

論文

Gradient Boosting の適用によるダム流入量予測の精度向上

天方 匡純¹ 藤井 純一郎² 梁田 信河²Accuracy Improvement of Dam Inflow Discharge Forecasting by Applying Gradient Boosting ,
Yachiyo Engineering Co.,Ltd.

Masazumi AMAKATA Junichiro FUJII Nobuka YANADA

近年、低頻度の外力現象の多発により、多くの水害・土砂災害が発生している。これらの被害を最小限に留めるべく、治水能力を有するダムは貯水機能を最大限に活用するため、事前放流操作や異常洪水時防災操作といった規定に基づき運用される。しかし、更なる効果的な貯水池運用を目指す場合、ダム流入量の予測精度向上は不可欠である。本稿では、Gradient Boosting Decision Tree と呼ばれるアンサンブル機械学習手法をダム流入量予測スキームに適用し、ニューラルネットワークモデルと比較して高い予測精度を確保できることを示す。

キーワード：ダム流入量予測、勾配ブースティング、アンサンブル、深層学習

1. はじめに

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、低頻度の外力現象の頻発により、多くの水災害・土砂災害が発生している。

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・安心をもたらしている。ただし、これらの運用はダム管理者の弛まぬ管理行為と豊富な経験に支えられたものであり、前述の低頻度外力の増大や団塊世代の引退等に伴う人員不足により、徐々に当たり前の光景では無くなりつつある。これは、現場において管理機能低下の危機感を募ら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が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や流量を水位に変換する河道水位追跡計算から構成される（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、既往出水に対してモデルを検証した上でパラメータを最適化する。しかし、取り扱うデータやモデルには不確実性が数多く含まれ、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る。これらの予測精度向上が水防上、そして、施設管理上の大きな課題である。

この不確実性に対応するためにダム流入量等を対象

¹ 八千代エンジニアリング株式会社 技術開発研究所、所長

² 八千代エンジニアリング株式会社 技術開発研究所、上席研究員

³ 八千代エンジニアリング株式会社 北日本支店、支店長

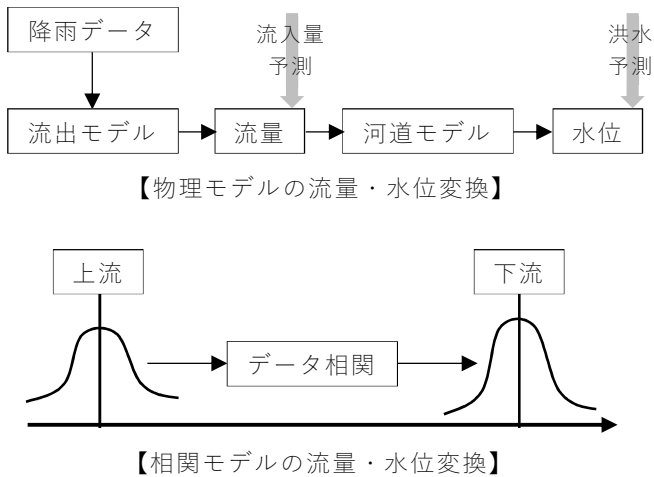


図-1 流量・水位算定方法

とした研究として、古くは単位図、貯留関数、タンクモデル等にカルマンフィルターを適用した事例^{1,2)}、近年では分布型流出モデルにカルマンフィルターを適用した事例³⁾等がある。しかし、カルマンフィルターは過去の観測値からモデル変数を最適化するため、雨量とダム流入量等の関係が急激に変化する際に予測精度が悪くなる等の問題があった。そこで、近年では粒子フィルターと呼ばれる現時刻観測値に状態方程式の変数を適合させ、現時刻から1ステップ先の予測精度を向上させる手法^{4,7)}が全国の一級河川等で取り入れられようとしている。この手法により現時刻、そして、1ステップ先の予測精度は飛躍的に向上することが期待される。

一方で、水文学的プロセスを追跡せず、データ間の繋がりから予測を目指す手法がある。これらは、雨量や上流側水位等の上流域水文学量から下流域水文学量を推定する方法(図-1 下段)であり、単相関、重相関、ニューラルネットワークといった手法がダム流入量予測⁸⁻¹⁰⁾、洪水予測^{11,12)}で適用されている。

本稿では、上流域水文学量から下流域水文学量を推定する手法として、機械学習のアンサンブル学習手法の一つである Gradient Boosting (勾配ブースティング) を活用する。Gradient Boosting は、教師データに対してモデルパラメータを決めていく教師あり学習を通して、どのような問題に対しても一番性能が良い万能なアルゴリズムは無いとする「ノーフリーランチ定理」に基づ

