

日本の水力発電効率化のための機械学習を用いたダム流入水量予測 Dam Inflow Prediction Using Machine Learning to Improve the Efficiency of Hydropower Generation in Japan

吉見 圭祐¹⁾ トリスタンハスケト¹⁾ 滝口 哲也²⁾ 高島 遼一²⁾ 大石 哲²⁾
Keisuke Yoshimi Tristan Hascoet Tetsuya Takiguchi Ryoichi Takashima Satoru Oishi

1. はじめに

日本は山がちな地形と多雨な気候から、水力発電のポテンシャルが高く、第二次世界大戦前まで、日本は水力発電に大きく依存していた。しかし、高度経済成長期になると、エネルギー消費量の増大が叫ばれ、化石燃料による発電は、急激な需要増に迅速かつ効率的に対応することができるため、この需要に応えるようになったのである。その後、オイルショックを機に、日本はエネルギーの自給自足という戦略的な理由から原子力発電を開発した。しかし、原子力発電所の事故や国際的な二酸化炭素排出量削減の要請により、現在のエネルギーミックスの長期的な存続が危ぶまれるようになった。太陽光発電や風力発電が推進される中、その断続的な性質から、単独では不十分との見方が広がっている。水力発電は低炭素であり、多様な水力発電所はベースロードとデマンドレスポンスの両方をもたらすことができる。そのメリットに対して、日本のダム運用の最適化を求める声が上がっている。

ダム運用の根本的な問題は、二つの相反する目的である。エネルギー生産と灌漑では、貯水池の水位を高く保つことで利益を得ている。水位が高ければ高いほど、位置エネルギーが蓄積されるため、一定の放流量に対してより多くのエネルギーを作り出すことができる。それに対して洪水調節では、貯水池の水位が低いほど、流入する水をより多く保持できるため、洪水の可能性を防ぐために高流量の水をよりよく緩衝することができるという利点がある。最適な運用方針とは、洪水のリスクを最小限に抑えながら、エネルギー生産量を最大化するものである。しかし、貯水池の水位を常に適切に保つには、貯水池に流入する水の量を事前に知ること、つまり正確な河川流量予測が必要である。特に、エネルギー生産のためにダム運用を最適化することは、ダム決壊のリスクを伴う。ダムの決壊とは、大量の水が貯水池に流れ込み、ダムの容量が一杯になったときに起こる現象である。このような場合、ダムのインフラが水の重みで崩壊するのを防ぐために、貯水池を残酷にも空にし、下流域を洪水にして住民に多大な犠牲を強いることになる。信頼できる流量予測がない場合、貯水池のゾーニング運用戦略がしばしば採用される。図1に示すように、貯水池のゾーニングは貯水池の各水位に異なる目的を割り当てることである。最低水位は、貯水池からの取水口の最低敷高で通常これよりも下の貯留水が利用できない水位。平常時最高水位は、平常時(非洪水時)にダムによって貯留することとした流水の最高水位である。

正確な河川流量予測が可能であれば、ダム運用者はダムの決壊を確実に防ぎながらエネルギー生産を最大化するために、下部の平常時最高水位という安全閾値を超えて貯水

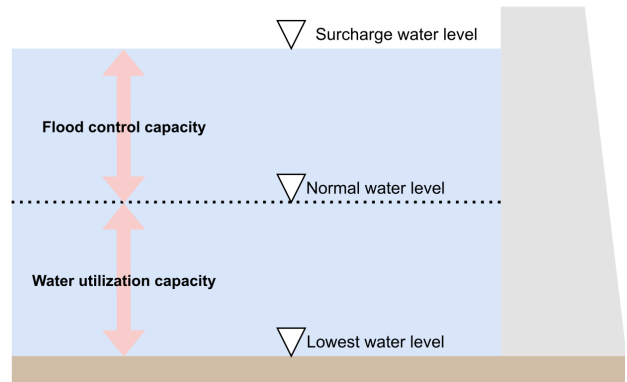


図1 ダムの運用例

池の水位を高く維持することができるようになる。一方、不正確な河川流量予測は、予期せぬ突然の流入時にダムの決壊を避けるために、ダム運用者に貯水池内の水位を安全閾値以下に維持させ、結果として効率の悪い発電政策となる。このように、ダムの発電効率は、入力される河川流量の精度によって本質的に制限されている。

