

ダム流入量予測への RNN 系モデルの適用

八千代エンジニアリング株式会社 技術創発研究所 *天方 匡純 AMAKATA Masazumi
 八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 藤井 純一郎 FUJII Junichiro
 八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 安野 貴人 YASUNO Takato
 八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 大久保 順一 OKUBO Junichi
 八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 嶋本 ゆり Shimamoto Yuri

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，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。

高水流量と低水流量に大きな差がある我が国にて，水運用の平滑化を目指し建設されたのがダムである。ダムの最大貯水容量は柔軟に変えられないが，運用の工夫により平滑化のポイントを変えることができる。現在も事前放流操作や異常洪水時防災操作といった運用規定により私達の生活に安全・安心や潤いをもたらしている。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，短時間強雨の発生頻度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無くなりつつある。

代表的なダム運用支援ツールとして，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，雨量を入力条件とし，流量を算定する流出解析モデルから構成される（図 1 上段参照）。

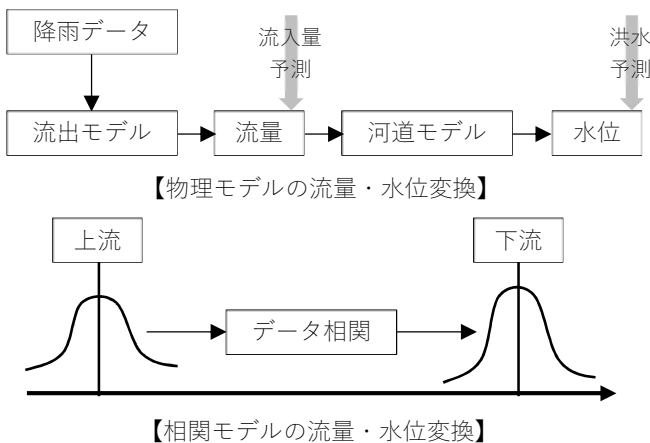


図 1 ダム流入量予測モデルの仕組み

これらのモデルは，既往洪水に対してモデル検証した上でパラメータ確定するが，各過程やモデルには入力・モデル条件等の不確定性を数多く抱え，リアルタイムでの流入量再現も難しい[1]。

そこで，上流水文量から下流水文量（ダム流入量）のデータ関係のみでダム流入量を予測する手法（図 1 下段参照）として，機械学習の一手法である DNN（Deep Neural Network）を活用する。各分野で適応される DNN 構造は様々であり，自然言語処理のように前後の言葉（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る。本稿では，RNN 系のモデルである simpleRNN，LSTM，GRU の 3 手法についてダム流入量予測への適用性を確認する。

2. RNN (Recurrent Neural Network)

RNN は図 2 の通り，時刻 $t-1$ 以前の hidden layer 情報を時刻 t の hidden layer に継承することができ，構造上は相当過去の情報も時刻 t に継承することができる。しかし，back propagation through time の勾配消失問題の影響を受け，数ステップ前の hidden layer 情報しか保持できない Short-Term Memory のネットワークとなっている。

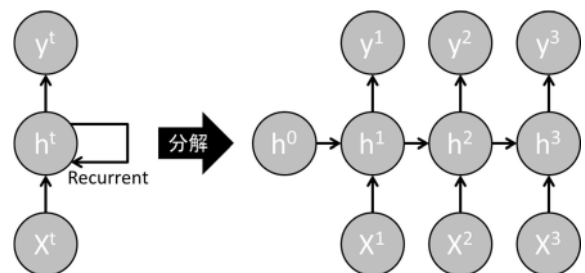


図 2 simpleRNN の特徴

3. LSTM (Long Short-Term Memory)

LSTM は、RNN の勾配消失問題を解決するべく、LSTM ブロックと呼ばれる3つのゲート(忘却ゲート, 入力ゲート, 出力ゲート)を有するネットワーク構造(図4参照)を抱え、Long-Term Memory を実現している。忘却ゲートは時刻 $t-1$ 以前の hidden layer 情報を次ステップに繋げるかを決定する。入力ゲートは時刻 t の入力情報をセルの更新情報として利用するかを決定し、出力ゲートはセルの更新情報を hidden layer 情報として次のステップに継承するかを決定する。

4. GRU (Gated Recurrent Unit)

GRU は LSTM を簡便化した構造(図5参照)であり、特性は LSTM と類似している。精度も LSTM よりも良いわけではなく、取り扱うデータによって良し悪しがある。変数の数は LSTM よりも GRU の方が少ないため、LSTM と GRU はバイアスとバリエーションの関係にあると考えられ、データ量が多い場合は LSTM の精度が良いと報告されている。

5. 適用流域

本稿にて対象とする道志ダム上流域は、図3の通り、神奈川県相模川水系道志川に位置し、流域面積が 112.5km^2 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(図1

下段参照)は概ね60~90分程度である[2]。

6. おわりに

発表では、図6に示すような RNN, LSTM, GRU の適応性を確認し、3つのネットワークの道志ダム流入量の予測精度の違いを報告する。

参考文献

- [1] 天方匡純, 渡邊優: ダム流入量予測技術の一つの考え方, 第72回年次学術講演会, (2017).
- [2] 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.339, pp.9-15, (2017).