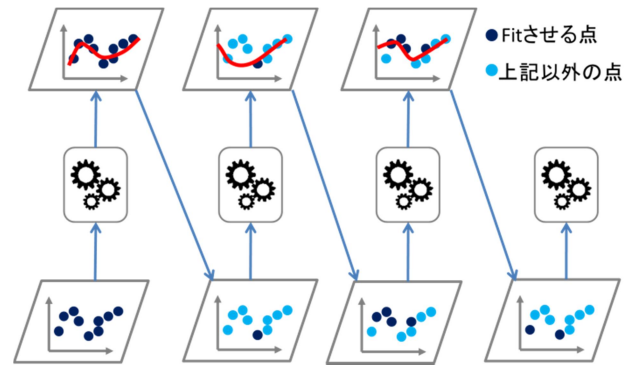


図-2 Boosting の sequential training のイメージ図

(Hands on Machine Learning with scikit-learn and Tensorflow に追記)

き性能の低い複数の識別器を組み合わせることで全体で識別性能を上げる手法(アンサンブル学習手法)の一つである。アンサンブル学習手法は、主に、バギング法とブースティング法が存在するが、Gradient Boosting は名前が示す通り、ブースティング法に分類される。

2. Gradient Tree Boosting の概要

ブースティング法は、幾つかの弱識別器を組み合わせ、より強い識別器とするアンサンブル学習手法の一つである。ブースティング法は、図-2 に示すように、識別器を継続的に順番に学習して、前に学習した識別器が残した教師データとの残差を次の識別器で補正しながら学習精度を高めていく方法である。そのブースティング法の中でも、本稿では Gradient Boosting と呼ばれる手法を活用する。また、この中で用いる弱識別器は決定木とする。決定木は分類問題での活用が良く知られているが、回帰問題にも活用できる。この決定木を活用した Gradient Boosting を Gradient Boosting Decision Tree (GBDT) と呼び、更に、そのアルゴリズムを高速化したものを XGBoost (eXtreme Gradient Boosting) と呼ぶ。本稿では、この XGBoost を活用し、ダム流入量予測の精度向上を図る。

k 個目の決定木を関数 f_k とする。そこで、入力 x_i に対する出力値(予測値) \hat{y}_i との関係は以下で記される。

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad \dots (1)$$

また、損失関数 $l(a, b)$ ($a = y_i$ (観測値), $b = \hat{y}_i$ (予測値)) を導入して、(2)式を定式化し、これ

$$\begin{aligned}
\min_{f_t} \sum_{i=1}^n l(y_i, \hat{y}_i^{(t)}) \\
&= \min_{f_t} \sum_{i=1}^n l\left(y_i, \sum_{k=1}^t f_k(x_i)\right) \\
&= \min_{f_t} \sum_{i=1}^n l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) \\
&\quad \dots (2)
\end{aligned}$$

を最小化する f_t を構築する。しかし、(2)式のみだと決定木の特性の影響を大きく受け、過学習が発生し、予測精度が低下してしまう可能性が高まるため、正則化の働きをする以下ペナルティ項を追加する。ここで、 T は決定木を構築したときの最終ノードの数（木の大きさ）であり、 ω は決定木 f_t が返す最適化された重みベクトルである。また、 γ , λ はそれぞれ、木の大きさ、重みの大きさを制御する値である。

$$\begin{aligned}
\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \\
\quad \dots (3)
\end{aligned}$$

そして、(2)式に(3)式を加えたものを $L^{(t)}(f_t)$ とする。

$$\begin{aligned}
\min_{f_t} L^{(t)}(f_t) &= \min_{f_t} \sum_{i=1}^n l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \\
&= \min_{f_t} \sum_{i=1}^n l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \\
&\quad \dots (4)
\end{aligned}$$

木の大きさ T を大きくすると過学習傾向が高まるため損失関数 l は小さくなり、目的関数 $L^{(t)}(f_t)$ も小さくなる一方で、 γT の部分は大きくなり、目的関数 $L^{(t)}(f_t)$ が大きくなるように機能する。結果として、 γ の値が大きければ大きいほど T の値が小さい、つまり、小さな木が好まれるようになる。 λ は ω の大きさを調整するパラメータであり、 λ を大きく設定すると、小さな ω が好まれるようになる。

更に、高速に最適化を進めるために、(4)式の f_t に関して2次のテーラー展開を実施する。

$$\begin{aligned}
\min_{f_t} L^{(t)}(f_t) &= \min_{f_t} \sum_{i=1}^n l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \\
&\approx \min_{f_t} \sum_{i=1}^n \left(l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \Omega(f_t)
\end{aligned}$$

ここで、 $g_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}$, $h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial^2 \hat{y}_i^{(t-1)}}$ である。最

