

中部地方整備局管内におけるAI技術を活用した ダム流入量予測の精度向上

浅野 雅博¹

¹中部地方整備局 河川部 河川管理課 (〒460-8514 愛知県名古屋市中区三の丸2-5-1)

中部地整管内のダムを上流ダムの放流操作の影響度等に応じて3パターンに分類し、パターン毎に当該ダムの流入量予測を目的とした深層学習モデルを構築した。予測波形の再現性や定量的な誤差指標により適用性を評価することで、AIを活用した深層学習によるダム流入量予測技術を他ダムへ展開する上での有用な知見を得た。面的な雨域の把握が可能なレーダ雨量を深層学習モデルの入力データとし、近年出水を学習対象出水としているが、低頻度・未経験洪水に対する予測精度確保が課題となるため、レーダ雨量が存在しない期間の大規模出水を地上雨量データを学習対象出水に追加することで、これらの洪水に対する予測精度を確保した取組について報告する。

キーワード AI技術の活用, ダム流入量予測, 入力データの選択, 深層学習

1. はじめに

近年、気候変動に伴う降雨強度・総雨量増大への対応、ハイブリッドダム及びカーボンニュートラルへの貢献等のため、流域内のリスクを最大限軽減でき、また未利用の水力エネルギーを有効活用するダムの放流操作が求められている。このようなダムの放流操作を実現するためには、ダム流入量の予測精度を確保することが重要であり、AIによる深層学習を適用したダム流入量予測技術に関する研究が数多く行われているが、学習データが限られる低頻度・未経験洪水に対して予測精度を確保できないなど、適用性に課題がある。

中部地方の治水機能を有する13ダムを対象に、上流に位置するダムの放流操作が当該ダム流入量に及ぼす影響度を分析し、上流ダムの放流操作の影響度等に応じて3パターンに分類した。各パターンから1ダムを選定し、特性に応じた深層学習モデルを構築することで、得られる予測精度やダム流入量予測への適用性を評価した。こうして深層学習モデルの構築過程を一般化することで、特性に応じた他ダムへの適用に寄与できると期待される。

ここで、面的な雨域の把握が可能なレーダ雨量を深層学習モデルの入力データとすることで、ダム流入量予測の精度確保が期待できる。そのため、低頻度・未経験洪水への対応を考慮して、雨量データの種別が混在する条件下で学習させたモデルを基本とするが、それによる予測精度への影響を分析することで、レーダ雨量全国合成開始以前に発生した大規模出水を学習データに追加す

ることの是非を確認した。

2. 中部地方のダムの分類と対象流域の選定

ダム流入量予測に対し、深層学習を適用する際、予測対象ダム（以下「当該ダム」という）上流域におけるダムの有無やそれらの放流操作の当該ダムへの影響度に応じて、最適なモデル構成が異なると考えられる。そこで、中部地方の治水機能を有する13ダムに対し、ダムの管理開始年以降に発生したピーク流入量が多い10出水程度を対象として、流下時間を考慮した上流ダムの調節量と当該ダムの流入量から、次式より上流ダムの放流操作が当該ダムの流入量に及ぼす影響量を算出した。

上流ダム放流操作の影響量 (%)

$$= \frac{\text{上流ダムの調節量}}{(\text{予測対象ダムの流入量} + \text{上流ダムの調節量})} \times 100$$

当該ダム上流域におけるダムの有無や放流操作の影響量をもとに、13ダムを分類した(表-1)。上流にダムが無いまたは有っても放流操作の影響量が小さい(各出水5%以下)場合をパターン①、放流操作の影響量が多い(5%超過の出水が存在)上流ダムが一つ有る場合をパターン②-1、複数有る場合をパターン②-2に分類した。

本論では、各パターンから新豊根ダム(パターン①)、横山ダム(パターン②-1)、丸山ダム(パターン②-2)

の3ダム(図-1)を選定し、ダム流入量予測を目的とした深層学習モデルを構築した。また、予測波形や誤差指標により予測精度を確認することで、モデル構築方法の妥当性やダム流入量予測への適用性を評価した。

