

【研究成果報告時/報告書（概要版）】様式-2-ウ

河川情報センター 研究助成 報告書

東京大学 准教授 澤田洋平

気候変動の進行に伴い「これまでに経験したことの無い水災害」に遭遇するリスクは近年高まっている。過去の水文データに当てはまるようにパラメータがキャリブレーションされた降雨-流出モデルでは未来の新しい気候レジームにおいて十分なパフォーマンスが維持できないことが危惧されている。本研究では44個の概念モデルを用いたマルチモデルアンサンブルにより降雨-流出過程モデリングにおける系統誤差を陽に考慮することで現在気候に対する過学習を防ぐ手法を新しく開発し、マルチモデルアンサンブルが気候や流況の変化に対して頑健であることを実証した。そのほかにも降雨-流出モデルの性能を測るための新しいベンチマークデータセットの開発・公開や、モデルパラメータの推定誤差を陽に考慮することによる流出解析の頑健性向上などの研究活動を行い、気候変動で大きく気候レジームが変化しても精度の低下しない「100年使える流出予測」の実現に貢献した。

Key Words : 降雨-流出解析、気候変動、キャリブレーション

1. 序論

気候変動の進行に伴い水災害の激化が進行し「これまでに経験したことの無い水災害」に遭遇するリスクは近年高まっている。洪水予測に用いられる降雨-流出モデルは過去の水文データに当てはまるようにパラメータチューニングされている。そのために各流域における降雨-流出モデルが気候変動による未来の新しい気候レジームに対しても現状のパフォーマンスを維持できるかどうかは疑問が残る。気候変動で大きく気候レジームが変化しても精度の低下しない「100年使える流出予測」の実現が求められる。

気候レジームが変化しても性能が劣化しにくい汎化性能を概念モデル型の降雨-流出モデルに持たせようとする試みは多数行われており、水文学におけるGrand Challengeの一つとなっている。代表的なアプローチはTrading space-for-time (TSFT) といわれるアプローチである。TSFTでは例えばある流域Aで現在気候における最適なパラメータをキャリブレーションで求めておき、将来気候と現在気候で大きく差が生じた場合、流域Aの将来気候に似た現在気候を持っていてかつ地質などの特質も似ている別の流域Bを探し、その流域Bでの最適なモデルパラメータを一部借用して将来気候における流出解析をするといった手法である^{1), 2)}。

気候レジームが変化すると性能が劣化する主たる原因は、現在気候でのパラメータチューニングにおいて過学習が生じているからである。過学習を避けるためには降雨-流出解析に内在するすべての不確実性を列挙し定量化する必要がある(図1)。過学習の原因がこのような不確実性、特にモデルの定

式化そのものに関連する系統誤差にあることは多くの研究で指摘されているが³⁾、TSFTなどの既存のアプローチにおいても系統誤差の存在を陽に考慮することはできていない。

本研究の目的は系統誤差をはじめとした流出解析の不確実性をマルチモデルアンサンブルによって精緻に考慮することで、気候変動に伴う流出解析の精度劣化を抑止できることを実証することにある。このような研究開発を通じて気候変動に対して頑健な「100年使える流出予測」の実現に貢献する。

本報告書(概要版)では、系統誤差に注目し44個の降雨-流出モデルを用いて上記の研究目的を達成し、学術誌に出版済みの研究⁴⁾について報告をする。2章で研究手法を述べ、3章で主たる結果をまとめる。4章で結論を述べる。5章では2-4章で報告したもの以外の2年間の研究活動についてごく簡単に概要を述べる。詳細については報告書本体で解説している。

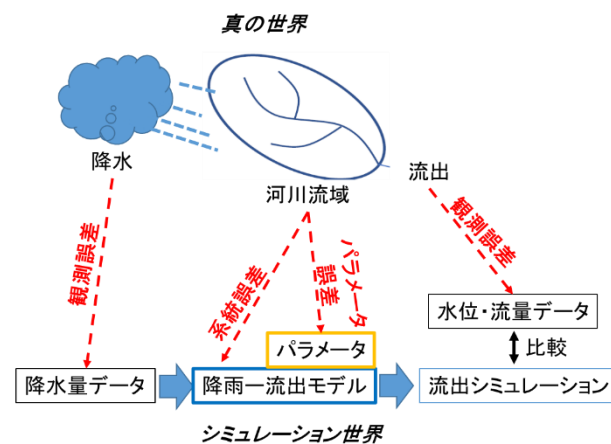


図1: 流出解析における不確実性の概要

2. 手法

降雨－流出モデルとして、オープンソースのライブラリであるModular Assessment of Rainfall-Runoff Models Toolbox (MARRMoT)⁵⁾を使用した。MARRMoTは既存の多様な降雨－流出モデルを分類し、それらを46種に集約することで、既存の降雨－流出モデルで扱われているスキームを網羅しようというものである。

