

流域属性に基づく流出モデルパラメータの 一般化に関する研究

株式会社建設技術研究所 ○谷岡広太郎
岡峰奈津美
日浦直紀
此島健男子
米勢嘉智

1. はじめに
2. 全国87水系におけるパラメータ探索
3. TabPFNを用いたパラメータ推定
4. 推定パラメータを用いた精度検証
5. まとめ

1.はじめに

【背景】

- ✓気候変動により豪雨の頻発化・激甚化が進行し洪水リスクが増大
- ✓洪水予測の高度化は、治水計画・避難判断を支える重要技術
- ✓中小河川や中山間地流域では観測体制が十分でない場合が多く、
観測情報の有無に依存しない洪水予測技術の確立が重要

【課題】

- ✓分布型流出モデルの、観測データに依存しないキャリブレーション
- ✓未観測流域においても、一定の精度を有する洪水予測を広域に展開

【本研究の目的】

- ✓全国規模で整理可能な地形、土地利用、植生等の流域代表属性に着目し、
土研分布型流出モデルのパラメータを一般化する手法を検討
- ✓観測データの有無にかかわらず一定の予測精度を有する広域洪水予測への
適用可能性を示すことを目的とする

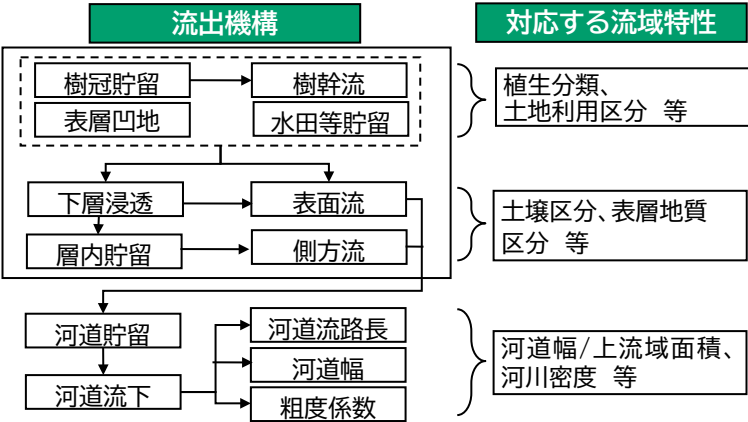
2.全国87水系におけるパラメータ探索

パラメータ探索を実施した355地点

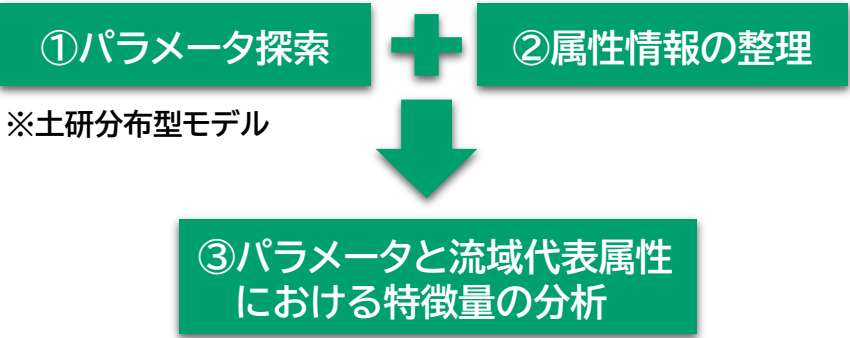
※地点毎に1～5洪水を選定



流出過程に応じた流域特性



パラメータ探索結果を用いた分析方針



収集した属性および出典

	属性の項目	出典
(1)	植生分類	環境省 自然環境局 第5回基礎調査植生3次メッシュデータ
(2)	土地利用区分	国土数値情報 土地利用細分メッシュデータ 令和3年度
(3)	土壌区分	国土数値情報 土地分類メッシュ
(4)	表層地質区分	国土数値情報 土地分類メッシュ
(5)	表層地質年代	国土数値情報 土地分類メッシュ
(6)	河道幅 (低水路幅・堤間幅)	全国流向マップ、川の防災情報
(7)	流路延長	国土数値情報 河川データ
(8)	流域面積比率 (流路延長/面積の平均値)	国土数値情報 河川データ、国土数値情報 流域界・非集水域データ、全国流向マップ
(9)	谷密度	国土数値情報 谷密度メッシュデータ
(10)	年平均降水量による 降雨分布 (国土数値情報 平年値メッシュデータ)	気象庁 メッシュ平均値図 降水量 (国土数値情報 平年値メッシュデータ)
(11)	残流域面積	国土数値情報 流域界・非集水域データ、全国流向マップ
(12)	地形勾配(斜面勾配)	国土数値情報 標高・傾斜度メッシュ
(13)	植生による被覆度	NASA EARTHData Normalized Difference Vegetation Index (NDVI) 2001-2023年での6-10月別平均値