適化に関係のない項、つまり f_t が関わっていない項を取り除いたものを $\min_{f_t} L^{-(t)}(f_t)$ とすると

$$\begin{aligned}
\min_{f_t} L^{-(t)}(f_t) &= \min_{f_t} \sum_{i=1}^n \left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \Omega(f_t) \\
&= \min_{f_t} \sum_{i=1}^n \left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \\
&= \min_{f_t} \sum_{j=1}^T \left(\sum_{i \in I_j} g_i f_t(x_i) + \frac{1}{2} \sum_{i \in I_j} h_i f_t^2(x_i) \right) + \gamma T \\
&\quad + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \\
&= \min_{f_t} \sum_{j=1}^T \left(\sum_{i \in I_j} g_i \omega_j + \frac{1}{2} \sum_{i \in I_j} h_i \omega_j^2 \right) + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \\
&= \min_{f_t} \sum_{j=1}^T \left(\sum_{i \in I_j} g_i \omega_j + \frac{1}{2} \sum_{i \in I_j} h_i \omega_j^2 + \frac{1}{2} \lambda \omega_j^2 \right) + \gamma T \\
&= \min_{f_t} \sum_{j=1}^T \left(\sum_{i \in I_j} g_i \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right) + \gamma T \\
&\quad \dots (5)
\end{aligned}$$

(5)式を $\frac{\partial L^{-(t)}(f_t)}{\partial \omega_j} = 0$ とすると、以下が求まる。

$$\begin{aligned}
\omega_j^* &= -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \\
&\quad \dots (6)
\end{aligned}$$

(6)式を(5)式に代入すると、以下が求まる。

$$\begin{aligned}
L^{-(t)} &= -\frac{1}{2} \sum_{j=1}^T \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \\
&\quad \dots (7)
\end{aligned}$$

(6)式は決定木構造の質を評価するために利用できる。普通、全ての決定木構造をしらみ潰しに列挙することは不可能であるため、逐次的に新しい木を加えていく。この際、どの決定木構造が良いかの評価は、分岐前と分岐後の(7)式の差分を取り、(8)式で算定する。この(8)式による値が最も大きくなる分割候補を選び、分岐することになる。

解析モデルは土木研究所モデル¹³⁾を用いた。土木研究所モデルは、3種類のタンクで構成される流出解析モデル(図-5)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、土地利用データ(出典：土地利用細分メッシュ図，国土地理院)，土壤データ(出典：土壤分類図，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100mメッシュに加工(図-6)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(1km²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、そのまま入力データとして用いた。また、実観測に伴う誤差も考慮していない。

4.3 出力データ

学習対象とする出力データも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。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、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(表-1)。

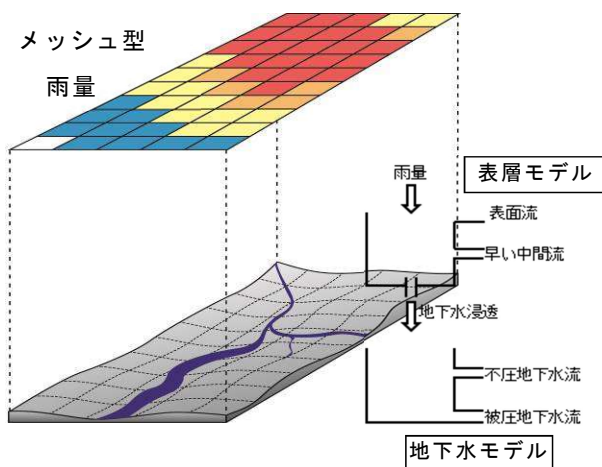


図-5 分布型流出解析モデルの構造

4.4 モデルの構築方法

入力と出力(教師データ)のデータを組み合わせ、XGBoostを介して、その関係性を学習することによりダム流入量予測モデルを構築する。ここでは、データ間の関係性を示し、モデル構築内容を説明する。大山橋(③)、大室橋(⑤)、両国橋(⑦)(各番号は図-6、図-7に示すモデル上の位置を示す)の3観測所の位置に対応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)(番号は図-6、図-7に示すモデル上のダム位置を示す)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時間先までのダム流入量を予測するため、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。

未学習出水に対する検証の際は、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。なお、比較のため、これまで私達が洪水予測モデル等に活用してきた3層のニューラルネットワーク^{11,12)}(以下、3NNと記す)による計算結果も記す。用いた3NNのモデル構造は表-2、図-8のとおりである。

5. 予測計算結果

52個の既往出水データを学習したのち、平成27年の3個の出水(7月3日、7月16日、9月9日)について予測精度を検証した。図-9～図-11の点線が計算値(観測値相当とする分布型流出解析モデルの計算結果)、実線がXGBoost、3NNの予測値である。図-9～図-11は、それぞれ現時刻で入手できる情報を利用して、0時間先、1時間先、2時間先、3時間先を予測した結果を示している。すなわち、0時間先のグラフは、現時刻の情報を入手し、現時刻のダム流入量を再現したものである。また、3時間先のグラフは、3時間前までの情報を入手し、現時刻のダム流入量を再現したものとなる。当然、情報入手と予測の時刻が離れるほど予測精度は劣化する。また、計算値と予測値の乖離度をRMSEで整理し図-12～図-14に示した。RMSEは(9)式より算定した。

