

Elastic Net を用いた大河川下流部の水位予測の研究

Research on water level prediction for downstream part of large-scale rivers using the Elastic Net

室蘭工業大学 ○学生員 若狭谷昇真 (Shoma Wakasaya)
 室蘭工業大学 学生員 山洞智弘 (Tomohiro Sando)
 室蘭工業大学 正員 中津川誠 (Makoto Nakatsugawa)
 室蘭工業大学 非会員 小林洋介 (Yosuke Kobayashi)

1. 研究背景

2016年8月、北海道において4つの台風が連続して上陸及び接近し、広域にわたって甚大な被害をもたらした。このような水害を事前に予測し被害を軽減させるためには、水防活動、避難場所の開設、住民の避難といった防災・減災対策に要する予測先行時間（リードタイム、以下、LT と称す）の確保が望まれる。特に市街地に近い河川では洪水時における早めの適切な判断が住民の安全に影響するため、LT が長く、水位上昇のタイミングやピーク水位について精度のよい水位予測が望まれる。

現在、国土交通省で導入が進められている水位予測手法は、対象流域をメッシュに分割して河道追跡と組み合わせた分布型流出モデル（土研モデル¹⁾）を基に、予測値と観測値の誤差修正を行うフィードバック手法（粒子フィルタ²⁾）が主流である。しかし、多数のパラメータの設定が困難であり、必ずしも満足の結果が得られないことが課題である。

そこで、最近では機械学習手法により、水位そのものを使用し予測する方法が注目されている。これまで、一言ら³⁾は深層学習、山田ら⁴⁾は Long short-term memory を用いた Recurrent Neural Network、坂本ら⁵⁾はランダムフォレスト（以下、RF）による水位予測を行い、その有用性を示している。中でも、ブラックホールの撮像に用いられたスパースモデリングが、少ない入力データから現象を高度に再現する方法として注目されている。山洞ら⁶⁾は Elastic Net を用いて未経験洪水事例の24h先までのダム流入量の予測を行い良好な結果を得た。

本研究では、過学習を抑制するために説明変数を選択することが可能な Elastic Net の利点を生かし、北海道を代表する大河川の水位予測を試みた。

2. 研究方法

2.1 対象地点と対象事例

本稿では、北海道を代表する大河川で、過去に大きな洪水を経験している石狩川、十勝川、鶴川、沙流川の4河川を対象に水位予測を行った。図-1に予測地点の位置図、左下には対象流域の北海道における位置図を示す。予測対象事例は、水位データを水文・水質データベース⁷⁾から取得し、ピーク水位が最大となった事例とした。とくに、水位予測の対象とした2016年の十勝川の洪水は、4つの台風が連続して上陸し、茂岩地点で計画高水位を超えた事例である。また、2003年の沙流川の洪水も、富川地点で計画高水位を超え、二風谷ダムの異常洪