2.全国87水系におけるパラメータ探索

パラメータ探索方法

- ✓ SCE-UA法によりパラメータ探索
- ✓ 複数の評価関数を用いる多目的最適化手法として順位合成法を採用
- 複数指標の合計順位が最小となるパラメータの組合せを最適と定義

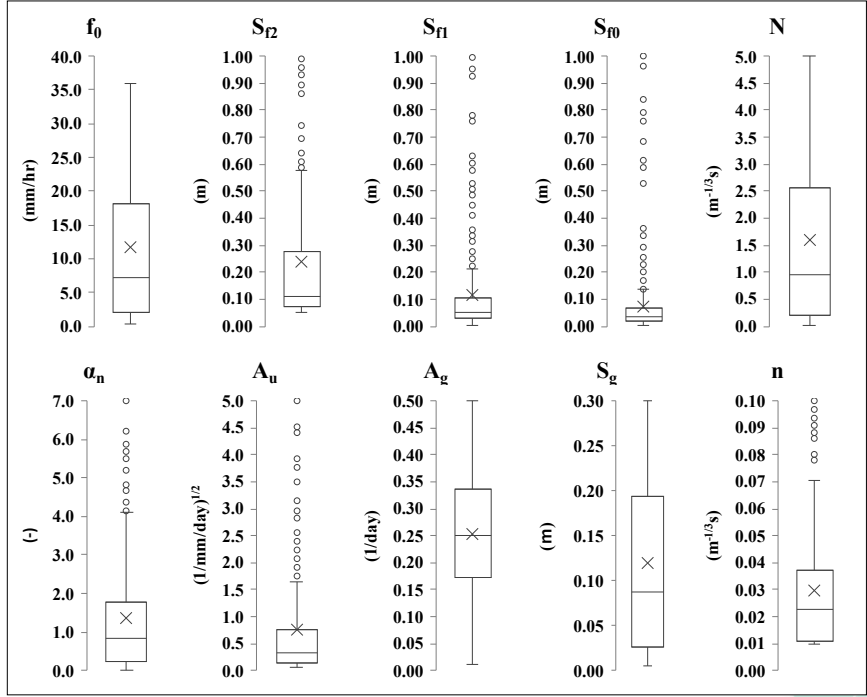
$$\Delta Q_{peak} = \left| \frac{\hat{Q}_{ipeak} - Q_{ipeak}}{Q_{ipeak}} \right|$$
 評価指標①

$$\Delta RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (Q_i - \hat{Q}_i)^2}$$
 評価指標②

ΔQ_{peak} :ピーク流量差, $\Delta RMSE$:二乗平均平方根誤差
 $\hat{Q}_{ipeak}$:予測ピーク流量, Q_{ipeak} :実績ピーク流量
 $\hat{Q}_i$:予測の流量, Q_i :実績の流量, n :データの総数

探索対象パラメータと探索結果

パラメータ		単位	探索範囲上下限
f ₀	表層から地下層への最終浸透能	cm/s	10 ⁻⁵ ~10 ⁻³
S _{f2}	表層の最大貯留高	m	0.05~1
S _{f1}	係数(表層早い流出発生高)	—	0.1~1
S _{f0}	係数(地下浸透の発生高)	—	0.1~1
N	地表面の粗度係数	m ^{-1/3} s	0.01~5
α _n	表層の速い中間流出係数	—	0.01~7
A _u	地下層の不圧流出係数	(1/mm/day) ^{1/2}	0.05~5
A _g	地下層の被圧流出係数	1/day	0.01~0.5
S _g	不圧地下水流出の発生高	m	0.005~0.3
n	河道の粗度係数	m ^{-1/3} s	0.01~0.1