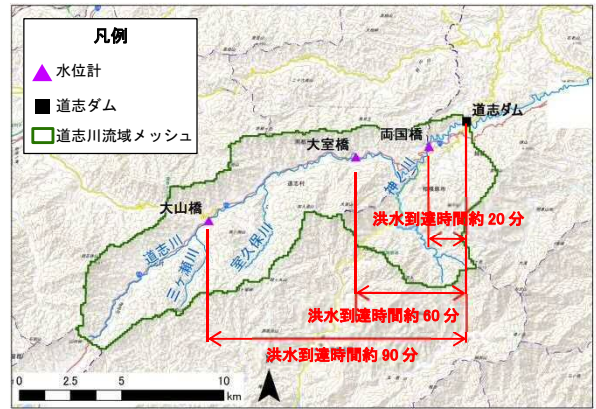


図3 道志ダム流域の概要

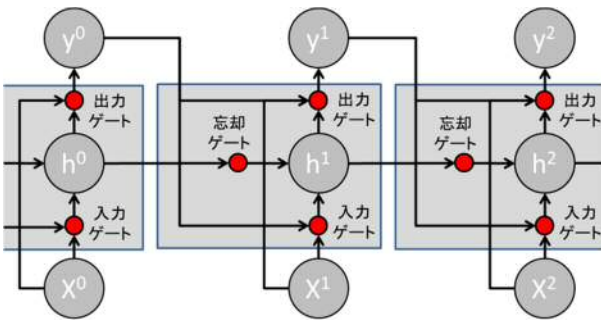


図4 LSTMの特徴

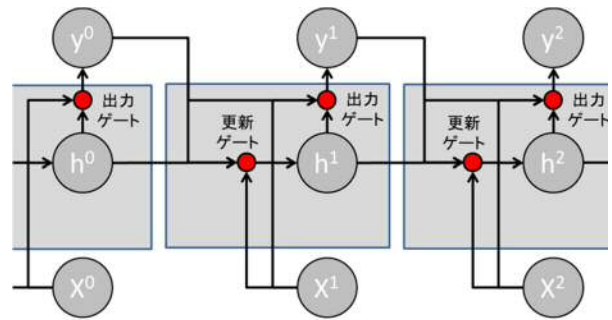


図5 GRUの特徴

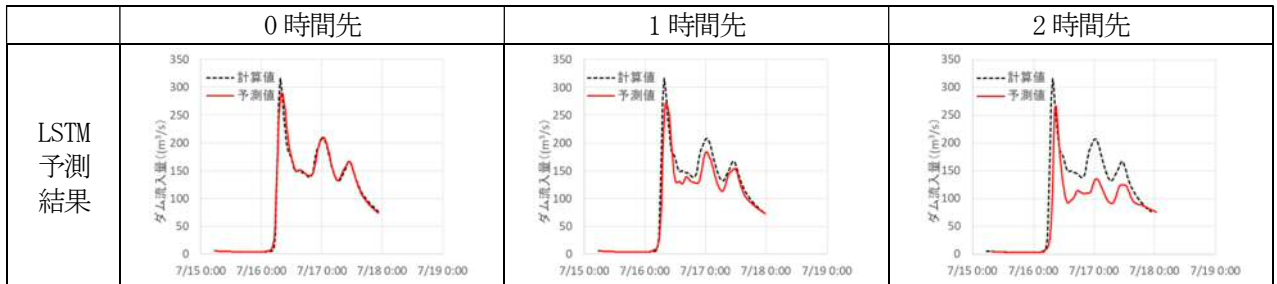


図6 LSTMの計算結果の一例

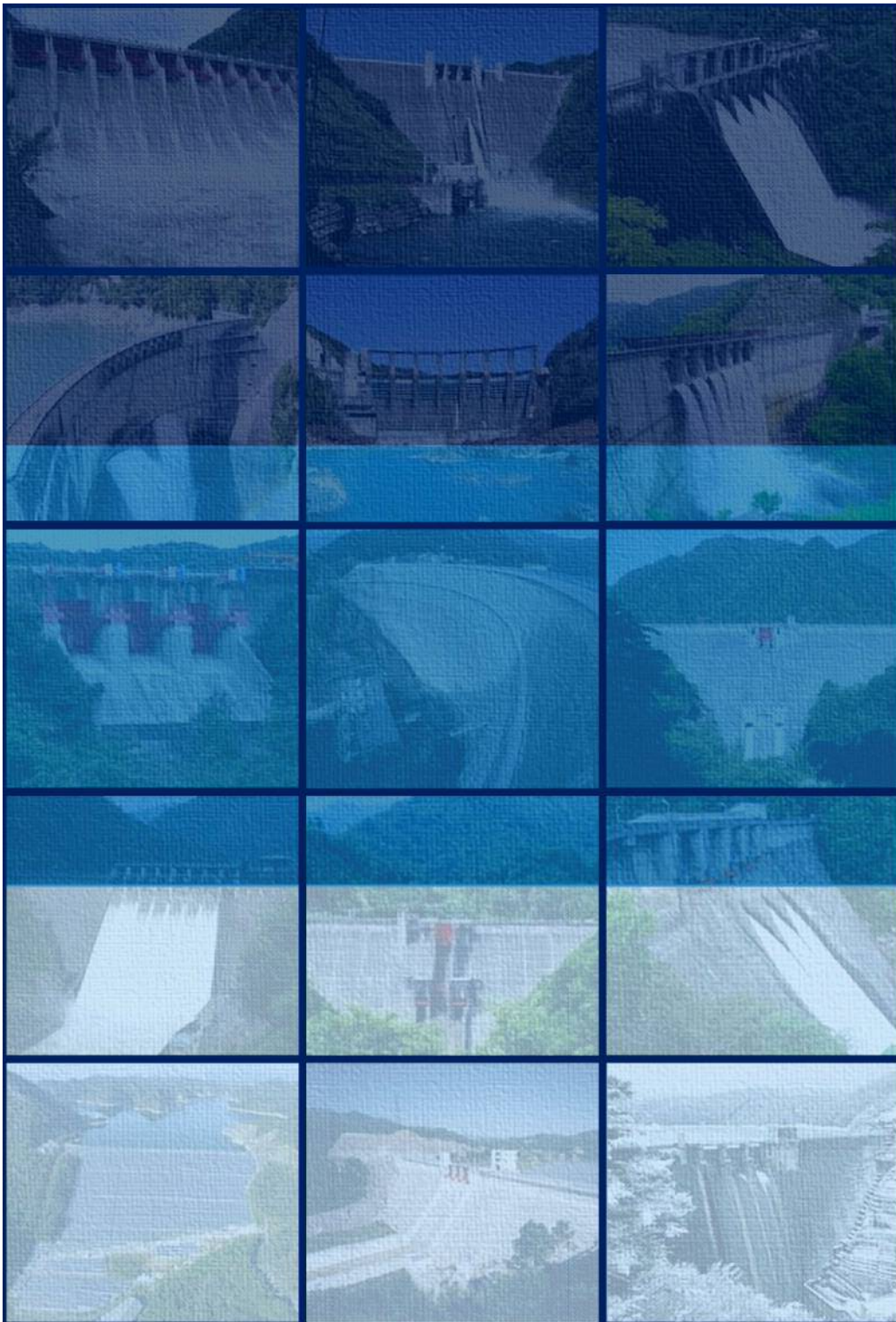
2019年3月15日(金)
日本オペレーションズ・リサーチ学会
2019年春季研究発表会
in 千葉工業大学

ダム流入量予測への RNN系モデルの適用

天方 匡純



yec 八千代エンジニアリング株式会社
<http://www.yachiyo-eng.co.jp/>



昨年7月の西日本豪雨のような記録更新の降雨が頻発するなか、河川氾濫を前提とした水防災意識社会が到来し、防災・避難のための予測情報提供の重要性が高まっている。しかし、**水位予測やダム流入量予測は当たらない。**

当たらない洪水予測

産経ニュース

閉じる

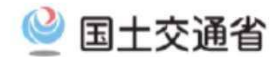
2017.7.2 07:29

河川の増水予測…外しまくりだった 国交省、モデル大幅見直しへ

大雨で増水した際に国土交通省などが特定の河川に対して発表する「指定河川洪水予報」について、過去5年間に行われた水位予測のうち、予測通りに水位が増減したのは約60例中で数例にとどまることが1日、同省関係者への取材で分かった。精度の低さが浮き彫りになった形で、同省は検討会を立ち上げて精度の大幅向上に乗り出している。

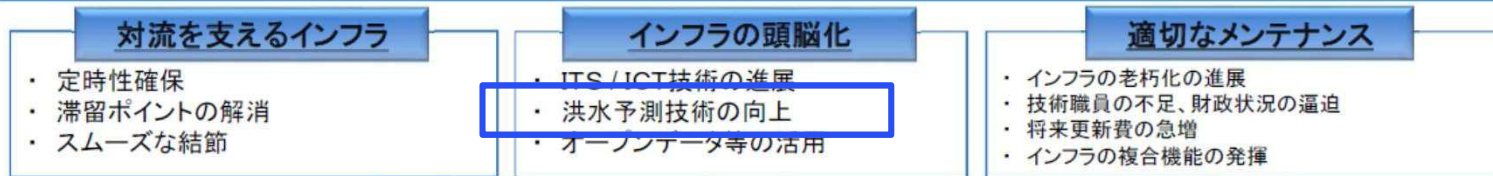
国交省は指定河川の水位予測状況を検証するため、各河川事務所を通じ、直近5年以内に発生した増水時の予測データを約60河川で1ケースずつ抽出し調査。

1時間ごとに出される1～3時間先までの水位予測を実測値と比べたところ、上がるとした予測より早く実測値が上がったり、下がるとした予測が実測値では上がるなど、予測と実測値が一致しないケースが多かった。判定基準はないが、おおむね一致したのは10ケース未満だったとみられる。



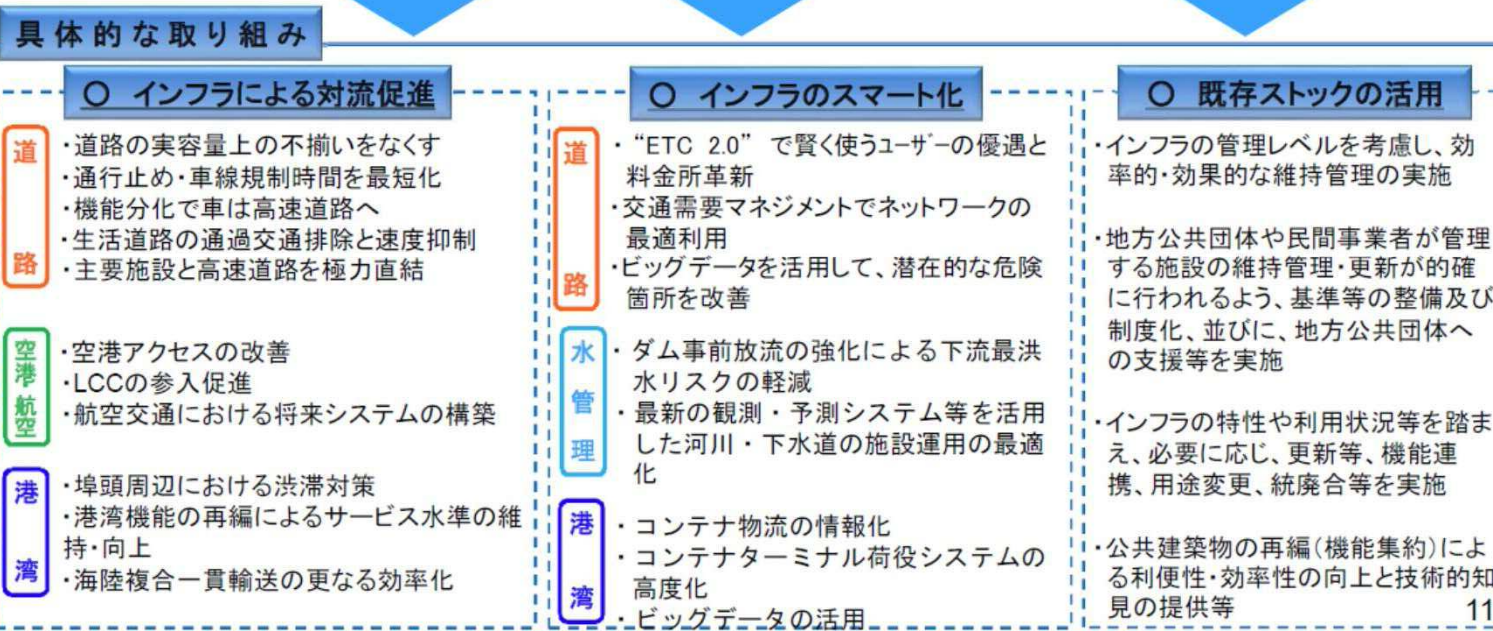
(9) インフラを賢く使う

○ インフラの**頭脳化**、滞留解消による**スムーズな結節**、既存施設の**適切なメンテナンス**により、ヒト、モノ、情報の**対流を促す“次世代インフラ”**を形成



方向性

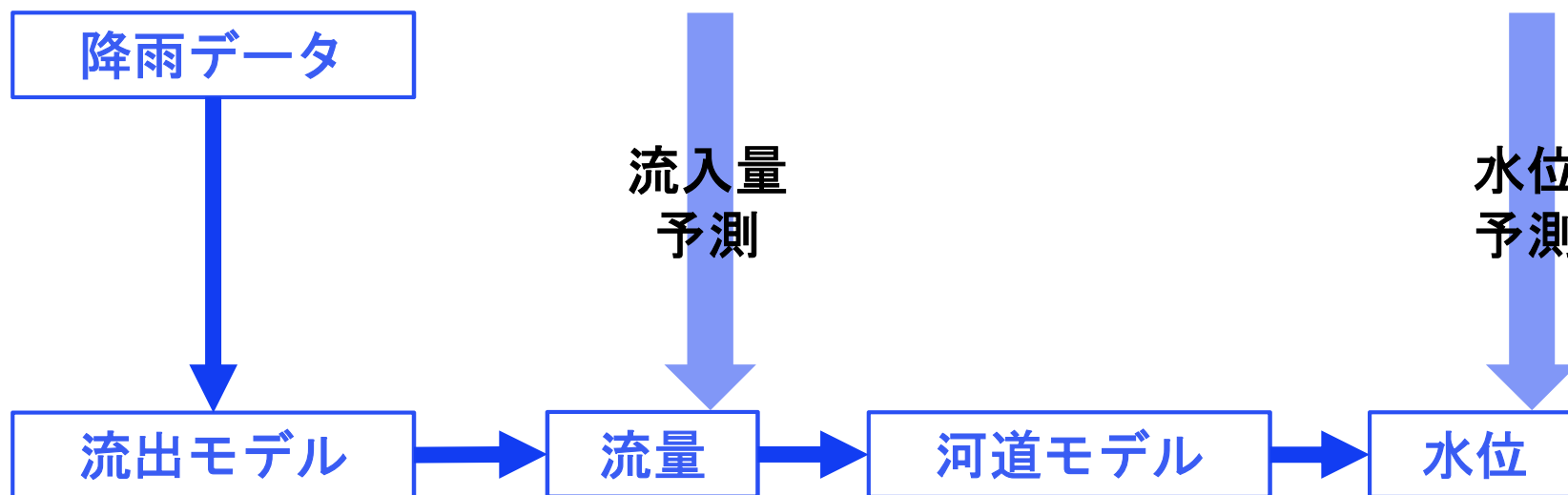
○ 高規格幹線道路、整備新幹線、リニア中央新幹線等の早期整備・活用等により、ネットワークの充実を図るとともに、対流、頭脳化、メンテナンスによりスムーズでスマートなインフラの形成