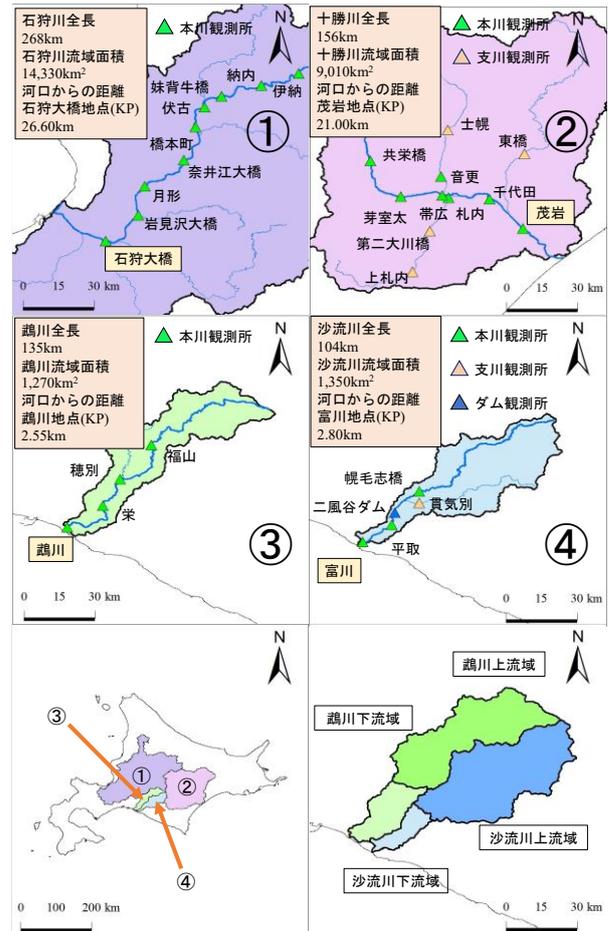


図-1 水位予測対象河川流域の位置と概要

水時防災操作が行われた特筆すべき事例である。

2.2 目的変数とその予測に用いる説明変数

本研究では6時間先及び12時間先の水位を目的変数とする。また、表-1に予測に用いる説明変数を示す。使用するデータは予測対象地点の上流部にある「観測水位」を用いるが、鶴川と沙流川は上流にある水位観測所が共に4か所のみとなっているので、「流域平均雨量」、「土壌雨量指数」も説明変数に加えた。この際、降雨の空間分布の影響が考慮できるように、支川と本川が合流する地点で2つの流域に分けて値を算出した。図-1の右下に流域を分割した図を示す。また、沙流川はダムによる影響を考慮し、二風谷ダムの「流入量」、「放流量」も説明変数に加えた。

以上で用いる「観測水位」、「流入量」、「放流量」

表-1 予測のための説明変数諸元

石狩川予測における使用説明変数の諸元					
予測対象地点	石狩大橋				
観測項目	観測所	KP(km)	データ概要	データ数	説明変数の数
観測水位 (m)	石狩大橋	26.60	t-6~tの時間値	7	63
	岩見沢大橋	44.50			
	月形	58.00			
	奈井江大橋	76.80			
	橋本町	93.90			
	伏古	104.78			
	妹背牛橋	113.50			
	納内	130.25			
	伊納	148.50			
全説明変数	63				
学習事例数	9事例				
十勝川予測における使用説明変数の諸元					
予測対象地点	茂岩				
観測項目	観測所	KP(km)	データ概要	データ数	説明変数の数
観測水位 (m)	茂岩	21.00	t-6~tの時間値	7	77
	千代田	37.60			
	帯広	56.70			
	札内	57.40			
	東橋	59.40			
	菅更	65.30			
	芽室太	71.10			
	第二大川橋	74.10			
	士幌	86.20			
	共栄橋	93.40			
	上札内	95.20			
	全説明変数	77			
学習事例数	9事例				
鶴川予測における使用説明変数の諸元					
予測対象地点	鶴川				
観測項目	観測所	KP(km)	データ概要	データ数	説明変数の数
観測水位 (m)	鶴川	2.55	t-6~tの時間値	7	28
	栄	25.98			
	穂別	41.10			
	福山	66.80			
流域平均雨量 (mm/h)	鶴川上流域 鶴川下流域		t-7~t-6, ..., t-1の時間値	7	14
土壌雨量指数 (mm)	鶴川上流域 鶴川下流域		tの時間値	1	2
全説明変数	44				
学習事例数	9事例				
沙流川予測における使用説明変数の諸元					
予測対象地点	富川				
観測項目	観測所	KP(km)	データ概要	データ数	説明変数の数
観測水位(m)	富川	2.80	t-6~tの時間値	7	28
	平取	15.55			
	貴賀別	33.40			
	幌毛志橋	34.30			
流域平均雨量 (mm/h)	沙流川上流域 沙流川下流域		t-7~t-6, ..., t-1の時間値	7	14
土壌雨量指数 (mm)	沙流川上流域 沙流川下流域		tの時間値	1	2
放流量 (m ³ /s)	二風谷ダム	21.20	t-6~tの時間値	7	7
流入量 (m ³ /s)	二風谷ダム		t-6~tの時間値	7	7
全説明変数	58				
学習事例数	9事例				

tは現在時刻を意味する。

*LT(リードタイム):何時間先の予測を行うかの時間、いわゆる予測先行時間

は、水文・水質データベース⁷⁾より取得し、「流域平均雨量」、「土壌雨量指数」は、レーダー解析雨量を(一財)気象業務支援センターから購入し、算出した。

次に、過去の洪水事例を表-2に示す。予測に用いたデータは1998年~2019年までを収集し、過去のピーク水位上位10事例を抽出し、第1位を予測事例、その他を学習事例として使用した。対象期間はピーク水位を観測した日の前後15日の計31日間とした。