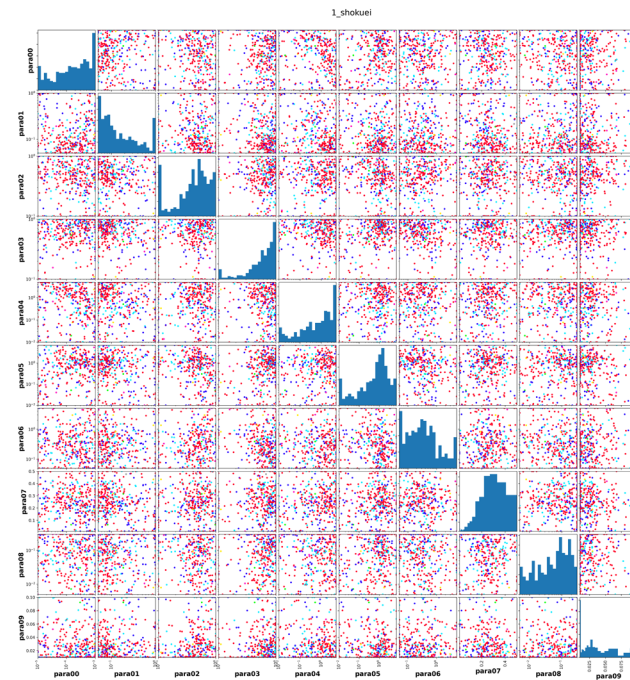
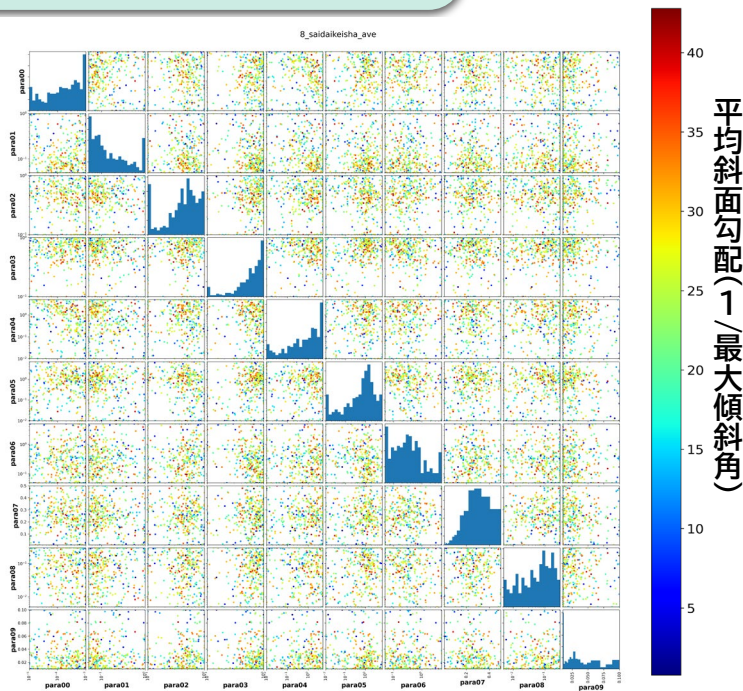
2.全国87水系におけるパラメータ探索

✓流域代表属性と最適パラメータとの間に、ばらつきを伴う弱い関係性は一部で認められたが、明確な1対1の対応関係は確認できない

【どうするか？】

属性間の複雑な相互作用の存在が想定されるため、属性とパラメータの相互作用を考慮したアプローチへ

特徴量の分析(Pairplot)



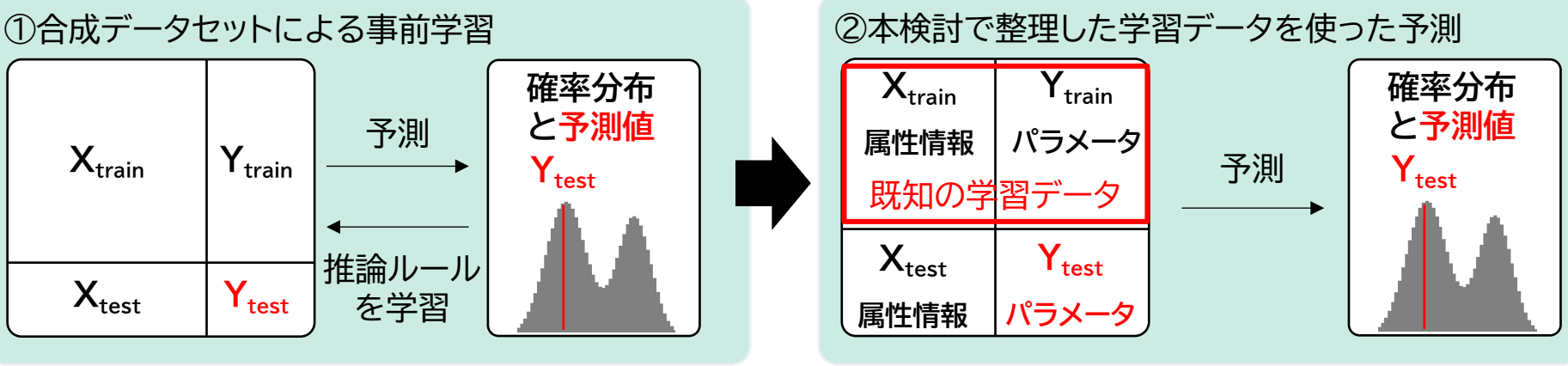
- 植生区分
- : 常緑針葉樹
 - : 常緑広葉樹
 - : 落葉針葉樹
 - : 落葉広葉樹
 - : 低木草本
 - : その他
 - : 欠測

3.TabPFNを用いたパラメータ推定

【本研究での設定】
X: 属性情報
Y: モデルパラメータ値

TabPFN(Tabular Prior-data Fitted Network)の概要

- ✓表形式データに対する分類・回帰問題を対象とした学習済み予測モデル
- ✓事前学習済みモデルであらゆるデータセットに共通する推論の仕組みを獲得しており，本検討で整理した学習データを入力する
- ✓本検討では複数の属性情報Xから個別のパラメータYを推定する



【表データのイメージ】

地点	属性情報				パラメータ
	X ₁	X ₂	...	X _n	Y
A	**	**	...	**	**
B	**	**	...	**	**
C	**	**	...	**	**
...	...	...	...	...	...
Z	**	**	...	**	??