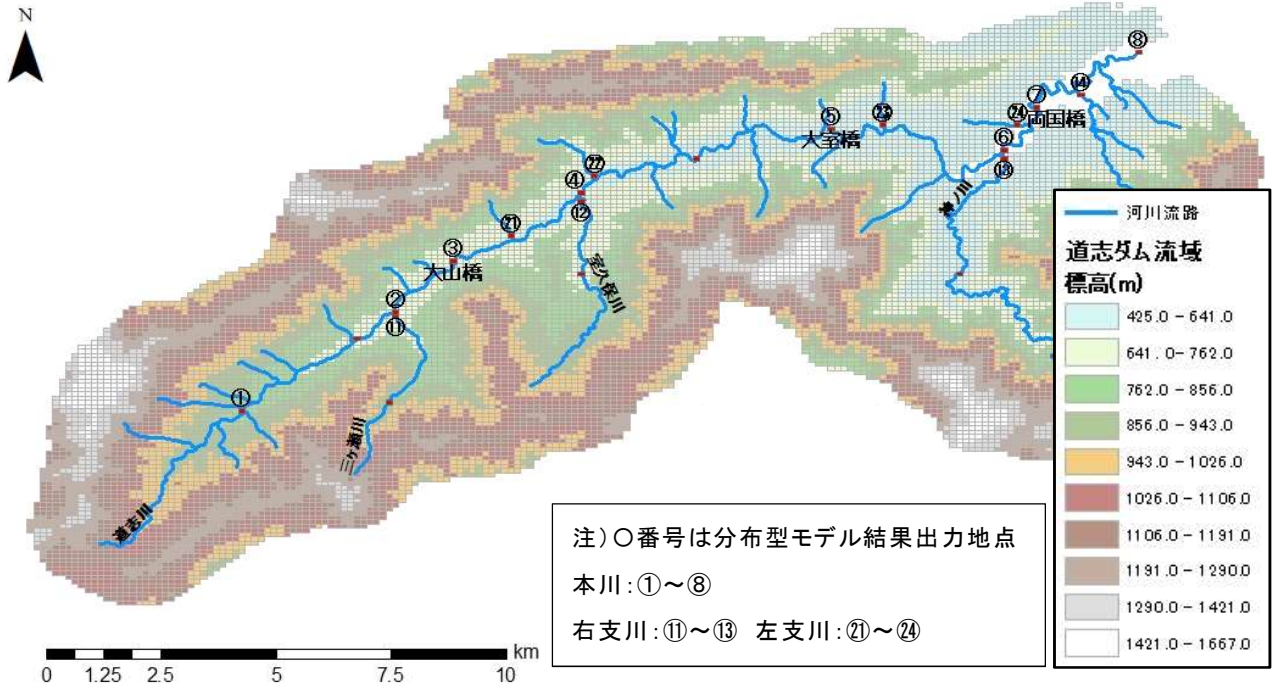


図-6 分布型流出解析モデルのメッシュ図

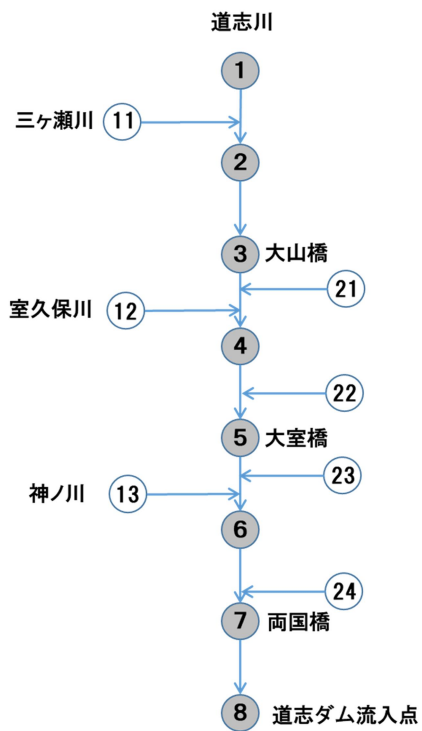


図-7 観測所位置対応図 (図-6 の○番号に対応)

表-2 3層 NN モデル構成表

| Layer | レイヤータイプ | 出力データ数 | パラメータ数 |
|--------------|---------|--------|--------|
| dense_1 | 全結合層 | 7 | 427 |
| activation_1 | 活性化関数 | 7 | 0 |
| dense_2 | 全結合層 | 1 | 8 |
| パラメータ総数 | | | 435 |

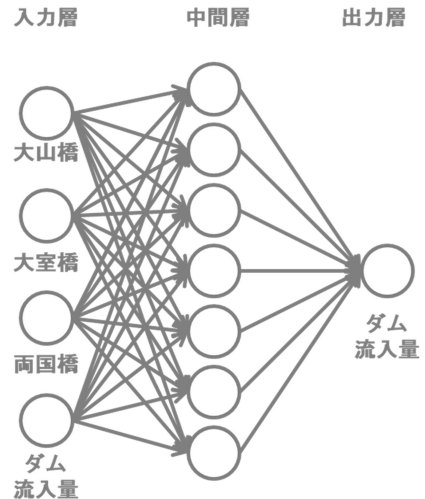


図-8 3層 NN モデル構造

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_c(i) - Q_p(i))^2} \dots (9)$$