出典) 国土のグランドデザイン2050

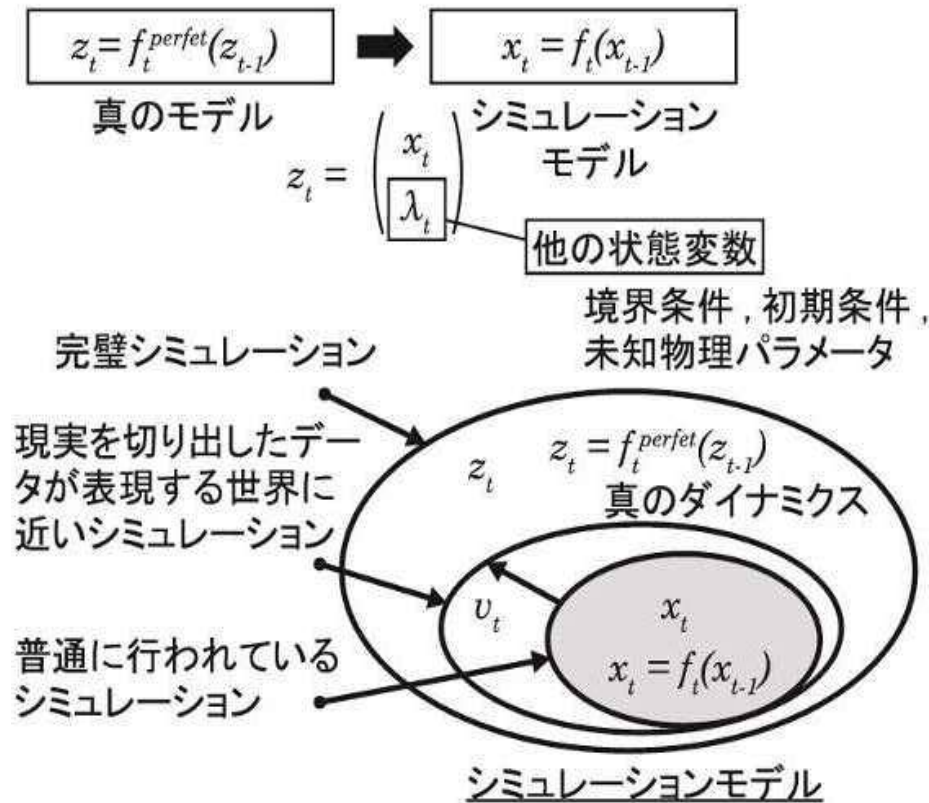
これまでのシミュレーションモデルを中心とした水位予測、ダム流入量予測は、データ、モデルに不確定要素が多く、予測精度確保が難しい。

これまでの水位予測・ダム流入量予測



物理モデルを実務に適用する場合、様々な簡便化が行われる。この結果、自然現象とシミュレーションモデルの再現現象との間には大きな乖離が発生する。この乖離を埋める統計的フィルタリングにも限界がある。