異なる流域間の類似性を考慮

学習データ

パラメータ推定

特徴量間の関連性を考慮

3.TabPFNを用いたパラメータ推定

TabPFNによるパラメータ推定

Case	学習パラメータ	絞り込み条件	データ数	平均R ²
1	最適パラメータ	なし	355	0.20
2	上位5位まで	なし	1,755	0.41
3	上位10位まで	なし	3,510	0.46
4	上位20位まで	なし	7,020	0.45
5	上位30位まで	なし	10,530	0.47
6	上位5位まで	探索範囲上下限の パラメータを持つ データを除外	722	0.71
7	上位10位まで		1,486	0.60
8	上位20位まで		3,221	0.55
9	上位30位まで		5,016	0.54

【グラフの見方】

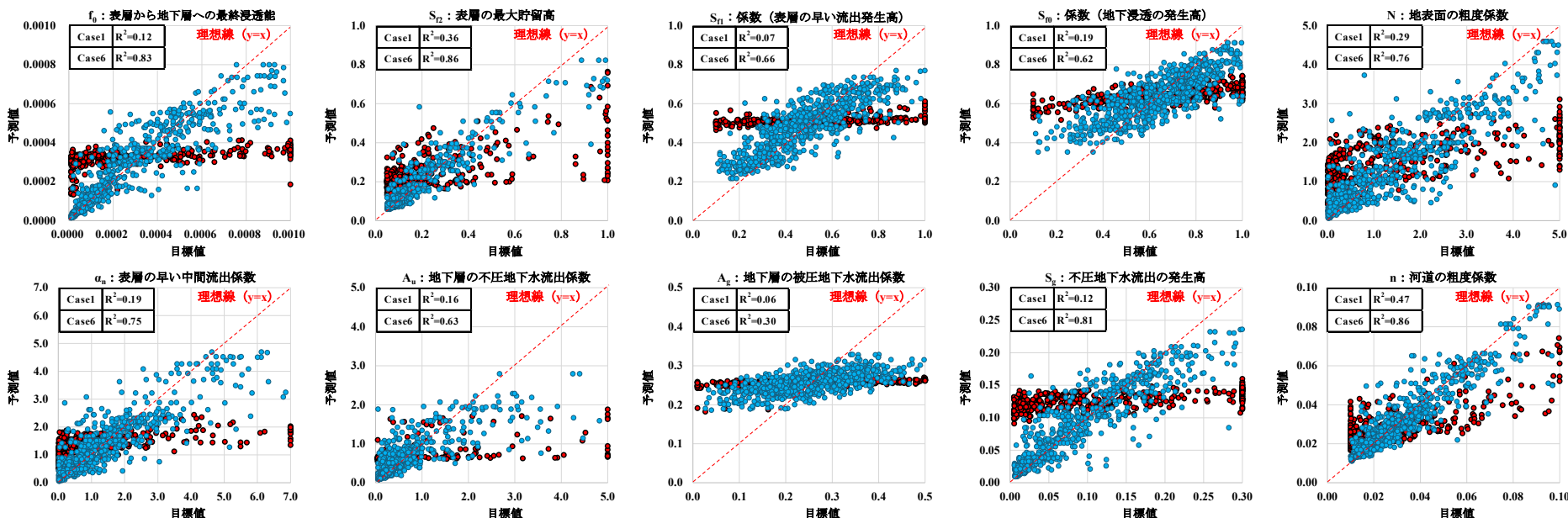
パラメータ予測値 y
(属性情報から推定)

決定係数

$$R^2 = 1 - \frac{\sum (x - y_{ave})^2}{\sum (x - x_{ave})^2}$$

予測値 y = 目標値 x
R² = 1.0

パラメータ目標値 x
(最適化済の値)



