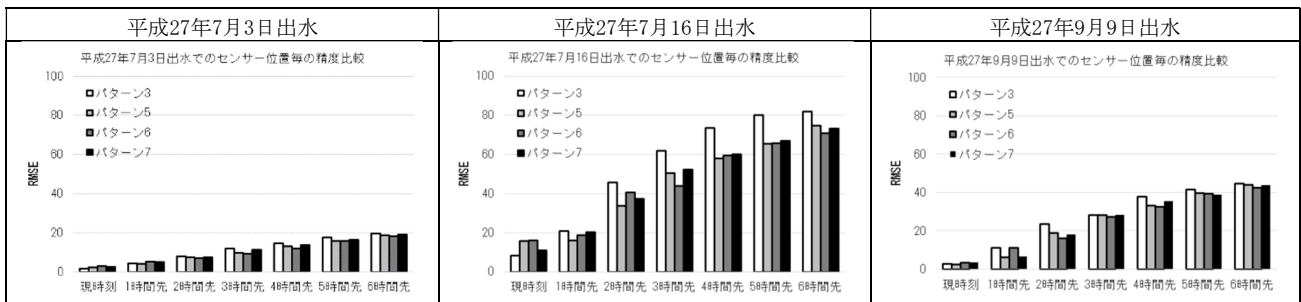


図-18 パターン毎の精度検証

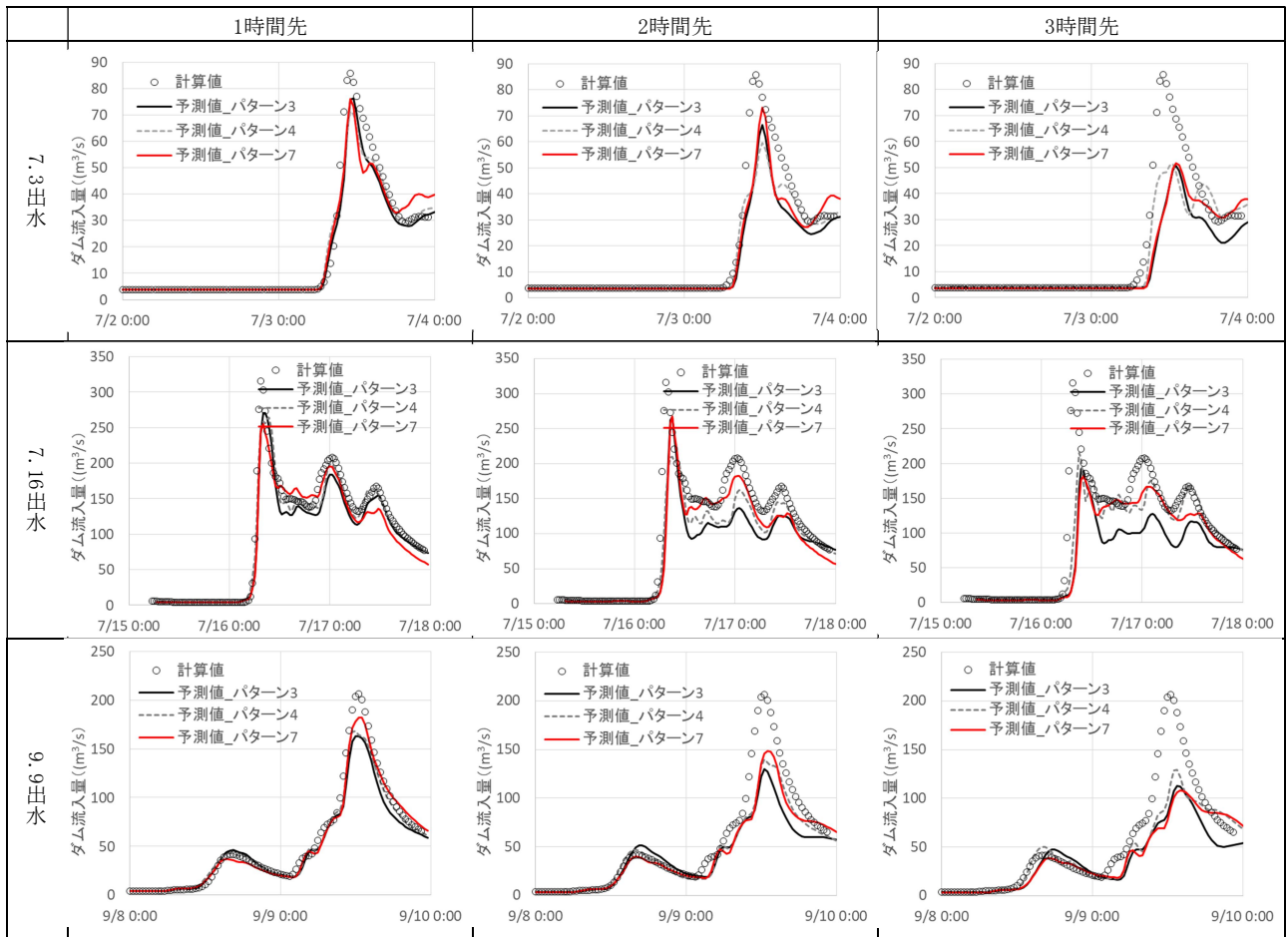


図-19 センサー設置によるダム流入量予測精度の違い

それ故にセンサー設置数に上限があること、また、資金にも限りがあること等から、本稿のようなシミュレーションを進めて、適切なセンサー配置を議論することは非常に重要である。