2.3 予測雨量・ダムの予測放流量の検討

これまでの報告により、石狩川及び十勝川における水位予測に際して、LTを延ばすにしたがって、回帰式を構成する重みの上位の説明変数が、より上流の「観測水位」に移行していることが確認されている⁵⁾。しかし、鶴川及び沙流川は、上流部の水位情報が少ないため、長時間の予測は難しいと予想される。

従って、この両河川のLT=12hの予測においては、表-3に示すように、実測の流域平均雨量を予測雨量に見立てて使用した「みなし予測雨量」を説明変数として加

表-2 過去の洪水事例

順位	石狩川 石狩大橋地点ピーク水位(m) (発生日時)		十勝川 茂岩地点ピーク水位(m) (発生日時)	
	事例年月	ピーク水位	事例年月	ピーク水位
1	2001年9月	6.28 09/12 07:00	2016年8月	12.68 08/31 11:00
2	2016年8月	5.33 08/21 20:00	2001年9月	9.64 09/12 12:00
3	2011年9月	4.97 09/04 17:00	2003年8月	8.93 08/10 13:00
4	2018年7月	4.82 07/04 11:00	2011年9月	8.48 09/06 12:00
5	2000年7月	4.13 07/27 08:00	1998年9月	8.46 09/17 04:00
6	2005年8月	4.03 08/22 21:00	2002年10月	8.18 10/02 17:00
7	2012年9月	3.68 09/12 19:00	2005年9月	7.03 09/04 15:00
8	1999年8月	3.63 08/03 00:00	2013年9月	6.89 09/17 06:00
9	2018年8月	3.32 08/17 14:00	2010年8月	6.79 08/13 05:00
10	2014年8月	3.26 08/06 15:00	2014年8月	6.71 08/11 22:00
順位	鶴川 鶴川地点ピーク水位(m) (発生日時)		沙流川 富川地点ピーク水位(m) (発生日時)	
	事例年月	ピーク水位	事例年月	ピーク水位
1	2001年9月	3.61 09/11 22:00	2003年8月	7.83 08/10 04:00
2	2003年8月	5.58 08/10 08:00	2006年8月	6.47 08/19 06:00
3	2016年8月	5.49 08/23 16:00	2016年8月	5.54 08/23 12:00
4	2006年8月	5.2 08/19 09:00	2010年8月	5.2 08/12 20:00
5	2000年8月	4.11 08/16 07:00	2018年7月	3.73 07/11 11:00
6	2010年8月	4.09 08/12 21:00	2005年8月	3.54 08/22 17:00
6	2005年9月	3.96 09/08 11:00	2013年9月	3.35 09/17 02:00
8	2012年9月	3.86 09/11 01:00	2009年7月	3.31 07/20 00:00
9	2013年9月	3.52 09/17 04:00	2008年7月	3.19 07/23 12:00
10	2018年8月	3.27 08/17 04:00	2014年8月	2.82 08/11 17:00

表-3 みなし予測雨量の説明変数諸元

鶴川 LT=12h				
観測項目	観測所	データ概要	データ数	説明変数の数
みなし予測雨量 (mm)	鶴川上流域	t~t+1, ..., t+11~t+12の時間値	12	24
	鶴川下流域			
沙流川 LT=12h				
観測項目	観測所	データ概要	データ数	説明変数の数
みなし予測雨量 (mm)	沙流川上流域	t~t+1, ..., t+11~t+12の時間値	12	24
	沙流川下流域			

tは現在時刻を意味する。

*LT(リードタイム):何時間先の予測を行うかの時間、いわゆる予測先行時間

表-4 みなし予測放流量の説明変数諸元

沙流川 LT=12h				
観測項目	観測所	データ概要	データ数	説明変数の数
みなし予測放流量 (m ³ /s)	二風谷ダム	t+1, ..., t+12の時間値	12	12

tは現在時刻を意味する。

*LT(リードタイム):何時間先の予測を行うかの時間、いわゆる予測先行時間

表-5 予測手法の分類とその対象流域

分類名	使用する説明変数	対象河川
分類I	表1の説明変数	石狩川, 十勝川, 鶴川, 沙流川
分類II	表1, 表3の説明変数	鶴川, 沙流川
分類III	表1, 表4の説明変数	沙流川

えたモデルを作成した。また、表-4に示すように、沙流川は実測の二風谷ダム放流量を予測放流量に見立てて使用した「みなし予測放流量」を説明変数に加えたモデルも検討した。以上を総括した予測手法の分類と適用した対象河川を表-5に示す。