本研究の最終的な目的は、機械学習を用いてダム運用におけるエネルギー生産を効率化することができるかどうかを明らかにすることである。ここで、ダムへの流入水量の予測精度が向上すれば、ダム運用の最適化に繋がると仮定することで、本稿では特に、機械学習を用いて日本のダムへの流入水量予測の予測精度を向上させることが可能かどうかという問題に限定して調査を行い、その結果を示す。

2. 関連研究

本研究は文献[1][2]に影響を受けている。文献[1][2]では、日本のすでにあるダムで運用を工夫することで、水力発電で発電される電力量を大幅に増大させることができると主張している。このような主張を受け、本研究は機械学習を用いて水力発電における発電効率を向上させることが可能か、ひいては、機械学習を用いてダムへの流入水量の予測精度を向上させることが可能であるか明らかにすることを目標としている。このプロジェクトは、水文モデリング、機械学習といったテーマを含む。以下、これらのトピックが交わる点で、関連する研究を紹介する。

水文学は地球上の水循環を主な対象とし、そのプロセスをモデル化する地球科学の一分野であり、特に、河川流量の計算などが研究されている。水文学のモデルは、地球規模か地域規模か、物理ベースかデータ駆動型か、という二つのスケールで分類することができる。全球河川流下モデル: Catchment-based Macro-scale Floodplain model (CaMa-Flood)[3]は、高度な物理シミュレーションと最小限のデータキャリブレーションに準拠しており、あらゆる場所で適用される分布型物理モデルである。また、集中型

1) 神戸大学大学院システム情報学研究所

2) 神戸大学都市安全研究センター

モデルは地域規模のモデルであり、水文学的な仮定のもと局所的に最適化されている。データ駆動型とも呼ばれる経験的モデルは、これらの物理シミュレーションを用いたモデルとは異なり、入力と出力の間の統計的な関係を用いる。その為、経験型モデルは入力の精度に大きく依存するという特徴がある[4]。

3. 水文モデル

ダムへの流入水量の予測を行う水文モデルを以下のように定義する。

$$F: X \rightarrow Y$$

X : 過去、及び予測時刻までの降水量、気温データ

Y : 三日後のダムへの流入水量[m³/s]

本論文でベースラインとして用いる線形モデルを図 2 に示す。モデルに与えるデータは、過去 20 日間、予測日時までの 3 日間の計 23 日間のものとする。使用する気象データは降水量・気温のみとする。ベースラインとする線形モデルでは、各時間ステップでの集水域内での気象データを一次元のベクトルに変換し、主成分分析により次元削減を行った後、標準化する。これを重回帰分析の説明変数として、三日後のダムへの流入水量を予測する。

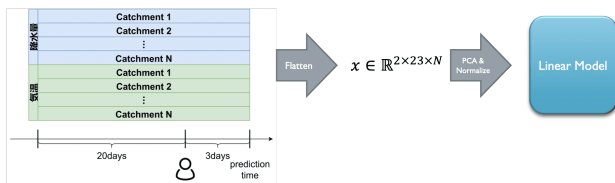


図 2 ベースラインモデルの概要

4. 水文モデルの入力に関する実験

本章では、線形モデルを用いて流入水量の予測精度に影響を与える要因を調べる。実験条件として、使用するデータのうち 2018 年 9 月 30 日までのおよそ 10 年間のデータを学習データ、それ以降のおよそ 1 年半のデータをテストデータとして使用する。また、日本にあるダムのうち 87 個のダムそれぞれに単一のモデルを学習させ、実験を行う。

4.1 融雪データの追加

本節では、線形回帰を用いた水文モデルの入力変数に、降水量と気温の他、雪に関するデータを加えて実験を行い、その影響について分析を行う。

線形モデルへの入力として、ベースラインモデルの入力として用いる降水量・気温の他、融雪量を加えて、実験を行う。降雪量のデータは気象庁にて公開されているメソ数値予報モデル GPV(MSM)のものを用い、融雪量には、TE-Japan の Web サイトで公開されている陸面課程モデル MATSIRO のシミュレーションによるものを用いる。また評価指標には、RMSE, NSE, KGE を使用する。

ベースライン、及び融雪量を追加した場合の予測精度を表 1 に示す。RMSE, NSE, KGE 全ての評価指標で融雪量データを加えた場合がベースラインを上回った。また図 3、図 4 にはそれぞれの場合における NSE の分布を示している。これを確認すると、ベースラインでは日本海側や北海道に位置するダムで予測精度が悪くなっているのに対し、融雪量を加えた場合では同地域で全体的に NSE が向上し