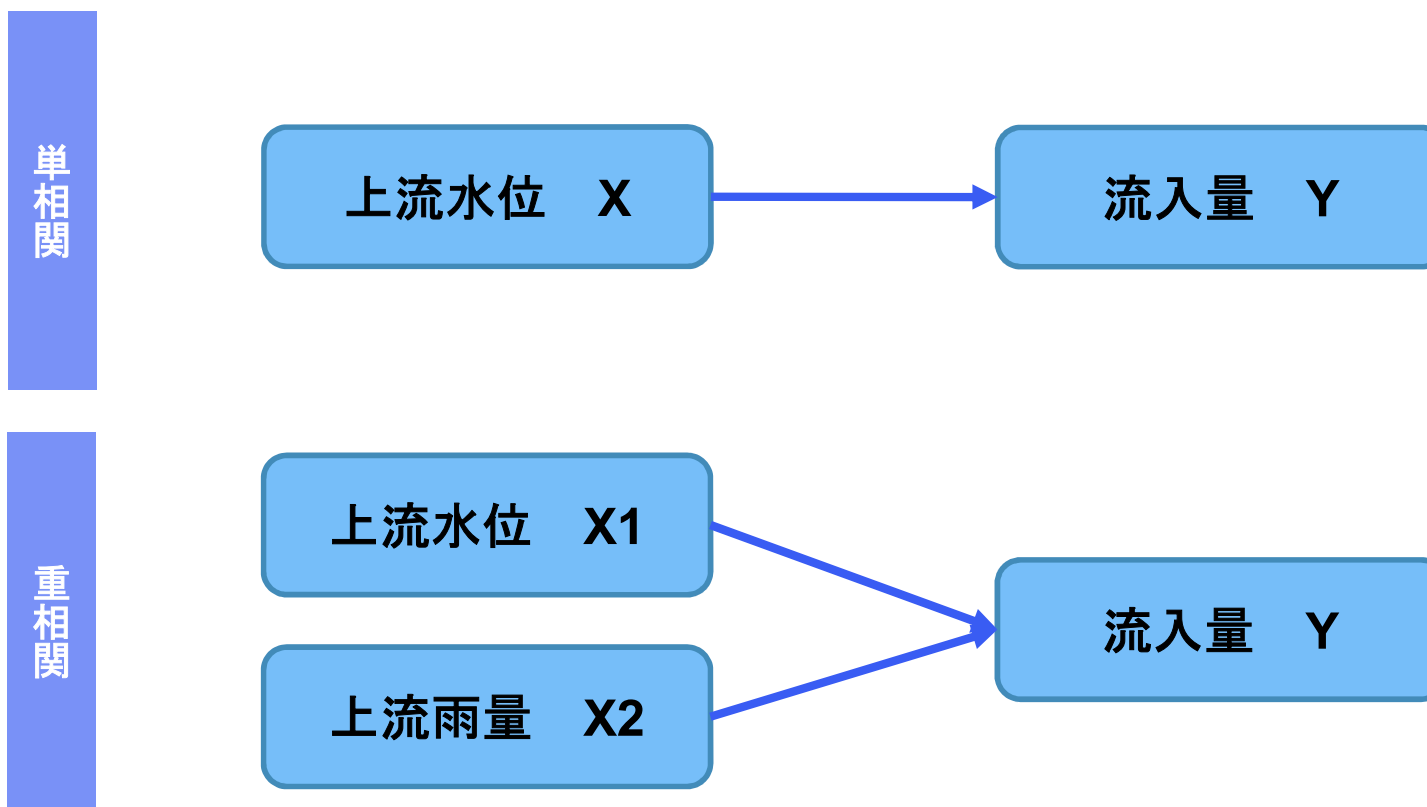
シミュレーションモデルの限界



出典) データ同化入門, 朝倉書店, 2011

上流と下流のデータの相関性を活用した予測は、**洪水到達時間分の先読みツール**として河川管理の現場で利用されている。

上流と下流の水位・流入量相関を用いた予測手法

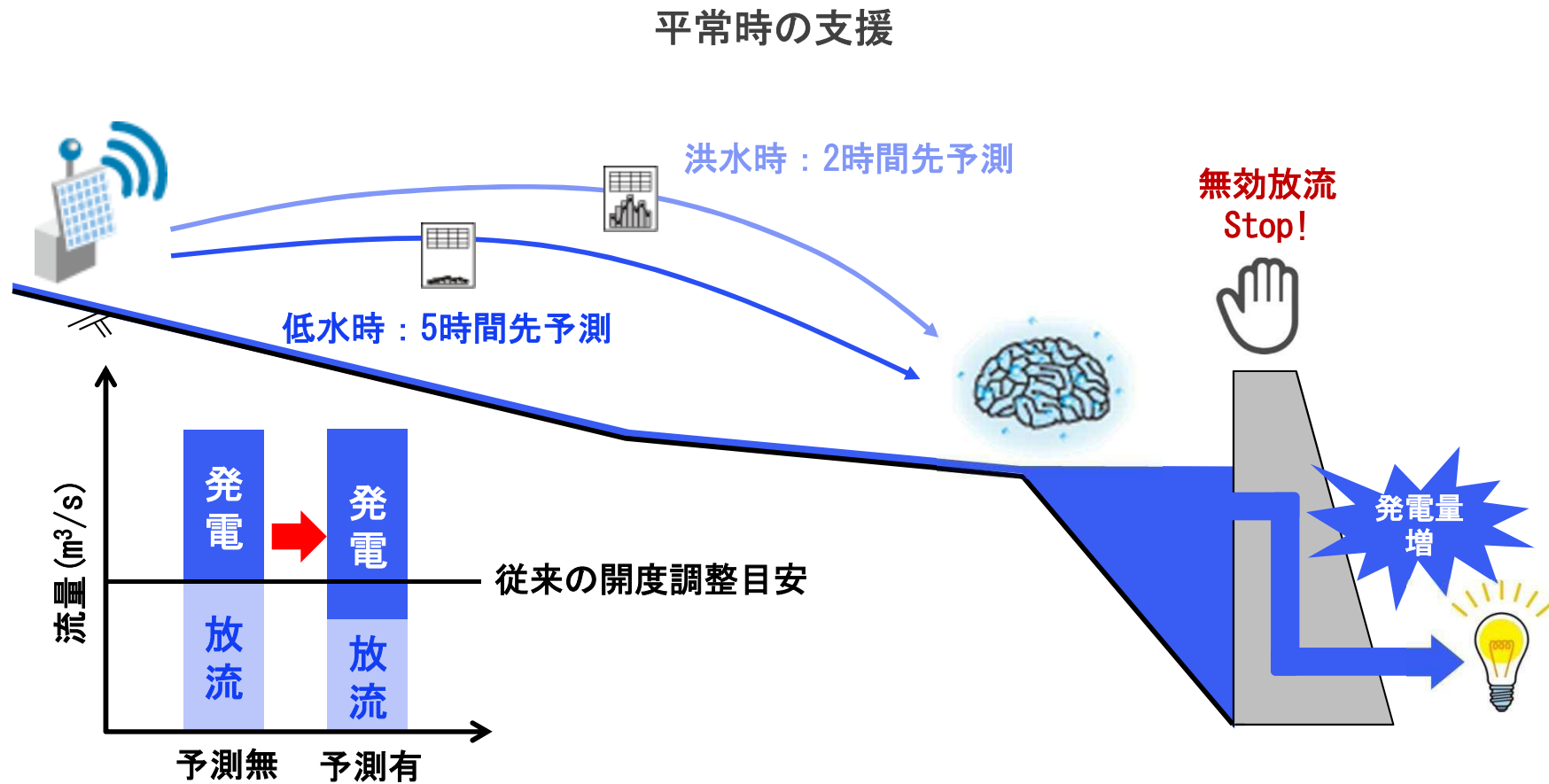


上流と下流のデータの相関性を活用した予測は、洪水到達時間をリードタイムとして、上流側のパルスから下流側のパルスを予測することになる。

上流と下流の水位・流入量の関係

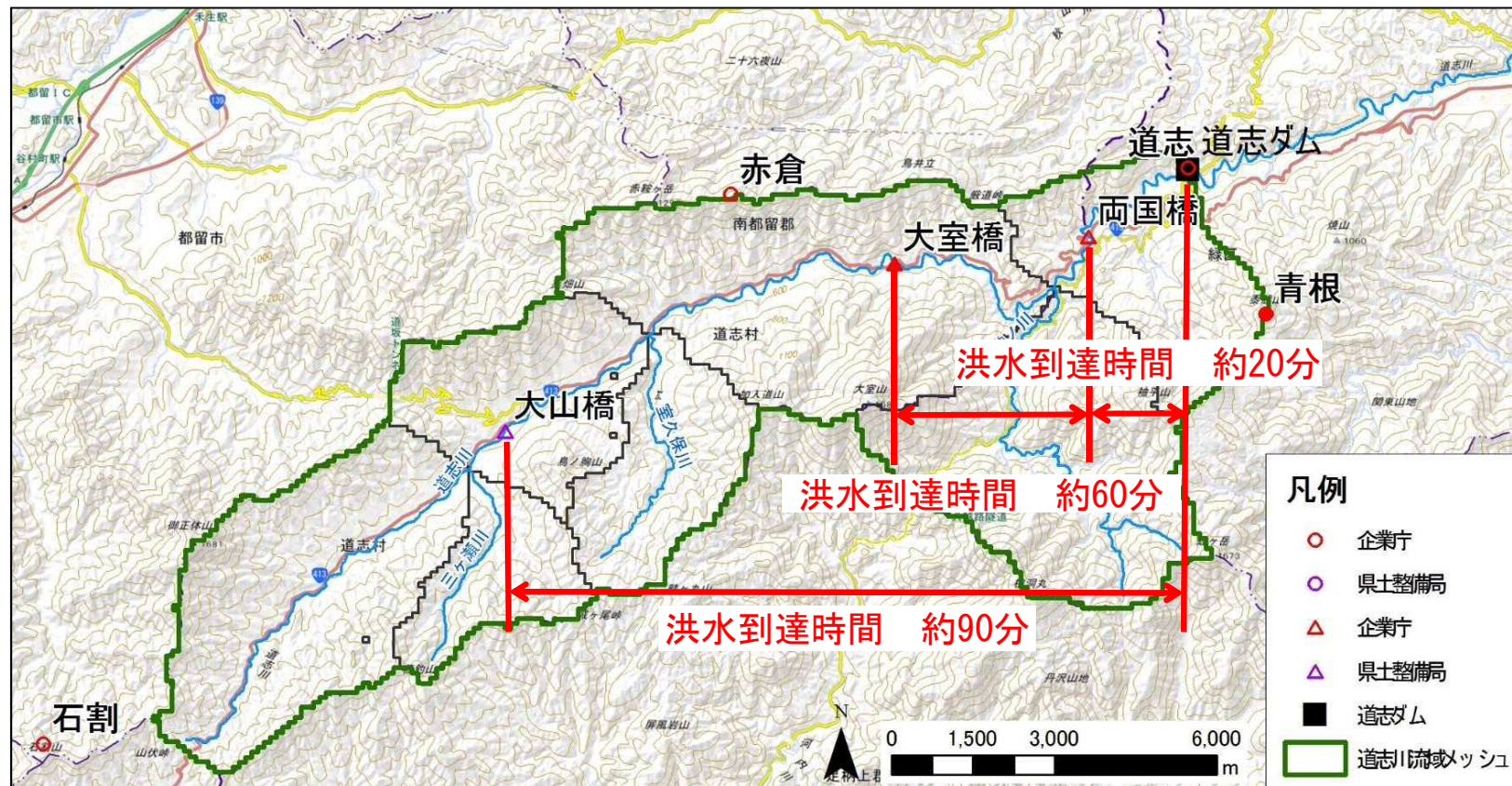


ダム流入量予測精度向上により、**低水時の積極的運用が可能となり、発電量をアップ**することも可能である。



対象流域は、神奈川県相模川水系にある道志ダム流域であり、流域面積100 km²で洪水到達時間が60～90分程度である。道志ダム上流域には、3ヶ所の水位観測所（大山橋、大室橋、両国橋）が設置されている。

洪水到達時間は60～90分



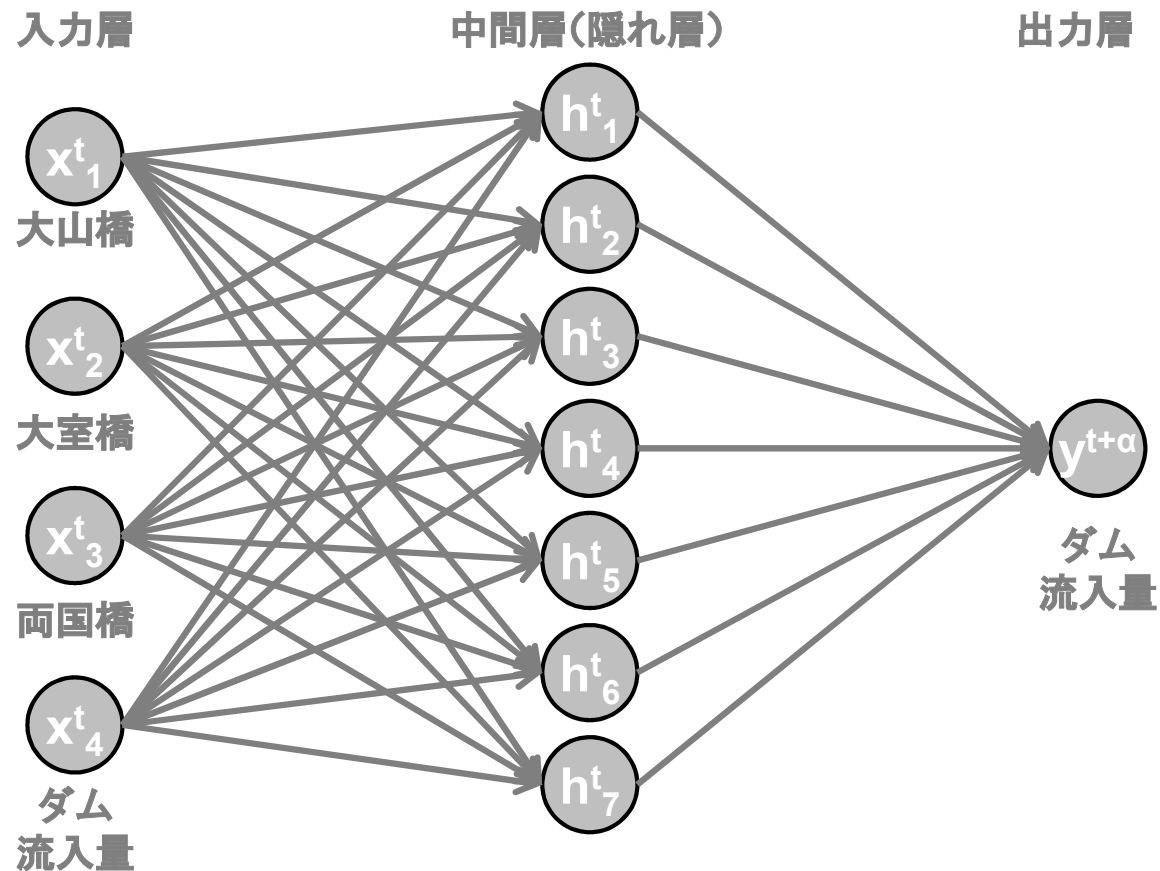
入力データは30分毎の上流側の流量データとし、1時間先、2時間先、3時間先等のダム流入量を予測する。

入力データセット

洪水 No	時刻	入力データ				出力データ			
		大山橋 流量	大室橋 流量	両国橋 流量	ダム 流入量	ダム 流入量			
		(m ³ /s)	(m ³ /s)	(m ³ /s)	(m ³ /s)	0時間 (m ³ /s)	1時間先 (m ³ /s)	・ ・	6時間先 (m ³ /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	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・	・ ・ ・