N: 総計算数、 $Q_c(i)$: i 時計算値、 $Q_p(i)$: i 時予測値

更に、実運用時の活用を想定し、図-9~図-11 の結果を利用して、現時点から毎時の 6 時間先、3 時間先までの XGBoost のダム流入量予測結果を図-15 に示した。実線が計算値、実線に丸のマーカが追記されている

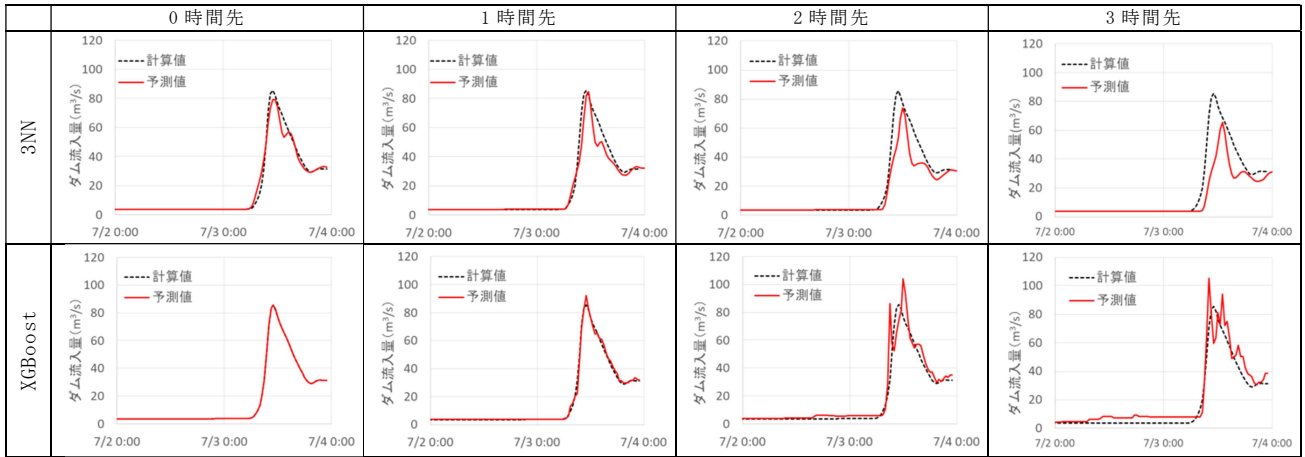


図-9 平成 27 年 7 月 3 日出水に対する各モデルの予測状況

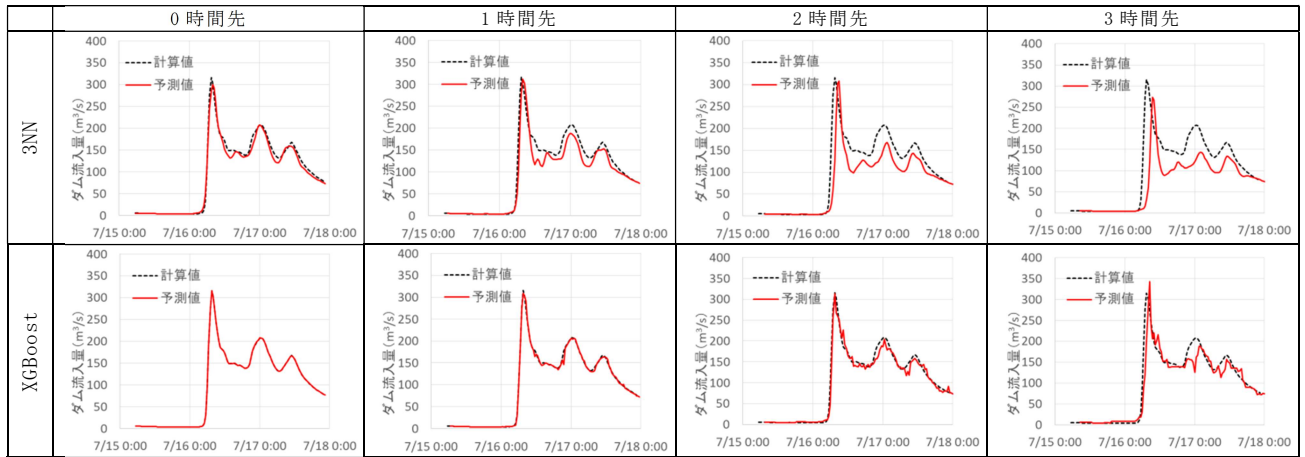


図-10 平成 27 年 7 月 16 日出水に対する各モデルの予測状況

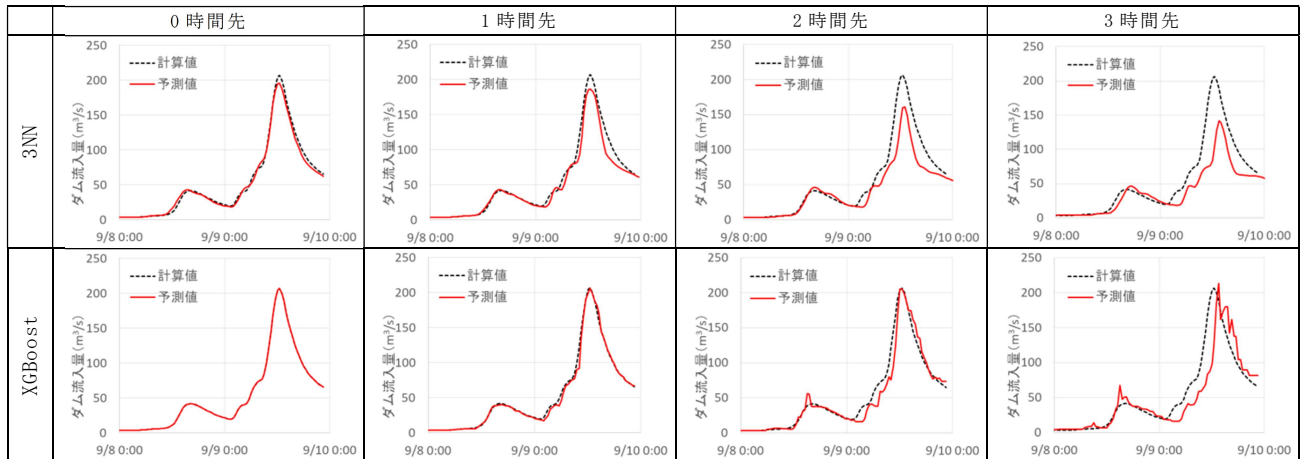