Case1: ● Case6: ●

Case1 (平均R²=0.20)
最適パラメータのみ

Case6 (平均R²=0.71)
上位5位パラメータ群

3.TabPFNを用いたパラメータ推定

TabPFNによるパラメータ推定

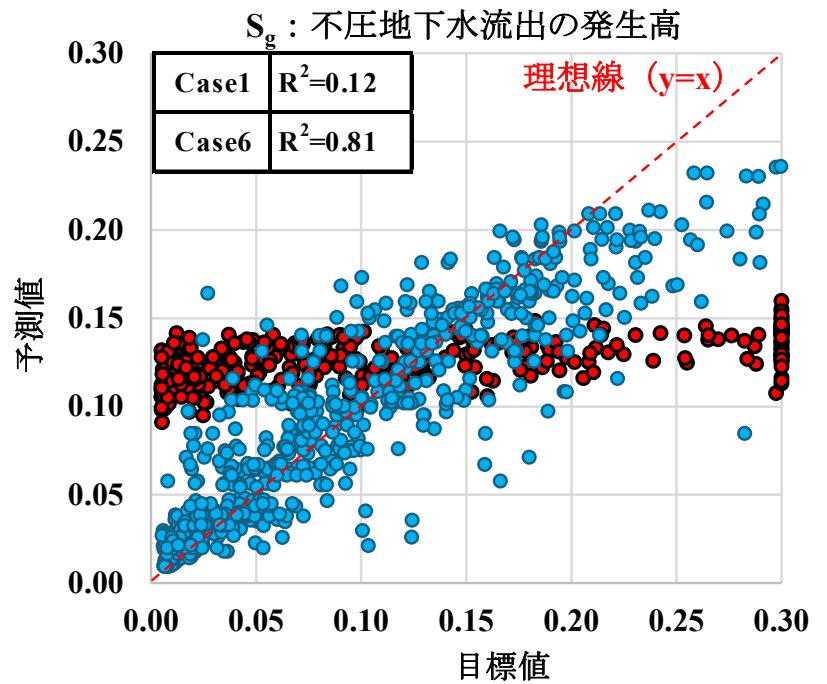
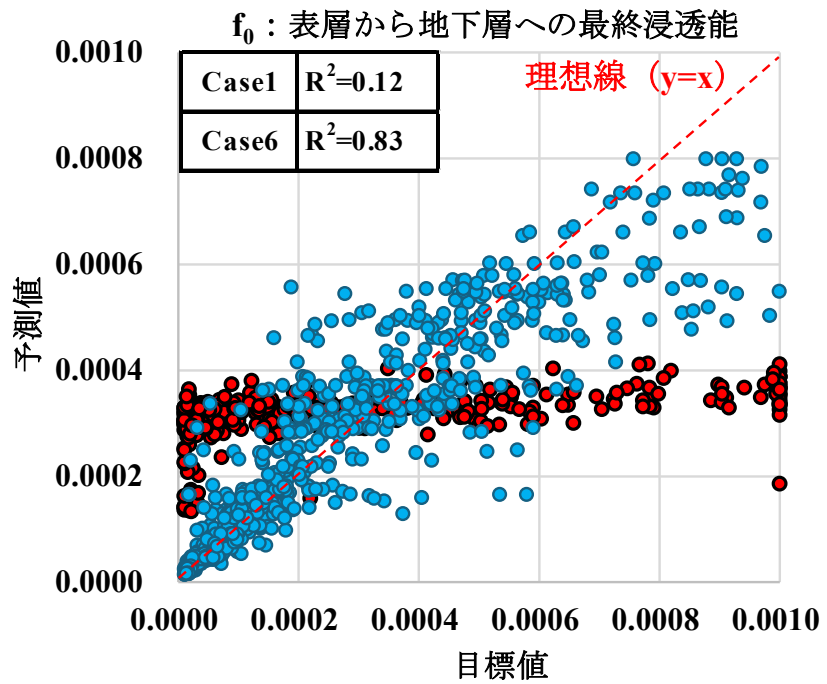
【グラフの見方】

パラメータ予測値 y
(属性情報から推定)

決定係数
 $R^2 = 1 - \frac{\sum (x - y_{ave})^2}{\sum (x - x_{ave})^2}$

予測値y=目標値x
R²=1.0

パラメータ目標値 x
(最適化済の値)



Case1 (平均R²=0.20)
最適パラメータのみ

→

Case6 (平均R²=0.71)
上位5位パラメータ群

Case1: ● Case6: ●

➡ CASE1の最適パラメータのみと比較して、CASE6のパラメータ群の多様な状態を取り入れることにより推定精度が向上がみられた

3.TabPFNを用いたパラメータ推定

キャリブレーション地点における精度検証

検証No	水系	地点	流域代表属性(カテゴリ値)						
			植生	土地利用	土壌	地下水涵養 貢献度	地質	地質年代	基底流量 貢献度
1	後志利別川	美利河ダム	落葉広葉樹	森林	褐色森林土壌Ⅱ	大	花崗岩質岩石	古生代	中
2	鳴瀬川	落合	落葉広葉樹	森林	黒ボク土壌	大	凝灰岩質岩石	新第三紀	中
3	荒川	浦山ダム	常緑針葉樹	森林	褐色森林土壌	大	砂岩	古生代	小
4	神通川	熊野橋	落葉広葉樹	森林	褐色森林土壌	大	各岩石の互層	新第三紀	中
5	雲出川	大仰	常緑針葉樹	森林	褐色森林土壌	大	花崗岩質岩石	中生代	中
6	大和川	番条	その他	建物用地	乾性褐色森林土壌	小	泥,砂,礫	中生代	中
7	佐波川	島地川ダム	常緑針葉樹	森林	褐色森林土壌	大	集塊岩及び凝灰角礫岩	中生代	中
8	渡川	秋田	常緑針葉樹	森林	褐色森林土壌	大	各岩石の互層	中生代	中
9	山国川	馬場	常緑針葉樹	森林	黒ボク土壌	大	安山岩質岩石	洪積世	大