分類パターン	水系名	ダム名
パターン①	上流ダムの放流操作が影響しない	木曽川 岩屋ダム、味噌川ダム、徳山ダム
		櫛田川 蓮ダム
		庄内川 小里川ダム
		天竜川 美和ダム、小渋ダム、 新豊根ダム
パターン②-1	上流ダムの放流操作が影響する	木曽川 横山ダム 、阿木川ダム
パターン②-2	複数の上流ダムの放流操作が影響する	矢作川 矢作ダム
		木曽川 丸山ダム
		大井川 長島ダム

赤字：AIモデルを構築するダム

表-1 上流ダムの放流操作の影響度を踏まえた13ダムの分類

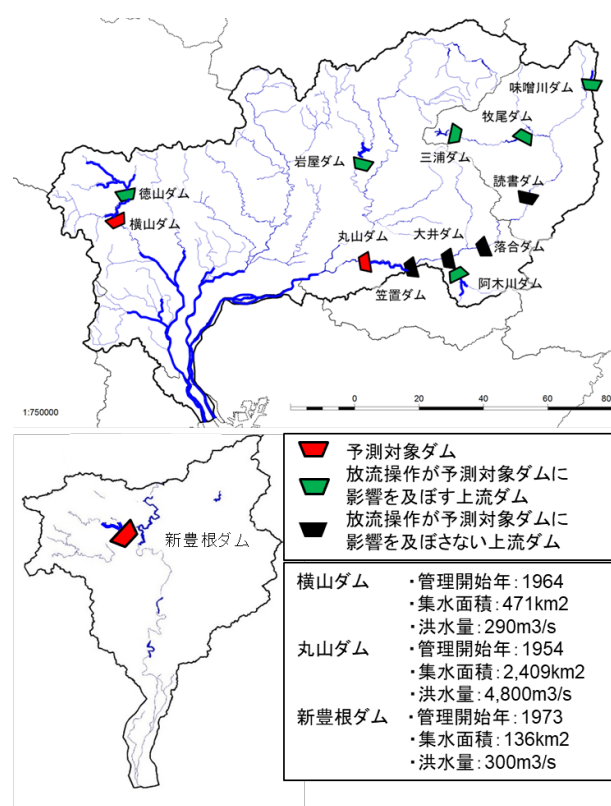


図-1 対象流域の概要

3. 深層学習によるダム流入量予測モデルの構築

選定した3ダムを対象に、事前放流の実施判断等にも資することを目的とした72時間先までのダム流入量予測が可能な深層学習モデルを構築し、予測波形や実績値との誤差を確認することで予測精度を評価した。学習に使用する雨量データは、面的な雨域の把握が可能なCバンド同時刻合成レーダ雨量(以下「レーダ雨量」という)とした。オンライン上で運用される洪水予測システムなどに深層学習モデルを導入する際、気象庁が配信する予測雨量(MSMやGSMなど)を入力データとするこ

とが一般的である。これらはメッシュ単位のデータであるため、データ特性がレーダ雨量と近く、実績雨量と予測雨量の時系列的な連続性の観点からも、学習には地上雨量ではなくレーダ雨量の採用が適切と考えられる。

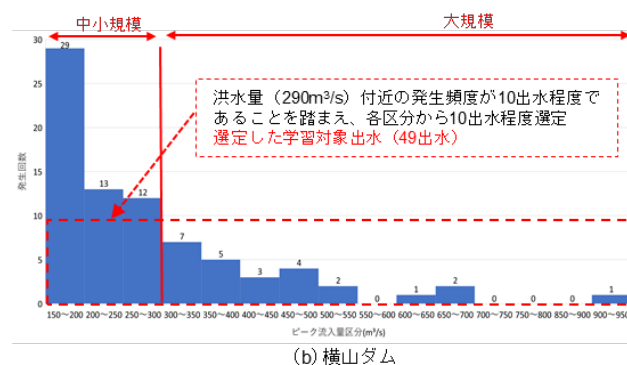
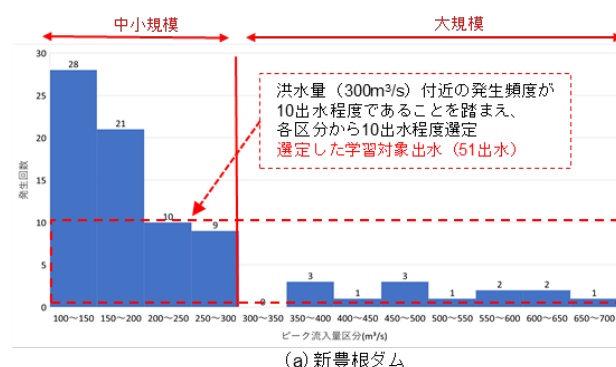
(1) 学習対象出水の選定