図-11 平成 27 年 9 月 9 日出水に対する各モデルの予測状況

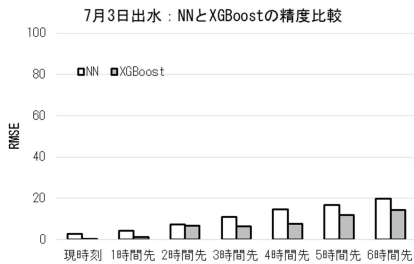


図-12 H27.7.3 出水の RMSE

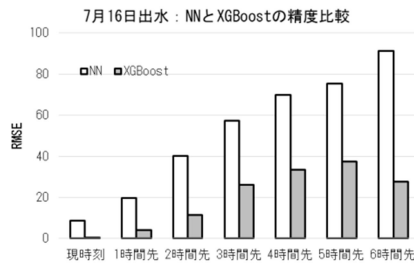


図-13 H27.7.16 出水の RMSE

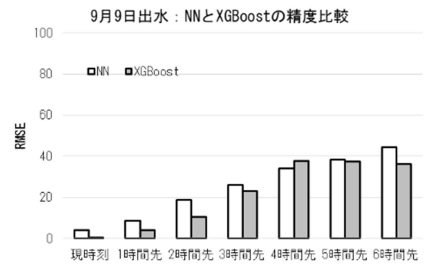


図-14 H27.9.9 出水の RMSE

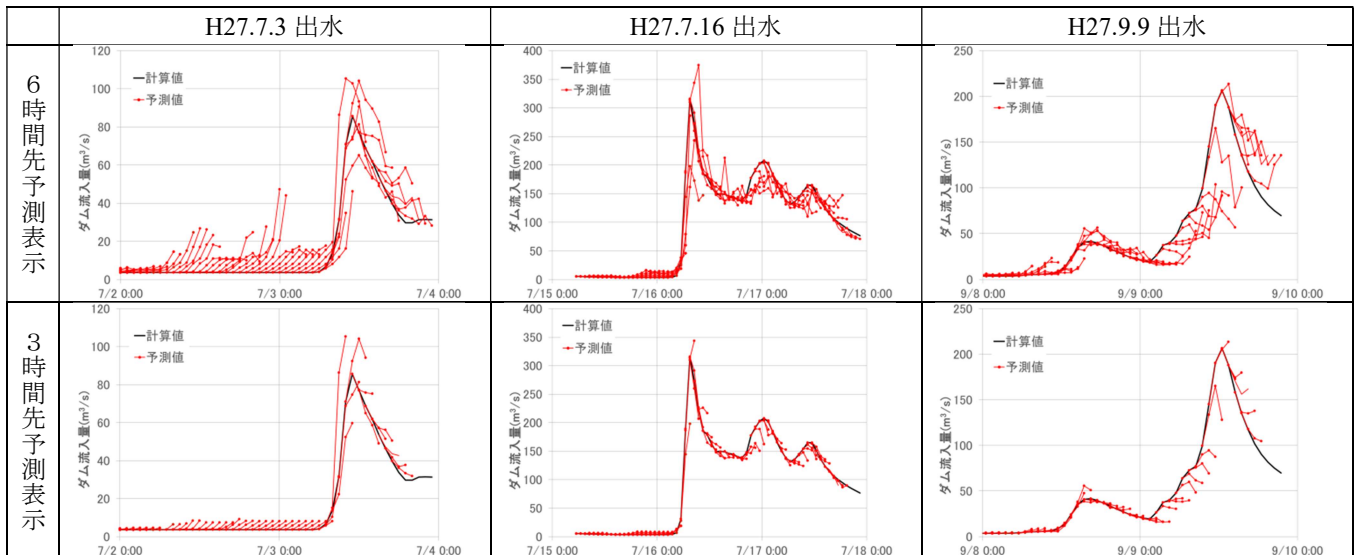


図-15 6時間先, 3時間先の運用を想定した表示 (図-9, 10, 11の表現を変えたもの)