検証No	水系	地点	流域代表属性(連続値)								
			低水路幅 (m)	堤間幅 (m)	集水面積 (km ²)	流路延長 (m/km ²)	谷密度 (本/km ²)	最大傾斜角 (平均)	最大傾斜角 (最大)	年平均降水量 (mm)	被覆度
1	後志利別川	美利河ダム	137.5	63.5	117.8	2,148	15.7	27.1	34.7	1,796	0.84
2	鳴瀬川	落合	76.5	60.4	199.8	2,752	16.9	15.1	23.3	1,548	0.76
3	荒川	浦山ダム	156.4	89.3	51.8	1,937	26.4	40.4	46.4	1,620	0.87
4	神通川	熊野橋	65.2	89.3	2,070.4	2,948	38.3	26.7	34.5	2,575	0.8
5	雲出川	大仰	58.7	64.2	304.5	3,347	25.9	27.6	35.7	1,948	0.81
6	大和川	番条	50.1	24.1	87.9	2,514	9.8	9.8	14.8	1,439	0.57
7	佐波川	島地川ダム	56.2	48.8	28.9	1,993	25.8	30.6	37.6	2,278	0.84
8	渡川	秋田	21.3	89.3	131.2	2,715	23.3	34.7	39.7	3,084	0.85
9	山国川	馬場	52.6	351.5	61.2	2,548	12.8	28.3	39.1	1,894	0.84



3.TabPFNを用いたパラメータ推定

キャリブレーション地点における精度検証

- ✓全国のキャリブレーション地点の中から，属性情報の異なる9地点を選定
- ✓複数の比較ケースを設定し， TabPFNによる推定パラメータの精度を確認
- ✓ (2)全学習条件の場合， (1) SCE-UAによって同定したパラメータと概ね同等の予測精度となる
- ✓ (3)検証地点を学習データに含めない場合， 個別に学習及びチューニングを行っていない(4)や(5)の 予測精度を上回る

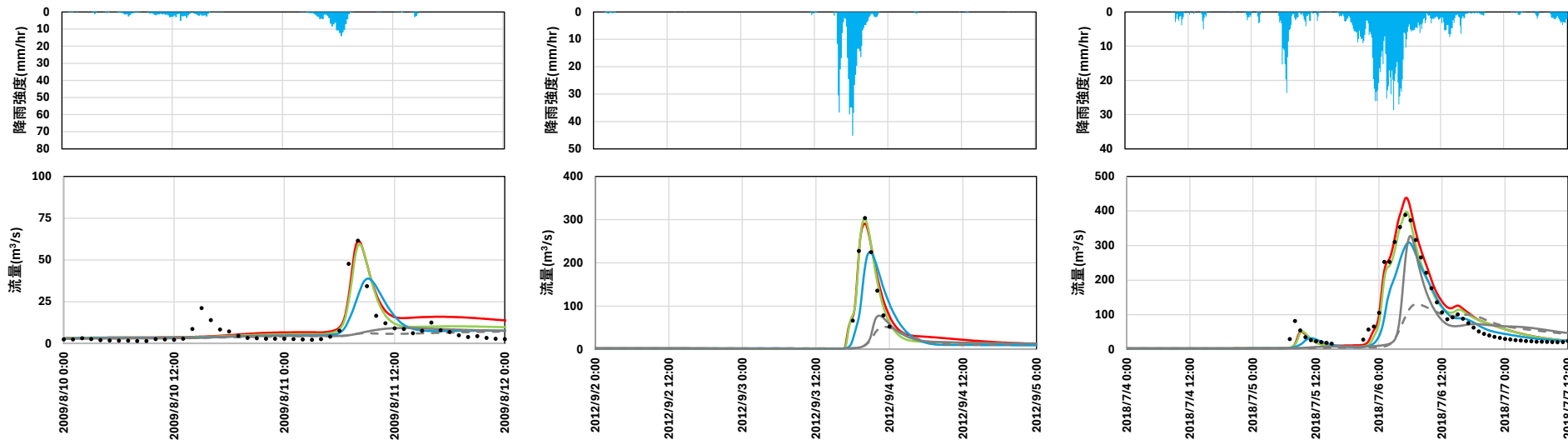
検証 No	RMSE(m ³ /s)					検証 No	ピーク流量絶対誤差率(%)				
	(1) SCE-UA	(2) TabPFN 全学習	(3) TabPFN 交差検証	(4) 既往検討 記載参考値	(5) 全SCE-UA 平均値		(1) SCE-UA	(2) TabPFN 全学習	(3) TabPFN 交差検証	(4) 既往検討 記載参考値	(5) 全SCE-UA 平均値
1	22.2	24.4	38.0	38.0	47.4	1	14.9	17.3	35.1	28.1	47.2
2	30.8	24.0	40.3	73.5	40.1	2	5.6	15.9	31.6	70.5	31.3
3	8.9	7.8	12.4	37.9	17.2	3	15.0	19.4	39.8	168.3	35.0
4	14.8	15.0	19.4	21.2	27.9	4	12.9	15.5	21.5	28.5	40.3
5	44.7	42.1	81.6	105.9	127.6	5	3.8	3.8	8.8	50.8	29.0
6	16.6	14.2	41.0	79.8	88.4	6	5.7	2.1	27.7	58.3	78.9
7	3.7	3.5	6.0	8.3	9.0	7	10.5	5.4	28.5	37.5	45.8
8	19.0	18.8	61.2	49.4	45.5	8	4.6	4.4	56.8	31.3	46.5
9	13.4	12.9	42.7	26.7	44.3	9	5.8	8.7	70.2	26.8	70.7
平均	19.3	18.1	38.1	49.0	49.7	平均	8.8	10.3	35.6	55.6	47.2