レーダ雨量の全国合成が開始された2003年7月以降の期間を対象に深層学習モデルの学習対象出水を選定した。ダム管理上、予測精度の確保が必要な出水規模は以下のいずれかを超過する規模と考えられる。

- ①ゲート放流操作の実施が必要な規模
- ②洪水警戒体制への移行が必要な規模

対象期間において、上記①、②を超過する規模の出水を選定した結果、新豊根ダム：81事例、横山ダム：79事例、丸山ダム：388事例であった。

ピーク流入量が洪水量を上回る出水を大規模出水、洪水量以下の出水を中小規模出水と定義し、発生回数を確認したところ、生起頻度に偏りが見られた。深層学習モデルは、学習データの分布に偏りがあると、データ数が多い領域に最適化されることが懸念される。そのため、生起頻度の高い中小規模出水に対し、降雨波形などの特性が類似する事例を省くことで、流量規模毎の出水数のバランスに留意して学習対象出水を選定した(図-2)。



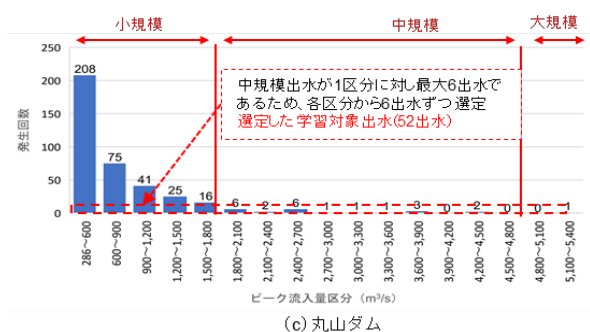


図-2 学習対象出水の選定

(2) AIモデルの学習

AIモデルでは、AIモデルの学習で更新される重みとは別に、学習過程そのものを規定するパラメータ（ハイパーパラメータ）が存在する。これらは、重みと異なり、学習過程で最適化されない。

そこで、複数のハイパーパラメータの組み合わせを事前に設定し、感度分析により最適と考えられるハイパーパラメータの組み合わせを検討した。

ハイパーパラメータの最適化を実施する際に、AIモデルの精度評価の対象とする出水を検証対象出水という。検証対象出水は、学習対象出水から、出水規模や降雨生起要因、降雨波形、降雨分布が網羅されるものを選定した。

その後それぞれのダムにおける検証対象出水に対し、設定した最適化手順に従って、新豊根ダムAIモデルのハイパーパラメータを最適化した。①活性化関数⇒②ミニバッチ数⇒③中間層の層数・ノード数の順に予測精度が改善していくことがわかる。

全てのハイパーパラメータを最適化した予測結果を整理したところ、新豊根ダムでは、2018.10.1.出水（検証対象出水で第一位）や2003.8.9.出水（検証対象出水で第二位）、横山ダムでは、2014.8.10.出水（検証対象出水で第一位）や2008.9.3.出水（検証対象出水で第二位）、2012.9.18.出水、丸山ダムでは、2021.8.15出水（検証対象出水で第一位）や2004.10.21出水、2011.9.20出水において、ピーク付近の予測精度が低い傾向がある。

このため、生起頻度が低く、学習データ数が不足する大規模出水を重複学習（同じ出水データを繰り返し学習）させることで、大規模出水に対する予測精度の向上を図った。（表-2）

ダム	規模	対象出水数
新豊根ダム	600m³/s以上	4出水
横山ダム	600m³/s以上	4出水
丸山ダム	3,00m³/s以上	13出水

表-2 重複学習の対象とする出水の規模と出水数

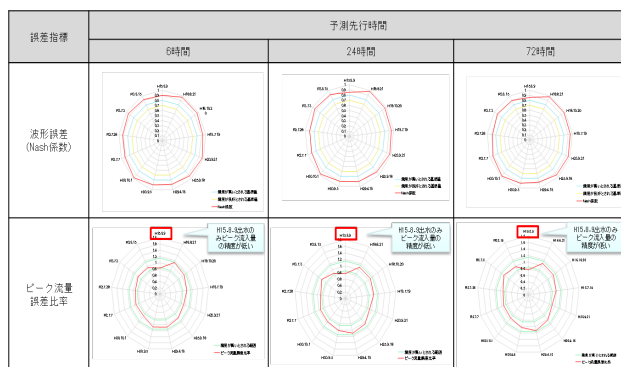
全てのダムにおいて、大規模出水に対する重複学習を実施することでレーダ雨量全国合成開始以降において最大規模の出水などに対して、ピーク付近の予測精度が改善する傾向がみられた。一方、丸山ダムのみ重複学習を実施したことで、予測精度が低下する出水が確認された。