2.4 Elastic Net[®]による統計モデル

スパースモデリング手法の代表的手法である Elastic Net は、正則化を用いた回帰分析手法である。ニューラルネットワーク等の手法と異なり、算出される説明変数の重み(係数)から、各要素がどれほど予測に影響するか考察することが出来る。また、不要な説明変数の重みがゼロになる正則化という特徴を持ち、従来の重回帰分析より過学習になりにくくといった特徴を持つ。Elastic Net のコスト関数であるJ(w)は以下の式で表せる。

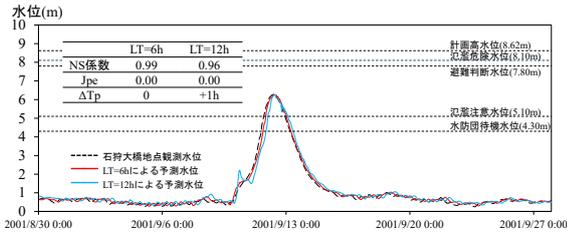


図-2 石狩大橋水位予測結果

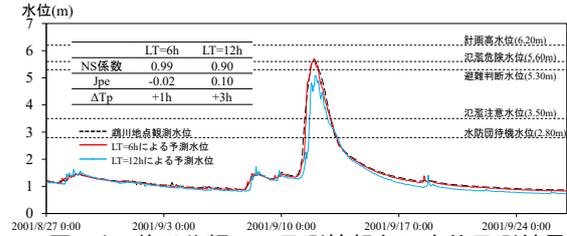


図-4 鶴川分類Ⅰ：予測情報なし水位予測結果

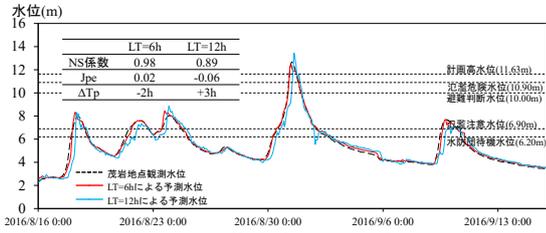


図-3 茂岩水位予測結果

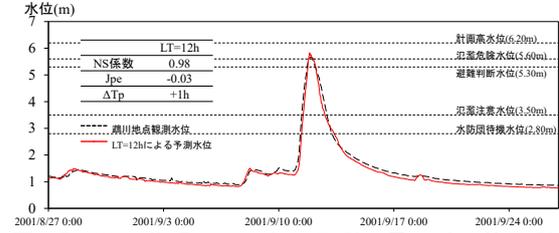


図-5 鶴川分類Ⅱ：みなし予測雨量あり水位予測結果

$$J(w) = \sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2 + (1 - \alpha)\lambda_1 \sum_{j=1}^m w_j^2 + \alpha\lambda_2 \sum_{j=1}^m |w_j| \quad (1)$$

ここで、 n は学習データの数、 m は説明変数の数、 $y^{(i)}$ は時刻*i*における実測値、 $\hat{y}^{(i)}$ は回帰式から求められる時刻*i*における予測値、 w_j は説明変数*j*の重みである。右辺の第1項は二乗誤差、第2項がL1ペナルティと呼ばれるLasso回帰で説明変数の削除を意図した制約条件、第3項はL2ペナルティと呼ばれるRidge回帰で選択する変数の個数に影響を与える制約条件である。Elastic Netは、この2つのペナルティ重み α とペナルティの大きさを制御する λ_1 、 λ_2 をハイパーパラメータ（機械学習アルゴリズムの挙動を設定するパラメータ）として最適化することで、少数のスパースな実測値からより良い回帰式を求める手法である。本手法は、学習範囲を最適化する深層学習を含む他の機械学習法がとる内挿問題最適化と異なり、線形結合式で表現されるため、外挿問題である未経験事例の予測に対して解の妥当性の検証が可能である。

本研究でのモデル学習では汎化性能を高めるため、学習データの交差検証（Cross Validation）を行なった。これはデータの分割を複数繰り返し、複数のモデルの学習平均値を最適化することで予測性能を高める手法である。またハイパーパラメータ値の最適化は並列分散最適化アルゴリズム Optuna で100試行の最適化処理で最適値を求めた⁹⁾。このように最適化されたハイパーパラメータから得られる各説明変数の重み w を用い、以下のような線形結合の回帰式を得る。

$$\hat{y}^{(i)} = w_0 + \sum_{j=1}^m w_j x_j^{(i)} \quad (2)$$

ここで、 $\hat{y}^{(i)}$ は時刻*i*における目的変数（時刻*i*+6及び時刻*i*+12の観測水位）、 $x_j^{(i)}$ は時刻*i*における説明変数*j*である。右辺の第1項は切片、第2項は回帰式の一般項を表している。