ものが予測値である。本検討対象流域の洪水到達時間を踏まえて予測時間（リードタイム）の信頼度を考慮し、6時間先と3時間先の2種類のグラフを表現した（両者の予測結果は同じもの）。

5.1 平成27年7月3日出水

本研究では、現時刻よりも先の予測は、上流側観測点とダム地点との洪水到達時間差によって担保されている。このため、図-9と図-12から、3NNでは洪水到達時間を超える先の予測となる2時間先以降で予測精度の劣化程度（グラフ形状やRMSE変化）が大きくなる。XGBoostも2時間先以降の精度劣化は大きいですが、予測値と計算値（観測値に見立てた分布型流出解析モデルの計算値）の概ねの hidro形状は類似し、総じて3NNよりもXGBoostの予測精度が良い。また、XGBoostは、2時間先、3時間先のピーク流量付近で予測が乱れながらも、3NNに比較してピーク流量規模の予測精度が高い点が特筆すべきところである。次に、図-15の6時間先、3時間先予測表示を確認する。図-15のグラフでは予測精度が100%の場合、予測値が計算値と重なる。概ねの傾向として、6時間先表示では予測の不確実性が高く、予測値が計算値と大きく乖離するが、3時間先表示では不確実性が低く、予測値が概ね計算値に追随する。これは、洪水到達時間により担保される情報信頼度に依存するためである。ただし、3時間先表示においても、流量が増加から減少に転じるピーク流

量付近は、計算値に比較して予測値が大きめに出てしまっている。これは、当該出水規模が小さく、流域内の流出に局所的偏りが発生し、局所的流出の学習が十分でないためと考えられる。

5.2 平成27年7月16日出水

3つの検証出水のうち最大規模の出水である。図-10と図-13から、3NNとXGBoostの予測精度の違いは平成27年7月3日出水と同傾向である。XGBoostは3NNに比較して、現時刻から6時間先までのあらゆる予測のRMSEが大幅に小さく、総じて3NNよりもXGBoostの予測精度が良い。また、図-15の3時間先表示では、平成27年7月3日出水に比較して、ピーク流量の予測精度が良く、図-10の結果と合わせてXGBoostの優れた特徴と捉えることができる。一方、6時間先表示と合わせて確認しても、ピーク流量発生後減水期の2山目、3山目を低めに予測する傾向がみられる。

5.3 平成27年9月9日出水

前半部に小さいピーク（1山目）があり、後半部に大きいピーク（2山目）がある出水である。図-11から3NNとXGBoost共に、1山目は2時間先も良い予測精度を維持しているが、2山目は形状を崩し、遅れて予測する傾向にある。同じく図-11から2時間先までであればXGBoostは2山目のピーク流量規模を適切に予測している。また、図-14からこの出水でも総じて3NNよ

りも XGBoost の予測精度が良いことが確認できる。一方、図-15 の XGBoost の 6 時間先表示, 3 時間先表示からは, 2 山目の出水に対して全体的に遅れて予測する,あるいは, 流量を低めに予測する傾向が見られる。他 2 出水に比較しても 2 山目の流量上昇部を低めに予測する傾向が顕著である。XGBoost により構築したモデルは他の機械学習モデルと同様にそのモデル特性は教師データ特性に依存する。当該出水にて, ピーク流量付近は汎化性が高く, 流量上昇部は汎化性が低い学習状況が確認されたが, 限られたデータから汎化性をどこまで高められるかは機械学習全般の大きな課題である。

6. おわりに

本稿では, 比較的小規模な流域において, Gradient Boosting と呼ばれるアンサンブル手法を 0~6 時間先のダム流入量予測に適用し, 3 層のニューラルネットワークよりも総じて予測精度が良い可能性を示した。本稿の対象流域の洪水到達時間を考えた場合, 2 時間先の予測が限界と考えられるが, Gradient Boosting は 2 時間先までの波形全体の予測精度の高さのほか, 特にピーク流量の予測精度の高さが特徴的である。数時間先とは言え, 出水時のピーク流量を想定できるのはダム管理者にとって有用と考えられる。更に, 大規模流域, すなわち, 洪水到達時間が大きな流域では, その分リードタイムが大きく取れるため, 本手法の意義も高まると想定され, 大規模流域での検証も進めていきたい。