### 3. センサー位置と予測精度の関係

山間部は谷が深く、急流である等から、本川筋にセンサーを設置できないことも多い。そこで、支川の橋梁に超音波式・電波式センサーを設置し、支川水位を観測することになる。超音波式は電圧が小さく手軽に扱えるが走査深度が小さく、小さな河川・溪流でしか利用できない。一方、電波式は走査深度が大きく、谷が深い山間部橋梁への設置に適するが、電圧が大きく電源確保が大きな課題となる。

上記特性も想定し、本川と支川にセンサーを設置する場合にて予測精度への影響を確認する。パターン3を基準に本川筋主体にセンサーを増やした場合の予測精度を図-17のパターン4に、支川筋主体にセンサーを増やした場合の予測精度を図-18のパターン5、7に示した。ただし、左支川は流域負担面積が小さいためセンサーを設定しなかった。これらの結

果より、①支川筋センサーは本川筋センサーの代替になり得る、②平成27年7月16日出水のように出水現象が複雑な場合、寧ろ支川筋にセンサーを付けた方が精度向上に繋がる、ことが確認できる。本川筋だけでなく支川筋をセンサー設置対象とすることでデータの多様性が増し、予測精度向上に寄与すると考えられる。この結果、センサー数が増えることになるが、今後のIoT技術の高度化・安価化により流域情報のデジタル化も徐々に進むと考えられる。

最後に、センサー数やセンサー位置による予測精度向上程度を図-19に示す。代表的な計算結果として、パターン3の現状観測体制、パターン4の現状観測体制に本川センサーを追加した観測体制、パターン7の現状観測体制を反映させず支川センサーのみの観測体制、を整理した。特に平成27年7月16日のような複雑な出水現象の場合、パターン3とパターン4のように初期の観測体制からセンサー数を増やすことで流量波形の再現性・予測精度が高まることが確認できる。また、同じ出水において、パターン3とパターン7を確認すると、2時間先、3時間先の精度の違いが大きく、センサーの位置を工夫すること

で予測精度が高まることが確認できる。支川条件は流域により異なるため、本稿のような試みによりセンサー位置を決定することは非常に有意義である。

## VI. おわりに

本稿では、流出解析モデルの計算結果を実流域の水位センサーデータに見立て、センサーの数や位置とダム流入量予測精度との関係を検討した。本稿で取り扱ったデータは、現場で測る水位データと特性上大きな違いがあるが、流域デジタルデータを増やすことによるダム管理の効率化・適正化等への有効な検討材料になると考えられる。更に、本稿の考え方は、分布型流出解析モデル等とデータ同化の組み合わせにも適用でき、例えば、どこにセンサーを設置することで分布型流出解析モデルによる予測精度を向上させることができるかといった問いに答えることが可能になると考えている。

なお、本稿では、ダム流入量予測に関わる Digital Twin の大枠や有効性を主に示し、上流域水文量から下流域水文量を予測する手法としての LSTM 利用の妥当性については紙面の都合上、詳述できていない。データ量（洪水数、水位観測所データのみ利用）や流域規模（長期データ依存性やデータ非線形性）から本研究対象流域では複雑なデータ関係を表現できる LSTM 選択の優位性は小さく、RNN や 3 層のニューラルネットワークと比較して特筆すべき精度向上は見られないことを別途発表により示している（天方ら、2019a）。ただし、この結論もデータ関係が複雑となる大規模流域では異なると考えており、別流域での研究も鋭意進めていきたい。

ここでは、本稿の課題についてまとめる。

- 1) 道志ダム流域は100km<sup>2</sup>程度の中小流域であり、比較的水文過程が単純な流域である。そこで、単純な LSTM アーキテクチャーを構築したが、道志ダム流域よりも大規模で複雑な流出機構を有する流域を扱う場合、水文現象も複雑となり、より多くのデジタルデータが必要になると共に、より複雑な LSTM アーキテクチャーが必要になると考えられる。
- 2) 本稿では、実観測水位データの代わりに分布型流出解析モデルの計算結果を活用している。実観測水位は水面振動、観測誤差等のノイズを含

み、データ特性が複雑になる。このため、実観測水位を用いて構築した LSTM モデルは、本稿で構築したモデルよりも予測精度が悪くなる（天方ら、2019b）。これを回避するため、実観測水位データを前処理してノイズを軽減する、モデルのバイアスとバリエーションの関係に留意し、正則化項を含めて LSTM アーキテクチャーを構築する等が考えられる。