ていることがわかる。このことから、日本におけるダムの流入水量予測は、降雪に大きく影響を受けることが明らかになった。

表 1 : 融雪データを追加した実験結果

Input data	RMSE	NSE	KGE
Baseline	12.776	0.440	0.505
+ snow melt	11.921	0.506	0.626

4.2 代替入力変数の提案

前節では、降雪量、融雪量といった雪に関するデータを入力変数として加えることで、線形モデルの予測精度を向上させられることを確認した。しかしながら、ここで用いた融雪量のデータは物理シミュレーションを用いた陸面課程モデル MATSIRO によって計算され、公開されているデータである。そのため、実用においてリアルタイムでデータを取得し運用するのは難しい。そこで本節では、融雪量に変わって雪の影響を考慮することのできる入力変数を提案する。

4.2.1 季節性の考慮

第一節にて線形モデルの入力として用いた融雪量のデータのうち、北海道の岩尾内ダムの集水域内におけるデータをプロットしたものを図 5 に一例として示す。これを見ると、どの年においても特定の期間でピークがあり、周期性があることがわかる。

そこで、流入水量の予測を行いたい日時における月を新たに線形モデルの入力変数として加えることを考える。これによってモデルは雪解けの影響を受ける時期において、

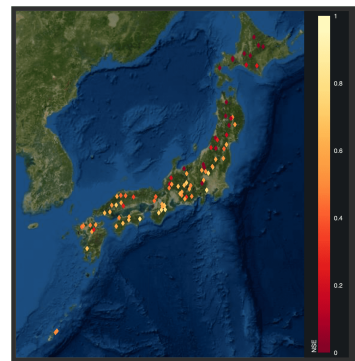


図 3 ベースラインモデルの NSE 分布

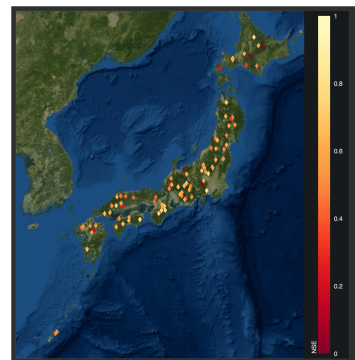


図 4 融雪量を加えた場合の NSE 分布

経験的に適切なバイアスを加え、予測精度を向上させることが期待される。ただし、月は 1 月から始まり 12 月と続き、再び 1 月に戻る。このように周期的な変動のある特徴量は線形なモデルではうまくその傾向を捉えられないと考えられる。そこで、図 6 に示すように、one-hot encoding したものを入力として用いることとする。

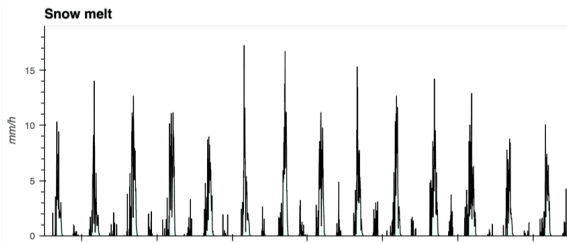


図 5 岩尾内ダムの集水域における融雪量

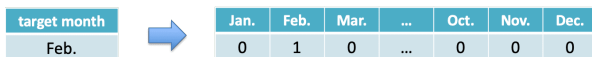


図 6 季節性の One hot encoding

4.2.2 過去の流入水量の利用

前項では、年単位での雪解けの傾向を確認するとともに、その周期性を用いた新たな入力変数を提案した。本項では、より狭い時間的スケールにおける融雪量の傾向を分析し、それに基づいた代替変数を提案する。ここで、北海道岩尾内ダムの集水域内における、2018 年 10 月 1 日以降の降水量と融雪量を面グラフにしたものを図 7 に示す。図より、やはり 4 月から 6 月頃にかけて雪解けの影響が大きく、ピーク時には集水域での水の収入の 9 割以上を雪解けによる水が占めていることも見て取れる。このことから改めて、雪解けの影響を考慮することが必要不可欠であると言える。また、融雪量の変動は降水量の変動に比べて比較的スムーズである。そのため、過去の予測を行う時刻において、その直近で観測された流入水量を入力に用いることで、モデルが融雪量による流入水量の変動を捉えることが期待される。ここで、使用するデータの日数は降水量、気温と同様に 20 日間分用いることとする。