※TabPFNの検証は前頁の最適ケースであったCase6(上位5位まで学習、上下限值除外)をベースにした全学習と交差検証の2条件

3.TabPFNを用いたパラメータ推定

検証結果の例:大和川水系番条

凡例
●:実績流量 —:(1)SCE-UA —:(2)TabPFN全学習 —:(3)TabPFN交差検証 —:(4)既往検討記載参考値 ---:(5)全SCE-UA平均値



誤差評価指標	検証洪水	(1)SCE-UA	(2)TabPFN全学習	(3)TabPFN交差検証	(4)既往検討記載参考値	(5)全SCE-UA平均値
RMSE (m³/s)	2009/8/11洪水	4.69	4.01	4.74	5.89	6.00
	2012/9/4洪水	22.00	17.56	78.87	155.82	160.47
	2018/7/6洪水	23.23	21.08	39.32	77.74	98.68
	平均値	16.64	14.22	40.98	79.82	88.38

誤差評価指標	検証洪水	(1)SCE-UA	(2)TabPFN全学習	(3)TabPFN交差検証	(4)既往検討記載参考値	(5)全SCE-UA平均値
ピーク流量 絶対誤差 (%)	2009/8/11洪水	0.21	1.99	22.68	52.19	53.87
	2012/9/4洪水	12.59	2.33	78.32	225.65	251.57
	2018/7/6洪水	49.60	9.02	79.74	60.91	257.60
	平均値	20.80	4.45	60.25	112.92	187.68