これは重複学習の影響以外に、異なる雨量種別の（地上雨量）の出水（著名洪水）を学習データに追加したことも影響した可能性が考えられる。

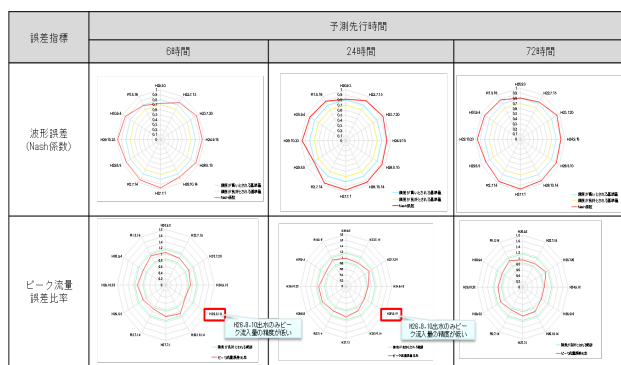
(3) 予測精度の評価

ダム管理に求められる精度が確保できているか確認するため、波形誤差（Nash係数）、ピーク流量誤差比率、洪水量到達時刻差、ピーク流量生起時刻差の誤差指標や予測波形により、AIモデルの予測精度を評価した。なお、評価対象とする出水はレーダ全国合成開始以降の検証対象出水とする。Nash係数およびピーク流量誤差比率については、既往研究において予測精度を定量的に評価することができる基準値が設定されている。既往研究で設定された基準値を基に予測精度を定量的に評価した。

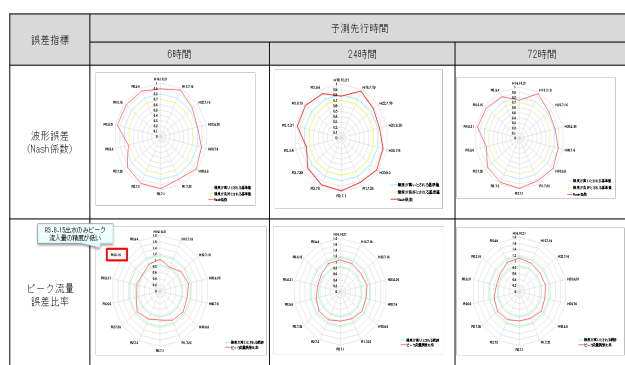
新豊根ダムでは2003.8.9出水、横山ダムでは2014.8.10出水に対し、ピーク流入量の予測精度が低い、その他の出水に対しては誤差指標の基準値を満足しており、予測精度は良好であった。丸山ダムでは、2021.8.15出水の6時間先予測のみピーク流入量の予測精度が低い、その他の出水に対しては誤差指標の基準値を満足しており、予測精度は良好であった。（図-3）



(a) 新豊根ダム



(b) 横山ダム



(c)丸山ダム

図-3 波形誤差とピーク流量誤差比率のレーダーチャート

4. 学習対象出水の拡張による感度分析

丸山ダムを対象に3項目において、感度分析を実施した。学習対象出水の拡張による感度分析については、大規模出水の精度向上を図ることを目的に4case（case1レーダ雨量全国合成以降の出水のみ、case2レーダ雨量全国合成以降の出水+既往最大洪水、case3レーダ雨量全国合成以降の出水+既往最大洪水+洪水量以上の著名洪水、case4レーダ雨量全国合成以降の出水+既往最大洪水+洪水量以上の著名洪水+洪水量以上の著名洪水の引き伸ばし）実施した。大規模出水では、case3で同等規模の大規模出水を追加することで、最も精度が向上する。一方、case4のように引き伸ばした降雨による流出計算結果を学習対象出水とすると、ピーク付近での予測精度が低下した。中小規模出水は、case1の予測精度が高く、case2以降では出水の立ち上がりで予測精度が低下したとの結果となったため、丸山ダムでは総合的な観点からcase3が最も最適であると考えられる。