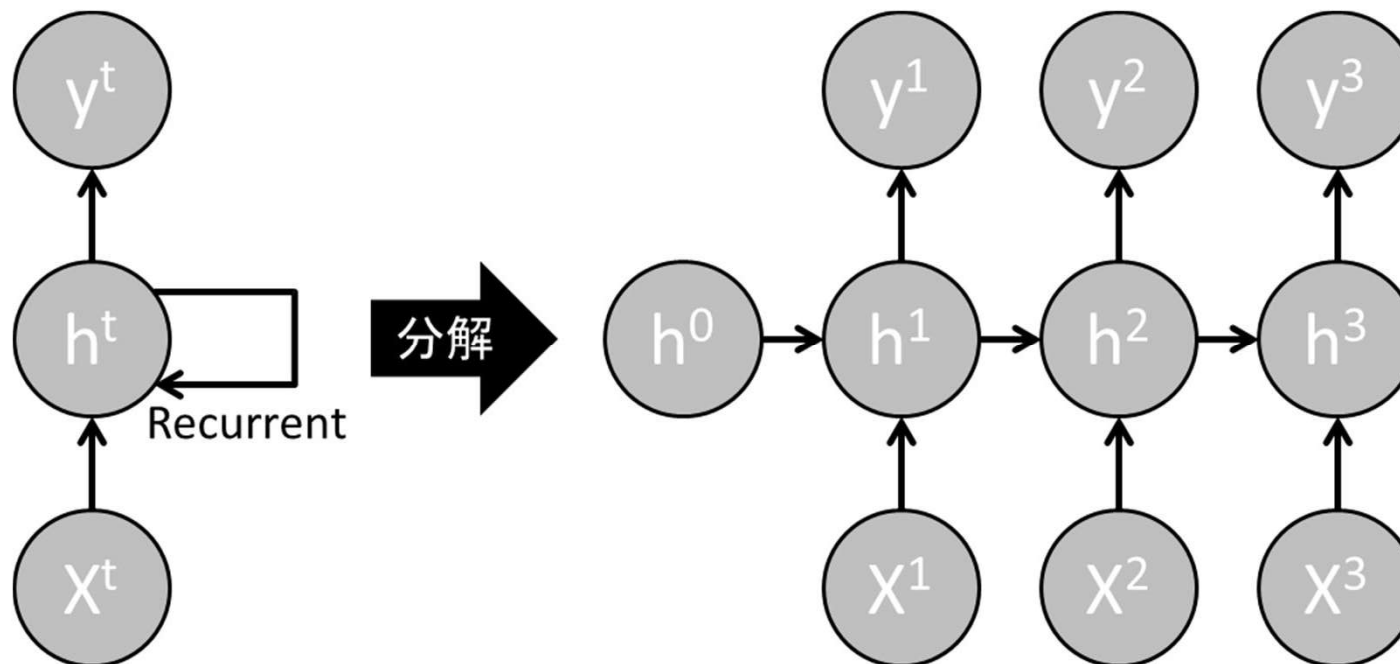
単相関や重相関よりもデータ間の非線形関係を表現できる手法としてニューラルネットワークがある。中間層を重ね、画像のような情報量の多い複雑な特徴量さえも抽出することができるが、再帰性を有していない。

NNモデルの概要



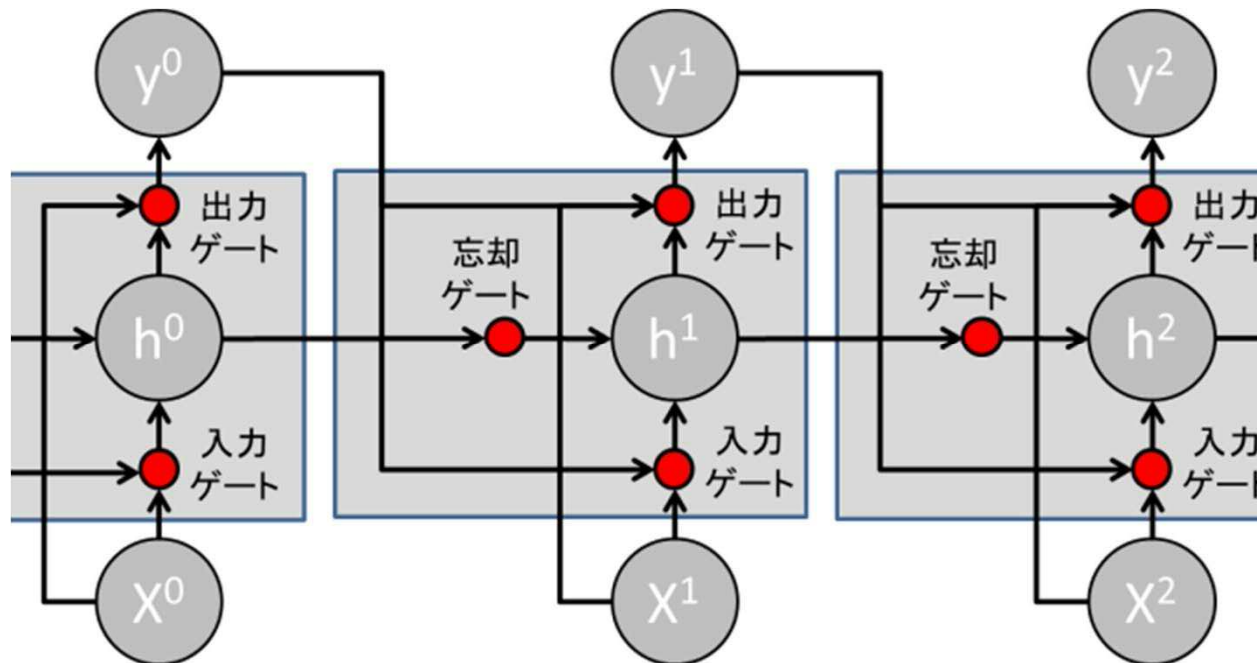
ニューラルネットワークとほぼ同様の構造で再帰性を有するネットワークをRNN（Recurrent Neural Network）と呼ぶ。具体的には、時刻 t の隠れ層に時刻 $t - 1$ 以前の隠れ層を関係付け、再帰性を持たせている。

RNNモデルの概要



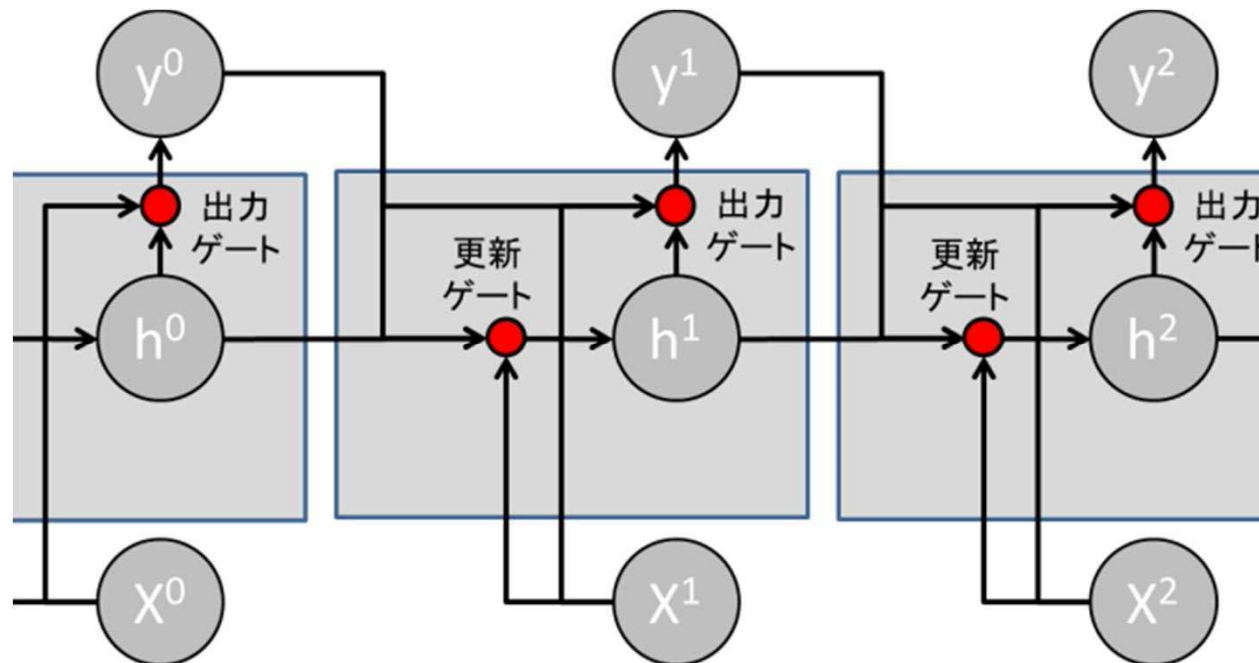
RNNの重み決定に当たっては、バックプロパゲーションの勾配消失問題により、現時刻から数ステップ前の情報としか関連付けられないことが確認されている。過去情報依存性を高めるため、LSTMブロックが開発された。