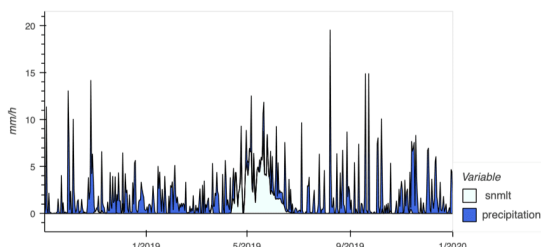


図 7 融雪量と降水量の変動の傾向

4.2.3 代替入力変数を用いた実験結果

線形モデルの入力変数として、実際に対象となる月、過去の流入水量を加えた場合の実験結果を表 2 に示す。どちらを加えた場合でも、ベースラインである降水量と気温を入力とした場合と比べて精度の向上が見られる。過去の流入水量を加えた場合では、NSE、KGE の二つの評価指標で第 1 節表 1 に示した融雪量を加えた場合よりもよい結果となった。さらに、対象の月と過去の流入水量をどちらも追加した場合は、全ての評価指標で融雪量を加えた場合を上回る結果となった。

表 2 代替変数を追加した実験結果

Input data	RMSE	NSE	KGE
Baseline	12.776	0.440	0.505
+target month	12.636	0.411	0.581
+past discharge	11.994	0.511	0.646
+target month, past discharge	11.829	0.515	0.651

4.3 降水量のデータソースによる影響

本節では、異なる手法で観測された降水量のデータそれぞれで実験を行い、その影響を調査する。

4.3.1 使用する降水量データ

- メソ数値予報モデル (MesoScale Model: MSM^{*1}):
メソ数値予報モデルは、気象庁で運用されている日本とその近海を計算領域とした気象モデルであり、観測されたデータを用いてリアルタイムのデータ同化を行なった数値予報を提供している。
- 雨量計 (gauge):
国土交通省の雨量計を用いて観測されたデータ^{*2}。ティーセン法による空間補完を行なっている。
- 衛星全球降水マップ (Global Satellite Mapping of Precipitation: GSMaP^{*3}):
GPM 主衛星に搭載された二周波降水レーダ (DPR) を中心に、複数の降水を観測する衛星や静止気象衛星を組み合わせることで開発された世界の雨マップである。雨量計補正版では、気象庁気象台の雨量計データをもとに補正されたデータが公開されている。

4.3.2 実験結果と考察

各雨量データを線形モデルの入力として用いた場合の実験結果を表 3 に示す。結果を見ると、雨量計によるデータを用いた場合で最も良い結果となった。また、GSMaP では、MSM や雨量計データを用いた場合よりも大幅に精度が落ちている。これは、GSMaP が全球降水マップであり、世界中で同一品質であることが求められているために、日本に特化したデータよりも観測精度が低いのではないかと考えられる。しかし、将来的に日本以外で本論文と同様にダムへの流入水量予測を行う場合、国によっては雨量計による観測データが公開されていない場合がある。こういった地域においては GSMaP での降水量データの使用が有効である可能性も考えられる。

*1 <http://database.rish.kyoto-u.ac.jp/arch/jmadata/gpv-original.html>

*2 <http://www1.river.go.jp>

*3 <https://sharaku.eorc.jaxa.jp/GSMaP/index.html>

表 3 異なる降水量データを用いた実験結果

Precipitation data	RMSE	NSE	KGE
MSM	11.829	0.515	0.651
Gauge	10.205	0.628	0.727
GSMaP	16.256	0.251	0.424
GSMaP(adjusted by gauge)	14.517	0.371	0.519

5. 非線形モデルを用いた流入量予測

前章では、線形モデルを使用して、ダムへの流入水量に入力データが大きく影響を与えることを明らかにした。本章では、非線形モデルを用いて実験・評価し、モデル性能が予測に与える影響について調査する。

5.1 実験条件

LightGBM, XGBoost, multi layer perceptron(MLP)を新たに用いて実験を行う。モデルの入力には第 4 章と同様に計 23 日間の降水量と気温のデータを使用する。ただし、降水量には、最も予測精度の高かった雨量計のデータを使用する。また、追加の入力として過去の流入水量と one-hot encoding した対象の月のデータを用いる。