学習出水数による感度分析については、学習出水数を変化させた学習データを複数作成し、予測精度の比較を行った。大規模出水では、出水数72のケースにおいて、ピーク付近の予測波形のずれが発生し、中規模出水では、出水数72のケースにおいて、ピーク流入量の低下が発生した。学習出水数を増減させると、出水規模別のバランスが悪くなり、結果として予測精度が低下したと考えられるため、選定する際には、流量規模毎の出水数のバランスが重要となり、過去に発生した出水の生起頻度を確認し、出水規模毎の出水数が同等となることに留意して学習対象出水を選定する必要があると思われる。（図-4）

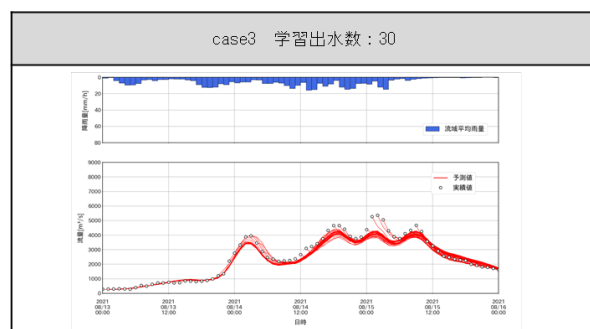
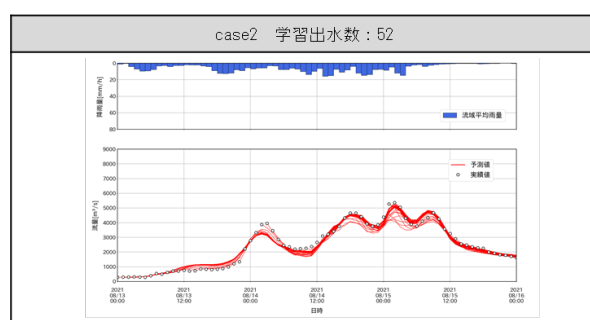
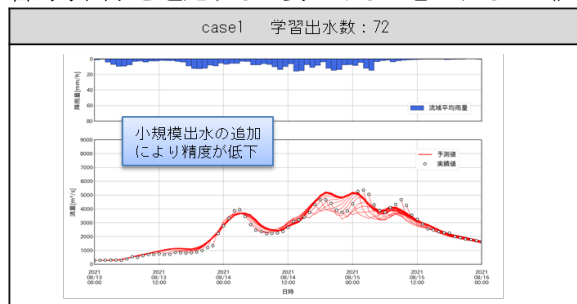


図-4 学習出水数による感度分析結果

降雨生起要因に応じた学習対象出水の分類による感度分析については、降雨生起要因別にAIモデルを構築することで、精度の改善がみられるかを検討することを目的に台風性及び前線性で実施した。

降雨生起要因が前線性であった出水に対して、「前線性の出水のみで学習したモデル」と「全ての生起要因の出水で学習したモデル」で予測を行った結果、前線性のみに学習した場合にピーク付近での予測精度が低下する傾向が見られた。降雨生起要因が台風性であった出水に対して、「台風性の出水のみで学習したモデル」と「全ての生起要因の出水で学習したモデル」で予測を行った結果、台風性のみに学習した場合にピーク付近での予測精度が低下する傾向が見られた。学習対象出水を同じ降雨生起要因に限定することで予測精度の向上を図ったが、予測精度の向上は見られなかったため、学習対象出水は、降雨生起要因に関わらず、流量規模毎の出水数のバランスに留意して選定することが重要であると思われる。

5. 既存流出計算モデルとの精度比較

AIモデルは、出水規模が大きい低頻度洪水に対して予測精度が低下する傾向にあり、今回対象とした3ダムでも同様であった。そのため、これらの出水に対しては既存流出計算モデルによる補完により精度向上の可能性を検討する必要があると考えられる。