本稿で紹介した XGBoost 及びニューラルネットワークは, 機械学習の一手法であり, 取り扱うデータセットの規模や形式, 特性等により手法優劣が異なり, 現時点では汎用的に優劣要因を特定することは難しい。一般的には XGBoost は表形式のデータ特徴量を捉えることに優れ, 小さなデータセットやデータ欠損等にも頑強な特徴量を抽出できるとされる。一方, ニューラルネットワーク, 特に深層学習技術は, 時空間データに代表される膨大なデータ量から特徴量を捉えることに優れている。以上より, 本稿で扱った表形式データが XGBoost に適していること, また, 画像等に比較して小さいデータ量であること等が XGBoost の効果を発揮させている要因と考えられる。なお, 同じ流域にて, 深層学習技術 (LSTM: Long Short-Term Memory) を用

いた検討も進め, 3 層のニューラルネットワークと同等の予測精度であることを確認¹⁴⁾している。

今後は, 様々な流域で XGBoost を適用し, 有効範囲(流域特性や出水特性)を確認すると共に, 未経験出水への対応可能性を確認し, XGBoost の汎化性を見極めていく必要がある。低水への適用可能性等も確認していく必要がある。

また, 本稿では, 分布型流出解析モデルの計算結果を活用し, 言わば, デジタル空間上に仮想流域を作成し, 上下流の水文関係を XGBoost により構築した。一方, 実際の流域では, 出水時の河道内では激しい水面振幅があり, ダム貯水池でも激しい水面振幅が発生する。このように扱うデータに様々な雑音が含まれる場合, ダム流入量予測の精度悪化に与える影響は大きい¹⁵⁾。観測データを扱う際の取得データの雑音の扱いは, ダム流入量予測精度向上に向けた損失関数 $l(a, b)$ や正則化項の工夫等により別途検討していく必要がある。

参考文献

- 1) 日野幹雄: 水文流出系予測へのカルマン・フィルタ理論の適用, 土木学会論文報告集, No.221, 39-47, 1974
- 2) 角屋睦, 田中丸治哉: 長短期流出両用モデルによる実時間洪水予測, 農業土木学会論文集, No.177, 31-41, 1995
- 3) 藤田暁, 大東秀光, 上坂薫, 椎葉充春, 立川康人, 市川温: 分布型流出モデルに基づくダム流入量実時間予測モデルについて, 水工学論文集, 第 45 巻, 115-120, 2001
- 4) 立川康人, 須藤純一, 椎葉充晴, 萬和明, キムスンミン: 粒子フィルター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 巻, 511-516, 2011
- 5) 工藤亮治, 近森秀高, 永井明博: 粒子フィルターを用いた実時間洪水予測システムの構築と河道任意地点における予測精度の検証, 水文・水資源学会誌, 第 24 巻, 第 3 号, 137-148, 2011
- 6) 高崎忠勝, 河村明, 天口英雄, 石原成幸: 粒子フィルター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.48, No.4, 511-516, 2012
- 7) 天方匡純, 梁田信河, 渡邊優: 粒子フィルターを活用したダム流入量予測に関する研究, 電力土木, No.389, 16-22, 2017
- 8) 天方匡純, 梁田信河, 田代克浩: SOINN (自己増殖型ニューラルネットワーク) を活用したダム流入量予測に関する研究, 電力土木, No.389, 9-15, 2017

- 9) 井本文雄, 中村正俊: 雨量情報に基づく数時間先行のダム流入量のオンライン予測法, 電学論 B, 113, 3, 213-220, 1993
- 10) 瀬野英二, 村上研二, 泉田正則, 松本普: 広域雨量データを用いたニューラルネットワークによるダム流入量予測, 電学論 B, 124, 4, 561-568, 2004
- 11) 竹村仁志, 眞間修一, 吉田武司, 石上尚: 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, 2013
- 12) 関基, 木村晃, 古山一志, 和田高宏, 金子祐, 穴水秀樹, 眞間修一: 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, 2013
- 13) 吉野文雄, 吉谷純一, 堀内輝亮: 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, 1990
- 14) 天方匡純, 藤井純一郎, 安野貴人, 大久保順一, 嶋本ゆり: 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, 春季研究発表会, 2019
- 15) 天方匡純, 安野貴人, 藤井純一郎, 嶋本ゆり, 大久保順一: データ非線形性とダム流入量予測精度に関する研究, 人工知能学会全国大会, 第 33 回, 2019

(2019 年〇月〇日 受理)

The strong rainfall which we have never assumed in the river design happens frequently and many water and sediment disasters happen. In order to minimize these damages, the dams which are able to control the flood damages are managed to maximize the reservoir utilization based on the regulations of Pre-discharge operation and Disaster prevention operation in the time of abnormal flood. But when we are going to aim at maximizing the reservoir utilization more effectively, it is indispensable to improve the accuracy of dam inflow forecasting. In this thesis, we will show that the ensemble machine learning method called Gradient Boosting is much more accuracy than the neural network method when we use them for the dam inflow forecasting scheme.

Key words: dam inflow forecasting, gradient boosting, ensemble, deep learning