LSTMモデルの概要

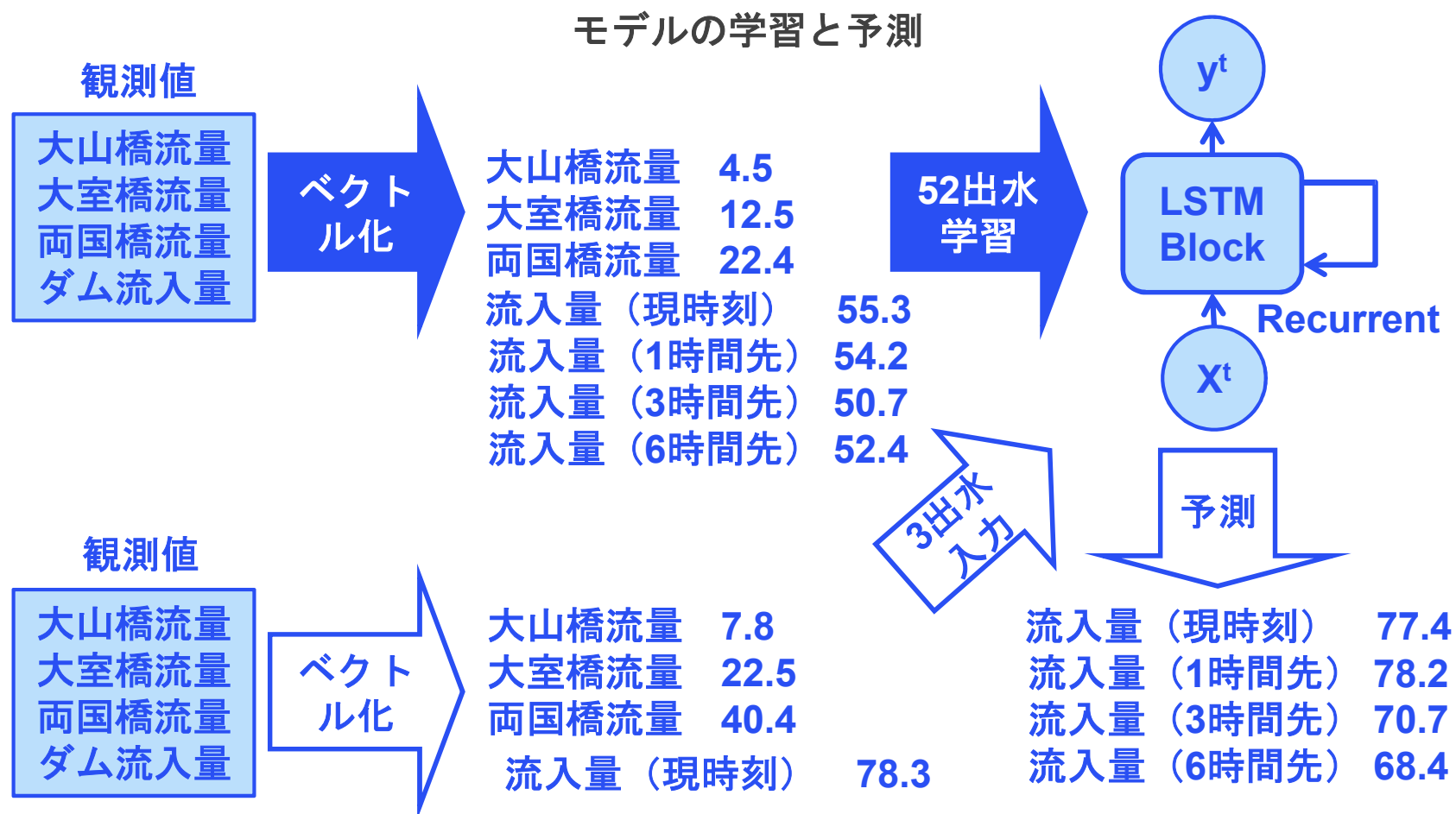


LSTMの計算過程を簡略化した手法がGRUである。パラメータが少なくなり、過学習性が弱まっていると言われている。

GRUモデルの概要

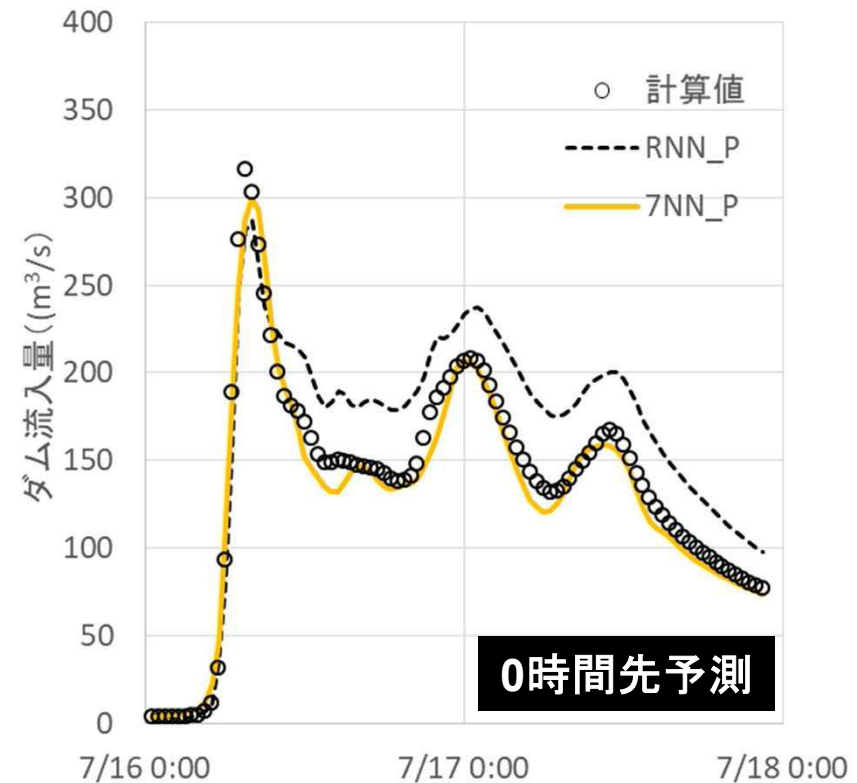
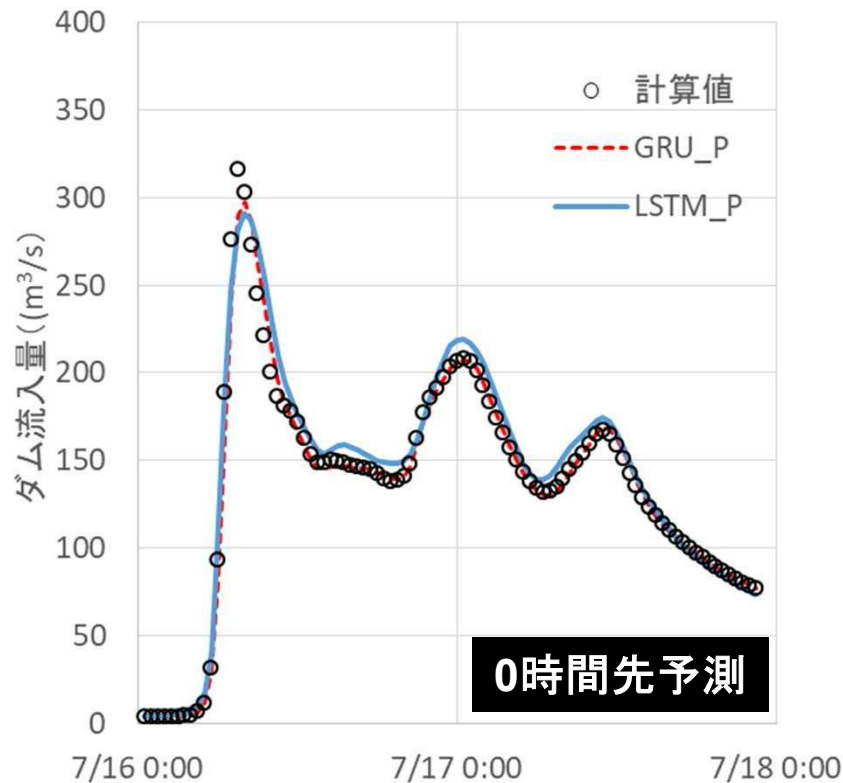


モデル構築に当たっては52出水を学習してパラメータを決定したうえで、平成27年の3出水でモデル精度を検証する。



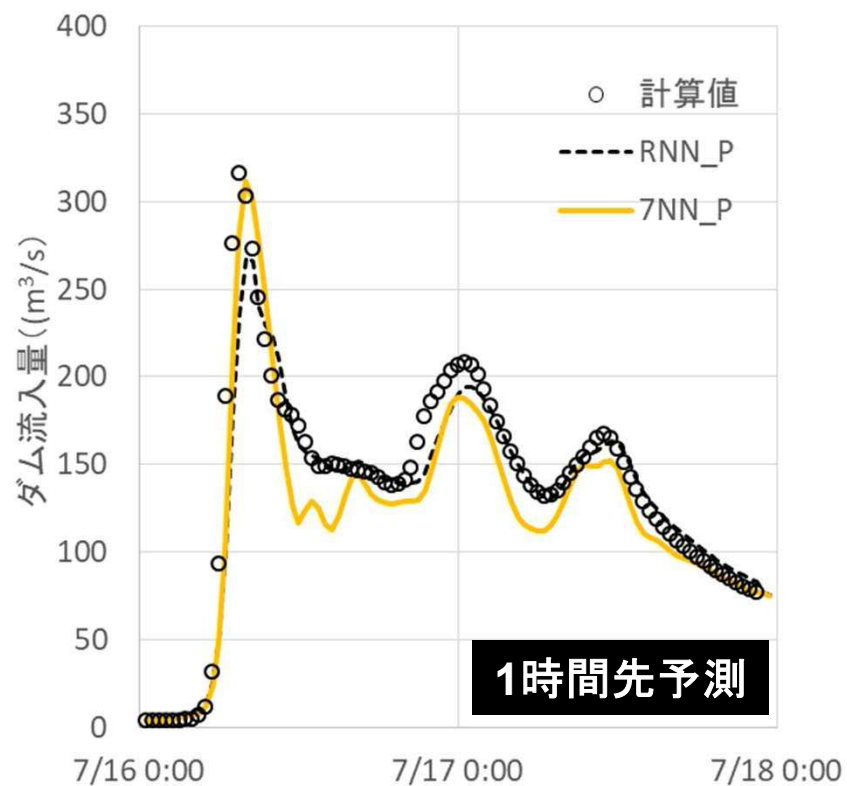
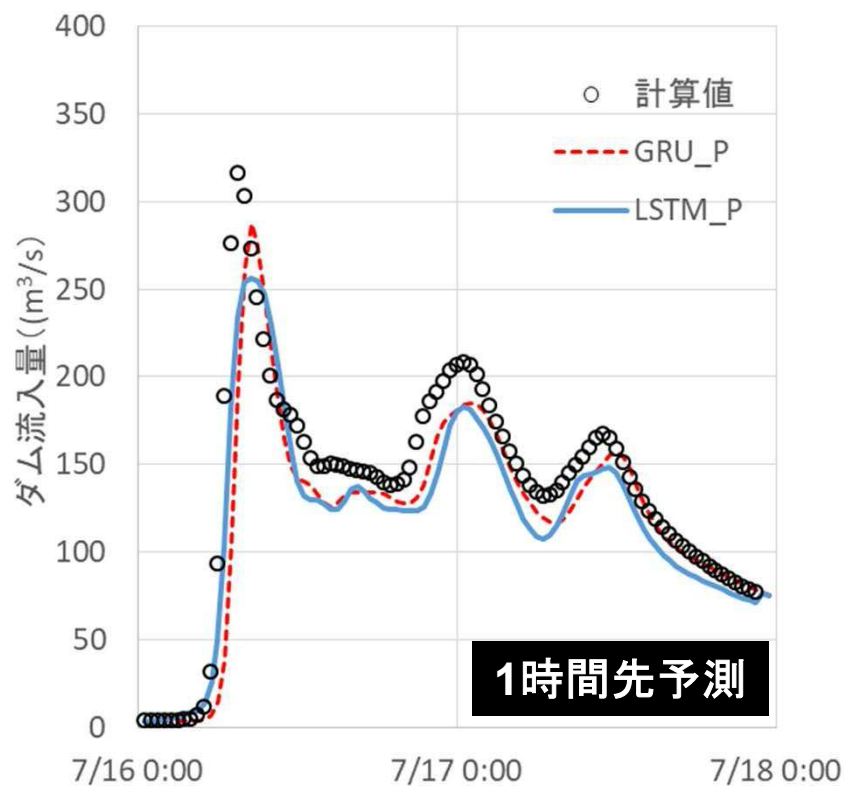
GRU、LSTM、RNN、7NNによる予測結果を記す。モデルパラメータ最適化に関する条件は全て同じである。また、中間層も全て7ニューロン、7ブロックとしている。RNNモデルが大きめの予測である。