- 3) 本稿では、高水データのみを学習し、人工知能モデルを構築した。一方、深層学習モデルのデータ特徴量を抽出する表現力を考慮すると、低水データも含めたモデル構築を行うことで予測精度は向上する可能性がある。低水データと高水データのデータ量の違いに着目した学習条件等に留意し、モデル構築を進めていく必要がある。
- 4) 状態方程式に粒子フィルターを適用する場合、現時刻から1ステップ先の予測精度は良好である。一方、LSTM のような相関モデルの場合、上下流観測地点間の洪水到達時間分の予測精度は良好である。しかし、両手法とも観測データとの関係が途絶える一定先の予測では精度確保が難しく、実運用上の大きな課題である。不確定性の大きい予測雨量（気象庁、2019）も加えて、実運用に耐えうる水位・流量予測の精度向上策の研究を進めていく必要がある。

今後は、上記4つの課題を克服するための研究を進め、ダム流入量予測精度向上に寄与する研究を進めていきたい。

## 引用文献

- 天方匡純・藤井純一郎・安野貴人・大久保順一・嶋本ゆり 2019a. ダム流入量予測へのRNN系モデルの適用. 日本オペレーションズ・リサーチ大会 2019年 春季研究発表会.
- 天方匡純・安野貴人・藤井純一郎・嶋本ゆり・大久保順一 2019b. データ非線形性とダム流入量予測精度に関する研究. 2019年度 人工知能学会全国大会（第33回）. DOI: 10.11517/pjsai.JSAI2019.0\_2Q1J203.
- Greff K, Srivastava RK, Koutnik J, Steunebrink BP, Schmidhuber J. 2017. LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems 28(10): 2222-2232. DOI: 10.1109/TNNLS.2016.2582924.

- Graves A, Schmidhuber J. 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks* 18(506): 602-610. DOI: 10.1016/j.neunet.2005.06.042.
- 一言正之・櫻庭雅明・清雄一 2016. 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: I\_187-I\_192 . DOI: 10.2208/jscejhe.72.I\_187.
- Hochreiter S, Schmidhuber J. 1995. Long Short-Term Memory. Technical Report FKI-207-95. Fakultät für Informatik. Technische Universität München.
- Hochreiter S, Schmidhuber J. 1997. Long Short-Term Memory. *Neural Computation* 9(8): 1735-1780. DOI: 10.1162/neco.1997.9.8.1735.
- 環境省・気象庁・国土交通省・農林水産省・文部科学省 2018. 気候変動の観測・予測及び影響評価統合レポート2018 ～日本の気候変動とその影響～. 環境省地球環境局: 37.
- 気象庁 2019. “降水短時間予報の精度について”, 降水短時間予報の精度検証結果, [https://ds.data.jma.go.jp/fcd/yoho/kotan\\_kensho/kotan\\_hyoka.html](https://ds.data.jma.go.jp/fcd/yoho/kotan_kensho/kotan_hyoka.html) (参照:2019/07/10) .
- 工藤亮治・近森秀高・永井明博 2009. 粒子フィルタを用いた河川流域における実時間洪水予測. *農業農村工学会論文集* No. 259: 17-25. DOI: 10.11408/jsidre.77.17.
- Li S, Li WQ, Cook C, Zhu C, Gao YB. 2018. Independently Recurrent Neural Network (IndRNN): Building a longer and deeper RNN. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition: 5457-5466. DOI: 10.1109/CVPR.2018.00572.
- Saddik A. El. 2018. Digital Twins: The Convergence of Multimedia Technologies. *IEEE MultiMedia* 25(2): 87-92. DOI: 10.1109/MMUL.2018.023121167.
- 関基・木村晃・古山一志・和田高宏・金子祐・穴水秀樹・眞間修一 2013. 河川特性を反映したニューラルネットワーク洪水予測の精度向上. *河川技術論文集* 第19巻: 295-300.
- 立川康人・須藤純一・椎葉充晴・萬和明・キムスンミン 2011. 粒子フィルタを用いた河川水位の実時間予測手法の開発. *水工学論文集* 第55: S551-S516. DOI: 10.2208/jscejhe.67.I\_511.
- 高崎忠勝・河村明・天口英雄・石原成幸 2012. 粒子フィルタを用いた都市貯留関数モデルによる実時間洪水流出予測特性. *土木学会論文集B1 (水工学)* Vol. 268 No. 4: I\_511-I\_516 . DOI: 10.2208/jscejhe.68.I\_511.
- 竹村仁志・眞間修一・吉田武司・石上尚 2013. 中流域に洪水調節ダムを有する河川における洪水予測システムの精度検証. *河川技術論文集* 第19巻: 247-252.
- 吉野文雄・吉谷純一・堀内輝亮. 1990. 分布型流出モデルの開発と実流域への適用. *土木技術資料* 32(10): 54-59.