2.5 予測精度の評価指標

水位予測の精度評価は、以下に示す Nash-Sutcliffe 係数（以下、NS係数）、ピーク水位誤差（以下、 J_{pe} ）、ピーク生起時刻差（以下、 ΔT_p ）の3つを用いて行った。

$$NS = 1 - \frac{\sum (H_{oi} - H_{ci})^2}{\sum (H_{oi} - H_{oave})^2} \quad (3)$$

$$J_{pe} = \frac{H_{op} - H_{cp}}{H_{op}} \quad (4)$$

$$\Delta T_p = T_{cp} - T_{op} \quad (5)$$

ここで、 H_{oi} は観測水位（m）、 H_{ci} は予測水位（m）、 H_{oave} は観測水位平均値（m）、 H_{op} は観測ピーク水位（m）、 H_{cp} は観測ピーク水位時の予測水位（m）、 T_{cp} は予測ピーク時刻、 T_{op} は観測ピーク時刻である。NS係数は1.0に近づくほどモデルの精度がよく、 J_{pe} は値が負になると安全側の予測であることを意味し、 ΔT_p はピーク時刻が早めに推定される場合を安全側とする指標である。ここでは、 $NS \geq 0.7$ 、 $J_{pe} \leq 0.05$ 、 ΔT_p はLT=6hであれば $\Delta T_p \leq 2$ h、LT=12hであれば $\Delta T_p \leq 4$ hを評価の良し悪しを判断する目安にした。

3. 結果と考察

3.1 石狩川、十勝川の水位予測結果

石狩川及び十勝川の水位予測結果を図-2、図-3に示す。石狩大橋地点と茂岩地点でLT=6h、LT=12hにおいて共に適切な結果が得られた。とくに、十勝川の2016年洪水は計画高水位を超える未経験の事例であり、RFではピーク水位を過小評価したが⁹⁾、Elastic Netでは良好な結果を得ることができた。

以上の結果より、上流の水位観測所が多い大河川では、水位情報のみで適切な予測が可能であると示された。

3.2 鶴川の水位予測結果

鶴川の現況値のみで予測した表-5の分類Ⅰの水位予測結果を図-4に示す。LT=6hにおいては高精度の予測が可能であったが、LT=12hにおいては $J_{pe} = 0.1$ とピーク水位の過小評価がみられた。次に、予測雨量を考慮した分類Ⅱの水位予測結果を図-5に示す。すべての指標からみて適切な予測結果となっていることがわかる。