人工知能による流入量予測結果



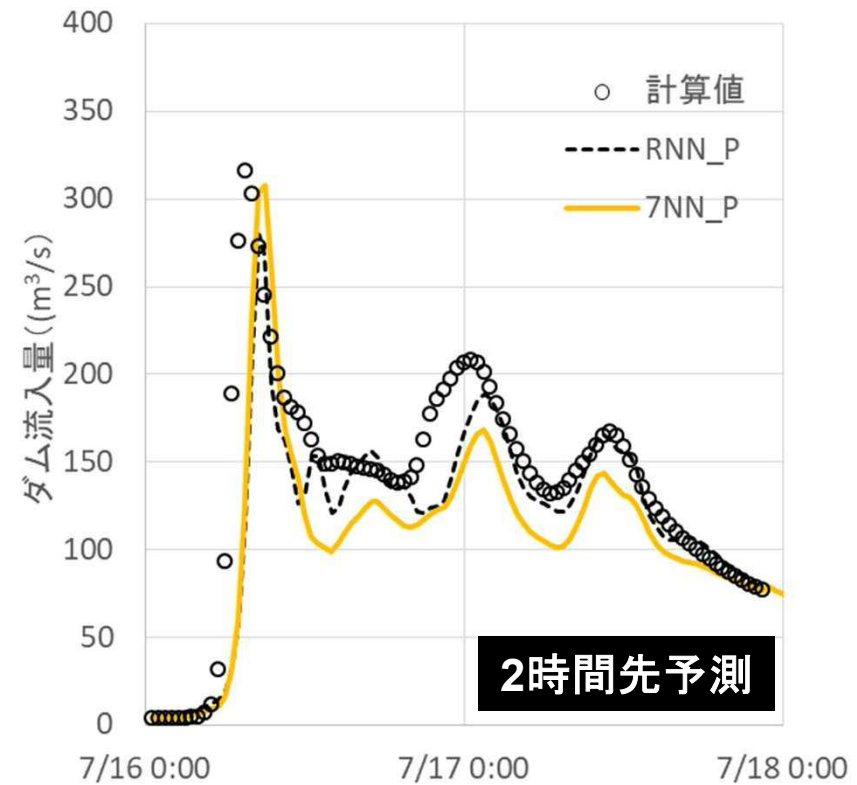
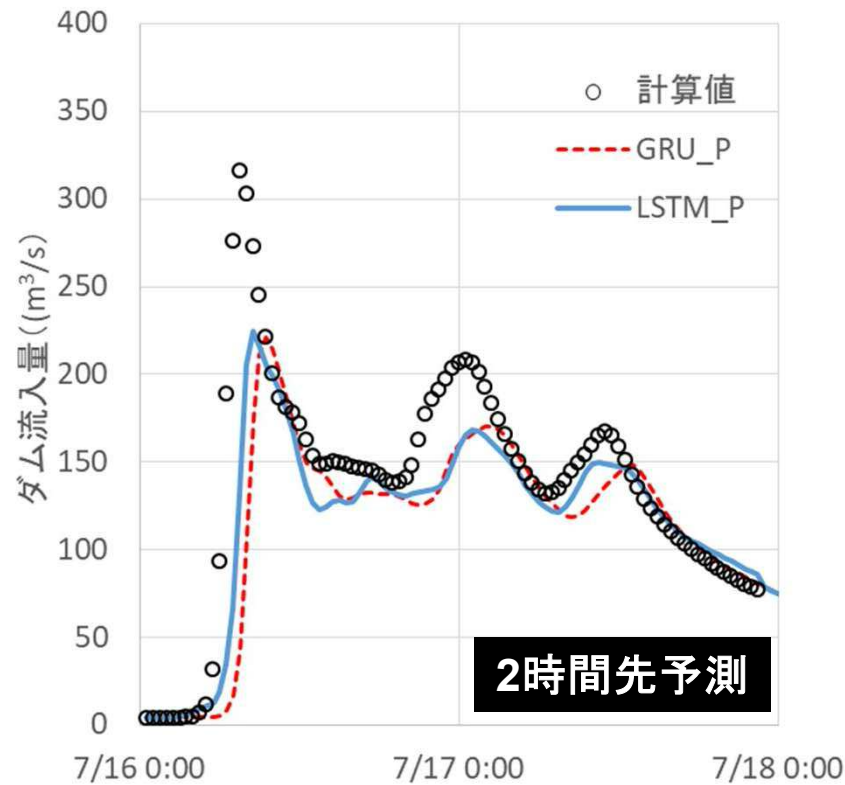
1時間先の予測では、いずれのモデルも同様の予測精度を有している。ピーク流量予測精度が高いのは、GRUと7NNである。

人工知能による流入量予測結果



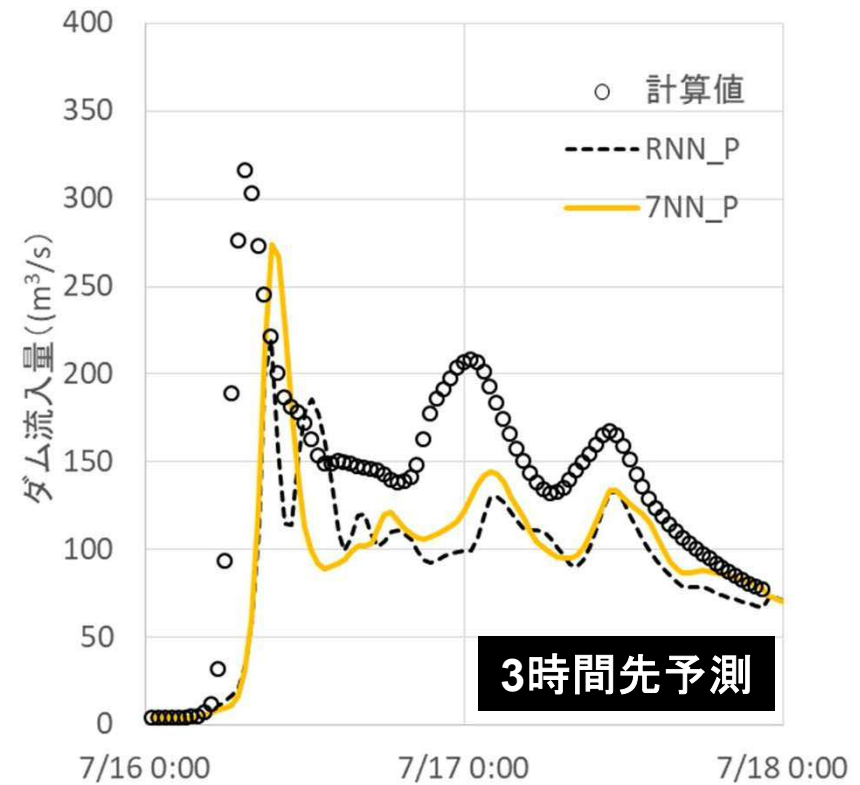
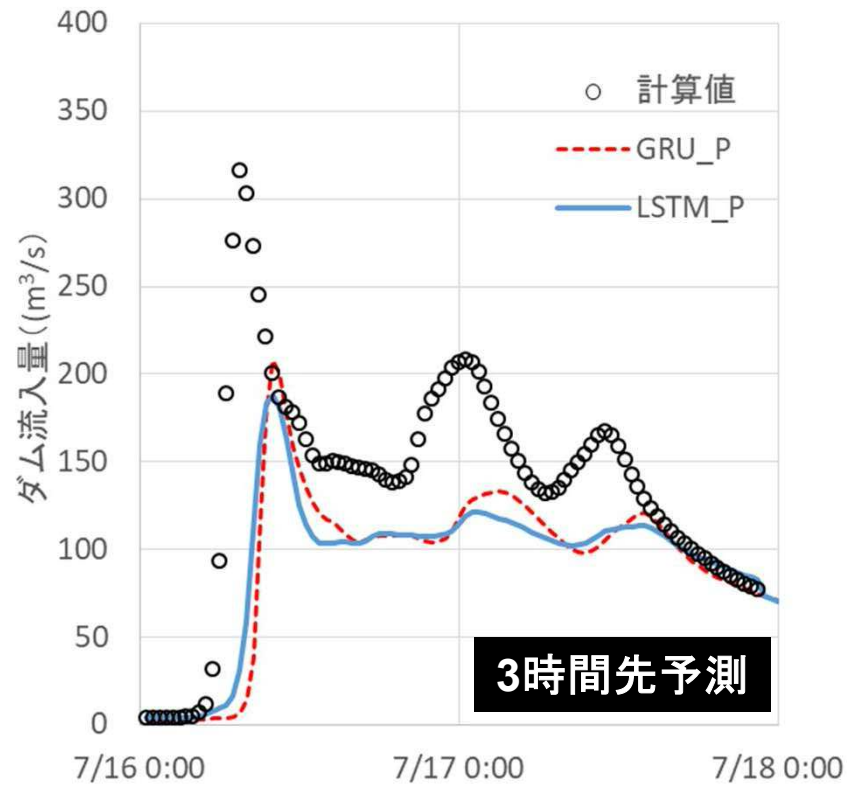
2時間先の予測では、いずれのモデルも再現性が低下し、全体的に計算値（観測値）よりも小さめに予測する傾向がある。