また、上記の出水以外にも、予測精度は良好であるが、予測流入量が実績の出水若干ピークに届いていない出水などは、既存流出計算モデルで補完することにより予測

精度が向上する可能性がある。誤差指標および予測波形により、予測先行時間毎の既存流出計算モデルの予測精度を確認した結果、6時間先までの予測精度は高いが、それ以降の予測先行時間では予測精度の確保が難しい結果となった。そのため、6時間先までであれば、既存流出計算モデルによる補完が可能であると考えられる。

そのため、既存流出計算モデルは、6時間先まで精度確保が可能であることを踏まえ、AIモデルと既存流出計算モデルの予測結果に対し、6時間先までの予測波形を比較した上で、誤差指標により、両モデルの予測精度を比較したところ、既存流出計算モデルはAIモデルと比較して、6時間先であれば予測精度が良好であることが確認できる。

以上の結果に基づき、流量規模に応じた予測モデルの選定方法を整理したところ、ピーク流入量が洪水量を超過しない規模の出水はAIモデルの予測結果を採用、ピーク流入量が洪水量を超過する規模の出水は6時間先までは、既存流出計算モデルの予測結果を採用、6時間先以降は、AIモデルの予測結果を採用することで予測精度の確保が可能であると推定される。

6. とりまとめ・課題

学習対象出水の選定や入力項目の選定、AIモデルの深層学習において、上記の工夫を行うことで、精度の高い予測が可能なAIモデルを構築できた。一方、一部の出水に対して、ピーク時の精度に課題があり、いずれも低頻度な大規模出水であり、AIモデルの学習データに同程度の規模の出水データが多く含まれていないことによる学習不足が原因と考えられる。

降雨生起要因に応じて学習対象出水を区分しても、AIモデルの精度確保は難しいと考えられる。学習対象出水は、降雨生起要因に関わらず、流量規模毎の出水数のバランスに留意して選定することが重要である。

対象ダム流域において、大規模出水の発生頻度が低い場合、大規模出水に対してAIモデルの予測精度が低くなる恐れがある。流量規模に応じて、既存流出計算モデルで補完するなどにより、低頻度・未経験洪水の予測精度を確保する必要がある。また、両モデルの併用方法を洪

水予測システムに実装することで、オンラインでの予測計算の信頼性を向上させることが望ましい。

治水協定の締結により、洪水調節容量を持たない多くの利水ダムにおいても事前放流を行うこととなり、治水機能を有するダムでは上流ダムの事前放流の影響を・効果を考慮した洪水調節を行うため、ダム流入予測精度の向上が必要である。上流に位置するダムが事前放流操作や洪水時操作（貯留・遅らせ）を実施する場合、放流操作に応じた期間（事前放流操作、洪水時操作）に分け、2種類のAIモデルの構築が必要となる。しかし、上流ダムの事前放流操作の実施頻度が低い場合、学習データが不足するため、AIモデルの予測精度の確保が困難となるため、学習データの蓄積を待つ必要がある。また、AIモデルを搭載した予測システムの運用開始以降、上流ダムが事前放流操作を実施した場合、上流ダムの洪水時操作がAIモデルの学習期間における洪水時操作と異なることが想定される。上流ダムの洪水時操作が学習期間と異なると、予測対象ダム流入量の予測精度に影響を及ぼす可能性があり、その影響をダム管理者が随時確認する必要がある。

今後の展開として、今回検討が対象ダムのみでのAIを用いた流入予測であるため、上流ダムの流入予測もAIを用いた予測とし、予測システムに組み込むことで、流域での予測精度が向上すると思われる。

また、本省において、AIを活用した流入量予測、操作支援機能を実装するにあたり、基本的な考え方、現時点で最低限押さえておくべきことをとりまとめた「ダム操作へのAI活用の技術的留意事項（骨子）」の「洪水時の流入量予測モデルの技術的留意事項」の構成に対し、AI技術の適用方法や留意事項を反映させ、中部地方整備局版としてとりまとめた。この検討により取りまとめた結果を地整内の各ダムに展開していくとともに、適宜ブラッシュアップを図っていくことで、さらなる地整内のダムにおける予測精度の向上につながっていくと推定される。

さらに、AIモデルの構築、実装後に発生する出水をモデルに追加学習することにより、構築時点はデータ不足となっていた低頻度・未経験洪水においても、データ不足による予測精度の低下を改善することができると考えられる。そのため、更なる予測精度の向上に向けて、構築後においても引き続き、モデルの改良を行っていく。