5.2 実験結果と考察

非線形モデルを使用した実験の結果を表 4 に示す。これを見ると、殆どの場合で非線形モデルの予測精度が線形モデルの精度を上回る結果となった。特に、MLP は RMSE, KGE の二つの評価指標で最も良い結果となった。

非線形モデルを用いることで予測精度を向上させることができたが、その影響は入力データ（特に降水量）を変えた時の影響に比べると小さい。このことから、ダムへの流入水量の予測においては、降水量の観測・予測精度が重要であると考えられる。

表 4 非線形モデルを用いた実験結果

Model	RMSE	NSE	KGE
Linear	10.205	0.628	0.727
LightGBM	9.667	0.662	0.706
XGBoost	9.960	0.671	0.738
MLP	8.823	0.637	0.751

6. データ分析アプリケーションの開発

本論文では、日本にある 87 個のダムで、流入水量の予測を行う水文モデルの実験を行ってきた。その予測精度は降雪に大きく影響を受けていることを第 4 章で示したが、降雪の他にも台風による大雨など、流入水量に影響を与える要因は存在することが考えられる。しかしこれらは多くの場合、ダムの位置する場所によりその影響度合いが異なる。そのため、ダムごとに、予測した流入水量やその時予測に用いた降水量などのデータをそれぞれ分析することがモデル性能を向上させる上で非常に重要である。そこで、このような分析をインタラクティブに行うことのできる図 8 のような Web アプリケーションを開発した。

作成した Web アプリケーションでは、流入水量の予測に用いるモデル、使用する降水量データの種類、モデルに追加で加える入力変数といった、様々な実験条件を指定することができ、この条件に応じた予測結果、評価指標の値などを確認することができる。これによって、高度な分析を効率的に行うことが可能となる。

7. 結論

正確な流入水量は、ダム運用の最適化を行う上で重要である。そこで本稿では、機械学習を用いたダムへの流入水量の予測において、その予測精度に影響を与える要因について調査を行い、以下のことを明らかにした。第一に、雪に関するデータを予測に用いることで高緯度の地域で予測精度を向上させることが可能であること。第二に、過去の流入水量や季節の情報を与えることが流入水量の予測で有効に働くこと。第三に、降水量のデータの品質が予測精度に最も影響を与えること。最後に、モデルの性能は予測精度に影響を与えるが、その影響は降水量のデータ品質による影響に比べると小さいことを明らかにした。また、効率的なエラー分析のためのアプリケーションを作成した。

参考文献

- [1] 竹村 幸太郎, “水力発電が日本を救う-今あるダムで年間 2 兆円超の電力を増やせる”, 東洋経済新報社 (2016).
- [2] 角 哲也, 井上 素行, 池田 駿介, 上阪 恒雄, 国土文化研究所, “今こそ問う 水力発電の価値 —その恵みを未来に生かすために,” 技報堂出版, (2019).
- [3] 山崎 大, 鼎 信次郎, 沖 大幹, “全球超高解像度水文地形データを利用した河川・氾濫原シミュレーション,” 水工論文集, 第 54 巻, (2010).
- [4] Sitterson, J., Chris Knights, R. Parmar, K. Wolfe, M. Mucbe, B. Avant, “An Overview of Rainfall-Runoff Model Types,” U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-17/482, (2017).

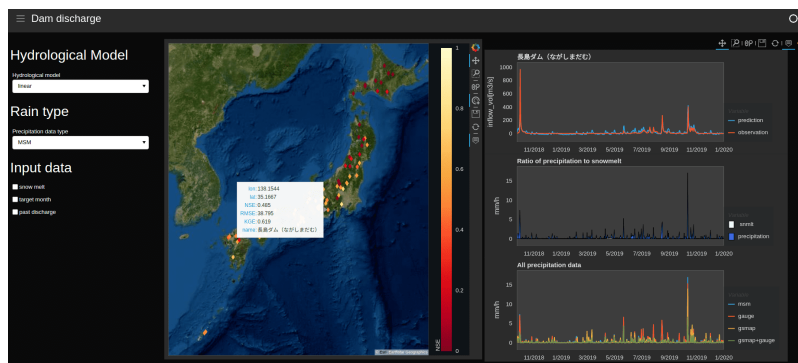


図 8 開発したアプリケーションの画面