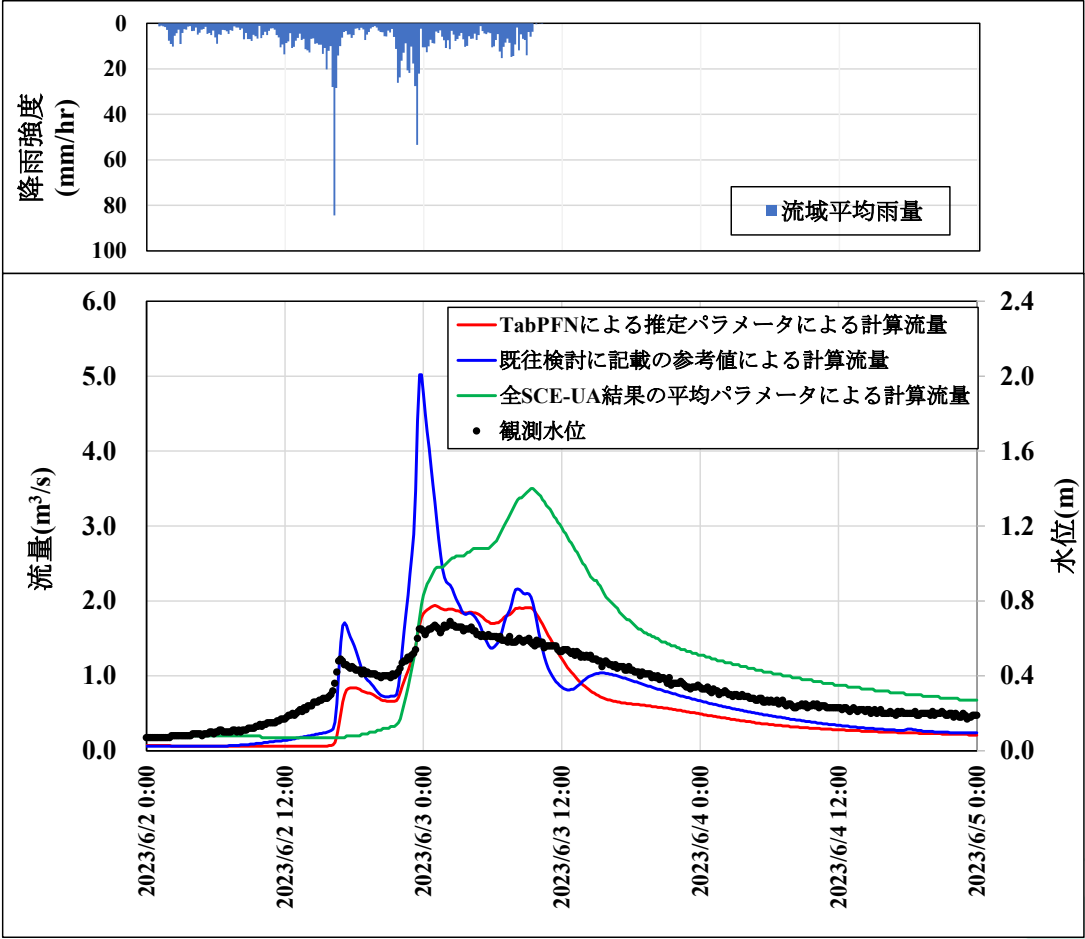
4.推定パラメータを用いた精度検証

- ✓公表されている観測データが存在しない中山間地において予測精度を検証
- ✓検証対象地点としたのは、多摩川水系A地点
- ✓比較ケースの中でTabPFNによる推定パラメータでの予測精度が最も高くなる結果となった

検証結果と属性情報(多摩川水系A地点)

植生	土地利用	土壌	地下水涵養 貢献
落葉広葉樹	森林	乾性褐色 森林土壌	小
地質	地質年代	基底流量 貢献度	低水路幅 (m)
砂岩	古第三紀	小	24.2
堤間幅 (m)	集水面積 (km ²)	流路延長 (m/km ²)	谷密度 (本/km ²)
—	1.5	960	35
最大傾斜角 (平均)	最大傾斜角 (最大)	年平均降水量 (mm)	被覆度
37	44.3	1,721	0.81

—:欠測



5.まとめ

- ✓パラメータ探索結果と流域代表属性を用いた特徴量分析により、既存の特徴量分析手法ではパラメータと流域代表属性の間に明確な関係性を見出すことは困難であった
- ✓パラメータ探索の結果と流域代表属性の整理結果を用いて、TabPFNによるパラメータ推定を実施した。その結果、学習済みの地点であれば個別に最適化したパラメータと同等の精度となることを示した
- ✓未学習の中山間地においても、既往の参考値等の結果よりも、流出波形の傾向を精度よく捉えられることを確認した

今後の課題

- ✓対象流域の拡張（全国109水系）による再現性及び広域適用性の向上
- ✓多様な流域への適用と適用限界の明確化
- ✓パラメータ調整によって得られるパラメータ群の活用方法の高度化
- ✓他手法との比較による推定手法の最適化
- ✓データ不足・不確実性を考慮したパラメータ推定の枠組み構築



未来につづく
安全・安心を

株式会社 建設技術研究所

【参考】キャリブレーション地点における精度検証

- ✓TabPFNによるパラメータ推定精度がCase6に次いで高かった2ケースでもキャリブレーション地点における精度検証を実施
- ✓検証結果はCase6と同様の傾向であり，比較ケースの中での相対的な予測精度の序列はパラメータ推定精度と概ね一致したがほぼ同等の精度

TabPFNの学習ケース

Case	学習パラメータ	絞り込み条件	データ数	平均R ²
1	最適パラメータ	なし	355	0.20
2	上位5位まで	なし	1,755	0.41
3	上位10位まで	なし	3,510	0.46
4	上位20位まで	なし	7,020	0.45
5	上位30位まで	なし	10,530	0.47
6	上位5位まで	探索範囲上下限のパラメータを持つデータを除外	722	0.71
7	上位10位まで		1,486	0.60
8	上位20位まで		3,221	0.55
9	上位30位まで		5,016	0.54

赤字:比較したケース

キャリブレーション地点における精度検証結果

Case	RMSE(m ³ /s)	
	TabPFN全学習	TabPFN交差検証
6	18.26	36.70
7	20.77	36.79
8	23.49	37.75

Case	ピーク流量絶対誤差率(%)	
	TabPFN全学習	TabPFN交差検証
6	10.6	32.6
7	12.0	35.2
8	13.6	33.9

※数値は全地点・全洪水の平均値