人工知能による流入量予測結果



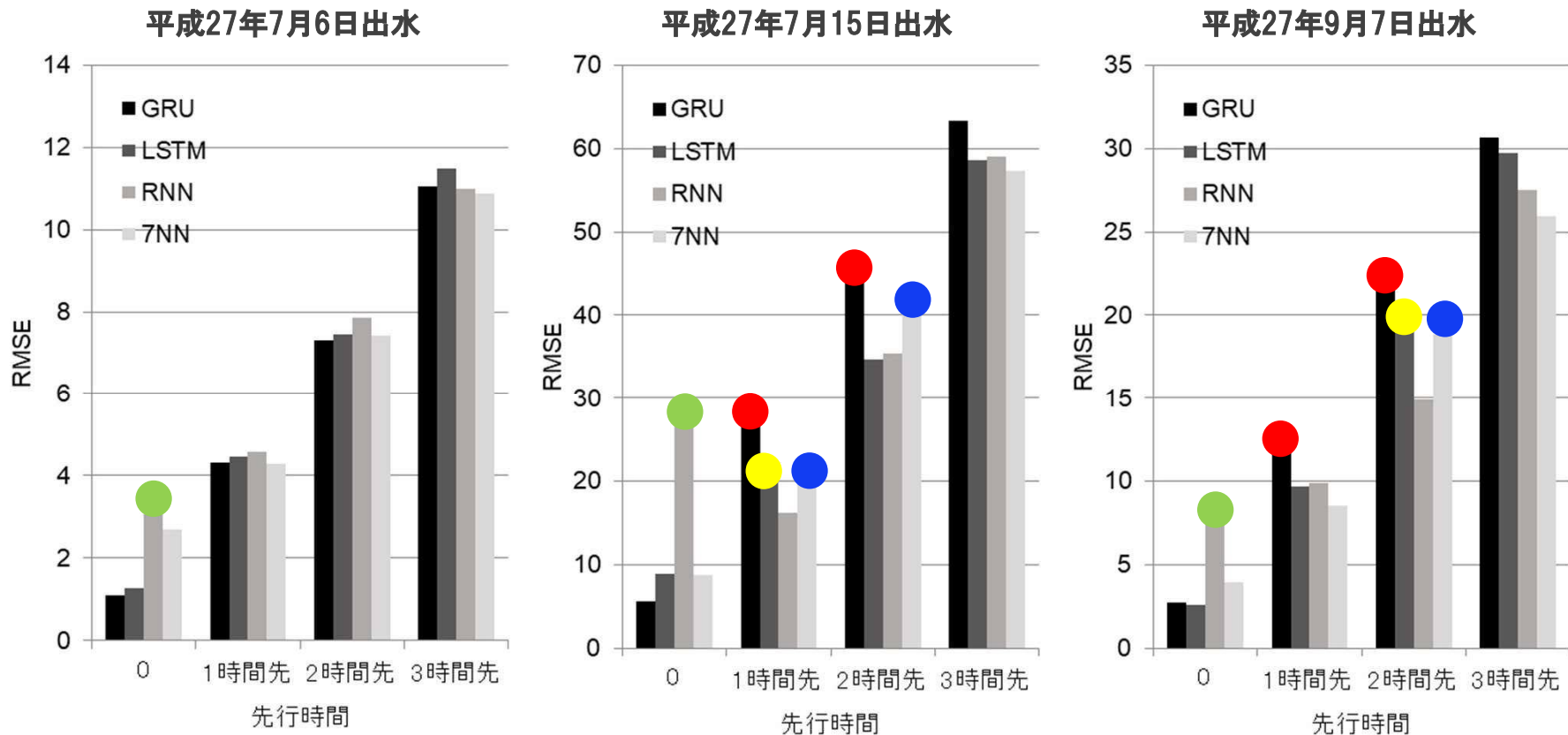
3時間先の予測では、再現性が大きく低下し、流量の大小起伏応答性が弱くなる（平滑化）傾向にある。

人工知能による流入量予測結果

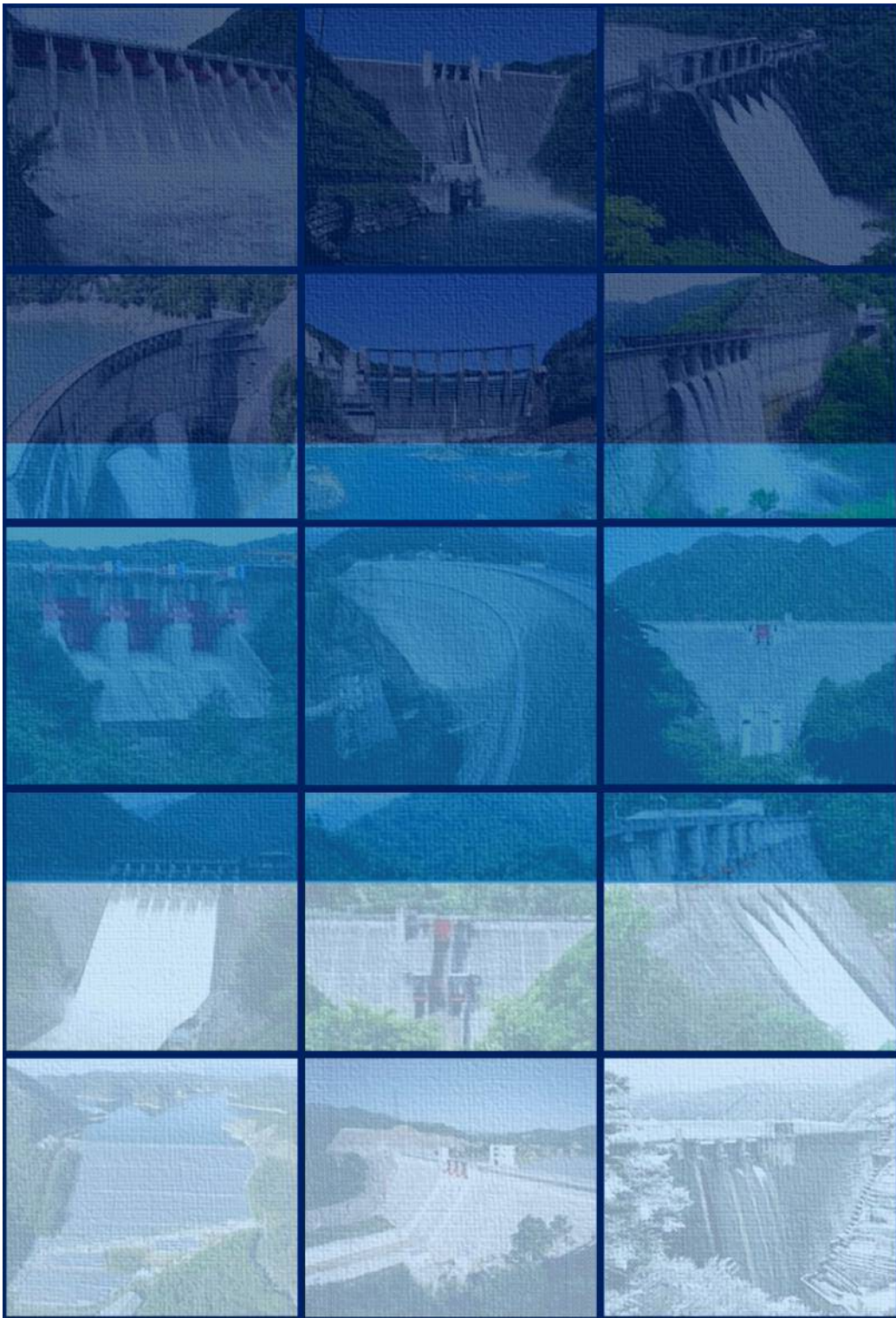


平成27年の3出水に対して実測と計算結果をRMSE（Root Mean Square Error）で比較した。出水規模によりRMSE値は異なるが、洪水到達時間を踏まえた2時間先までのモデル精度は、GRU < RNN < LSTM と評価できる。

異なる出水での精度（RMSE）比較



- RNN系深層学習モデル間で予測精度に大きな違いは無かった。100km²程度の流域では、水文データ間の非線形性が弱く、また、流出時間が早く、現象そのものの過去依存性（再帰性）が小さいため、従来のニューラルネットワークでも十分に再現可能であったことに起因すると考えられる。
- 3出水の2時間先までの予測精度について、特に悪い結果を示さなかったのはLSTMである。RNNと7NNは精度が良い出水もあるが、精度が悪い出水もあり、安定性に欠けた。
- 出水時だけでなく平常時での活用を念頭に置くと教師データも多くなり、精度も適用性も向上する可能性がある。
- 大規模流域であれば上流と下流で水文データ間の非線形性も高まり、深層学習モデルの有効性を確認できる場となり得る。
- 水位や流量等の水文データセットは、現在深層学習モデルが最も得意とする画像データと比較すると極めて小さなデータとなり、過学習を誘発する。このため、メッシュ雨量データ等を加えて、バリエーション性を低減していく手立てが必要である。



ありがとうございました。