以上の結果より、LTの長い予測には上流端付近に水位観測所のない場合、予測雨量を説明変数に加えることで水位予測が可能であることが示された。従って、モデ

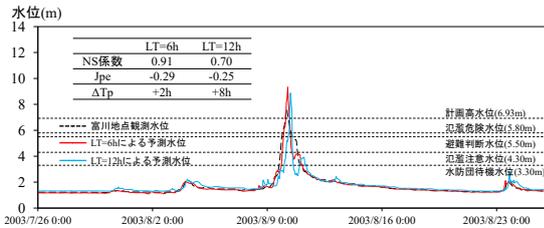


図-6 富川分類Ⅰ：予測情報なし水位予測結果

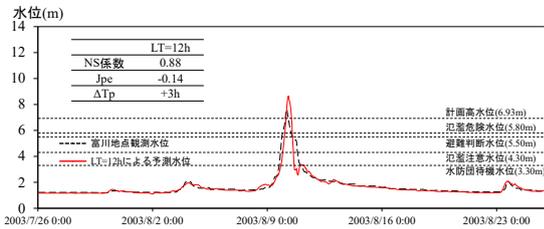


図-7 富川分類Ⅱ：みなし予測雨量あり水位予測結果

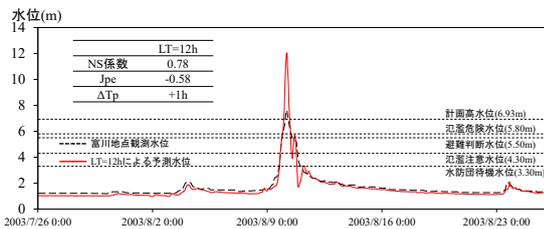


図-8 富川分類Ⅲ：みなし予測放流量あり水位予測結果

ルに組み込めるような予測雨量の精度向上が必要である。

3.3 沙流川の水位予測結果

沙流川の各分類の水位予測結果を図-6, 図-7, 図-8に示す。現況値のみで予測した分類Ⅰは、LT=6 hでは概ね適切な予測が可能となっているが、LT=12 hにおいて $\Delta T_p=8$ hと大幅な遅れがみられた。次に、予測雨量を考慮した分類Ⅱ、予測放流量を考慮した分類Ⅲの予測結果は共に遅れが改善され、適切な予測が可能となったが、分類Ⅲのほうがピーク水位の早期予測が可能であった。沙流川の2003年洪水は二風谷ダムの異常洪水時防災操作が行われた未経験の事例であったが、予測放流量を入れることで良好な水位予測が可能となることがわかった。

次にElastic Netにより学習された各説明変数の重み上位5つを表-6に示す。分類Ⅰのモデルと比べ、分類Ⅱのモデルは上流の水位情報の重みが大きくなり、分類Ⅲのモデルは将来のダム放流量の重みが大きく影響していることがわかる。

以上の結果から、ダムの予測放流量は予測雨量よりも下流部の水位予測に大きな影響を及ぼすことが示された。これは、ダムの人為的操作に影響を受けるため、ダム情報の重みが大きくなると推察できる。そのため、ダム下流部の水位予測を実務に実装するには、ダム操作ルールを考慮した放流量の予測情報が必要である。

4. 結論

本研究において得られた結果を以下にまとめる。

- 1) Elastic Netにより、上流部に水位観測情報を有する石狩川や十勝川のような大河川の水位予測において既往第1位の事例をそれ以下の情報で予測することが可能であることが示された。

表-6 沙流川における分類別説明変数の重み

順位	予測情報なし	みなし予測雨量あり	みなし予測放流量あり
1	富川現在水位	富川現在水位	富川現在水位
2	現在放流量	幌毛志橋現在水位	12時間先みなし予測放流量
3	上流域現在流域平均雨量	平取2時間前水位	幌毛志橋現在水位
4	平取2時間前水位	貴氣別現在水位	10時間先みなし予測放流量
5	平取現在水位	平取現在水位	3時間先みなし予測放流量

- 2) 鶴川、沙流川はLT=12 hの予測は、現況の情報のみでの予測は難しいが、「予測雨量」、沙流川では「ダムからの予測放流量」を説明変数に加えることで、適切な予測が可能であることが示された。上流部の水位情報が少ない河川においては、雨量やダム放流量の予測情報が重要であることが示唆された。よって、これらの精度向上を図ることが今後の課題である。

謝辞：本研究は、JSPS 科研費（課題番号 JP20K04698, FY2020~2022）及び（公財）電気通信普及財団の研究調査助成（FY2019~2020）、（一財）北海道河川財団を受けたものである。ここに記して謝意を表す。

参考文献

- 1) 伊藤弘之, 飯野光則, 平塚真里子, 大浪裕之: 洪水予測を目的とした分布型流出モデルの同定, 土木技術資料, 56-3, p2~5, 2014.
- 2) 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, Vol.67, No.4, I_511-I_516, 2011.
- 3) 一言正之, 櫻庭雅明, 清雄一: 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, 2016.
- 4) 山田恒輝, 小林洋介, 中津川誠, 岸上順一: リカレントニューラルネットワークを用いた2016年の常呂川洪水事例の水位予測, 土木学会論文集(水工学), Vol.74, No.5, I_1369-I_1374, 2018.
- 5) 坂本莉子, 中津川誠, 小林洋介: ランダムフォレスト法に基づく大河川の低頻度洪水の水位予測に関する研究, 土木学会論文集B1(水工学), Vol.74, No.5, I_1375-I_1380, 2018.
- 6) 山洞智弘, 中津川誠, 小林洋介, 坂本莉子: 未経験洪水事例に適用できるElastic Netによる24時間先までのダム流入量予測手法の提案, 土木学会論文集B1(水工学), Vol.76, No.2, I_835-I_840, 2020.
- 7) 国土交通省: 水文・水質データベース, <http://www1.river.go.jp/> (閲覧日: 2020/12/10)
- 8) Hui,Z. and Trevor,H. :Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society*, Series B67, pp.301-320, 2005.
- 9) Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. :Optuna: A Next-generation Hyperparameter Optimization Framework, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD'19, pp.2623-2631, 2019.