MARRMoTは入力データとして降水量、可蒸発散量、地上気温を用い、流出を出力データとして計算する。これら3種の入力データについて米国での数値実験においてはCatchment Attributes and Meteorology for Large-sample Studies Dataset (CAMELS)⁶⁾から582流域(図2)のデータを抽出した。各流域の流量観測もCAMELSから取得した。

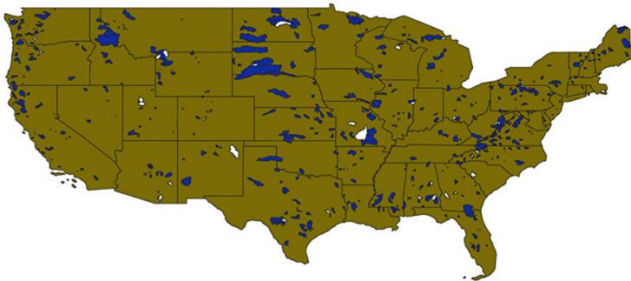


図2: 米国を対象とした数値実験における対象流域

MARRMoTの46個の降雨－流出モデルのうち、私たちの計算環境では安定してすべての流域を解くことができなかつた2つのモデル(SMARとPRMS)を除外した残りの44個のモデルを各流域においてキャリブレーションした。観測とシミュレーションの当てはまりの良さの指標としてはKling-Gupta Efficiency (KGE)⁷⁾を用いた。KGEを最大化するパラメータを探索するためにNelder-Mead シンプレックス法を用いた。

パラメータ最適化は1980年10月1日から1990年9月28日の3650日のデータを用いて行った。最適化したパラメータがこのキャリブレーション期間を超えて流量を再現できるかを確かめるために2013年1月から2014年12月31日までの730日間をバリデーション期間とした。キャリブレーション期間とバリデーション期間で10年以上の間を空けたのはこの間に流域においてある程度の気候・流況の変化を期待したからである。本章後半ではキャリブレーション期間とバリデーション期間の間に大きな流況の変化が見られた流域とそうでない流域とで分けて性能を解析することで、気候変動に対する頑健性が大量モデルを使うことで高まるかどうかを分析している。

3. 結果: マルチモデルアンサンブルの気候変動に対する頑健性

図3は44個のそれぞれのモデルの582流域でのKGEを箱ひげ図にまとめたものである。灰色で示した箱ひげ図の1つ1つがそれぞれのモデルの582流域でのKGEの分布を示している。44個のモデルがキャリブレーション期間においてもそれぞれに多様なKGE分布を示しており、十分にパラメータを最適化しても多様な系統誤差が存在することがうかがえる。パラメータ最適化に使用していない期間のデータでの検証では全体的にKGEが落ち込んでいることがわかる。これはよく知られた挙動であるが、KGEの落ち込み方にも44個のモデルで多様性があることがわかる。

44個のモデルによる流出予測を平均するマルチモデルアンサンブルは、この多様な系統誤差を平準化し、予測性能を向上させることが期待できる。図3の色付きの箱ひげ図が示すように、マルチモデルアンサンブルのKGEはキャリブレーション期間、バリデーション期間共に個々のモデルによるKGEを上回っている。一方で、マルチモデルアンサンブルに参入するモデルをうまく絞り込んで性能を向上させることは難しいこともわかる。図3上のキャリブレーション期間でKGEが大きくなるようにうまくモデルを選択しても、図3下のバリデーション期間ではKGEの上昇は見られない。このように系統誤差からモデルの低い汎化性能が生じるプロセスは非常に複雑であり、汎化性能が高まるようなマルチモデルアンサンブルを作ることは簡単ではない。いずれにしてもマルチモデルアンサンブルをとることで汎化性能は向上しており、大量モデルを使うことによって系統誤差の定量化を行うことは流出予測の頑健性を高めることが分かる。

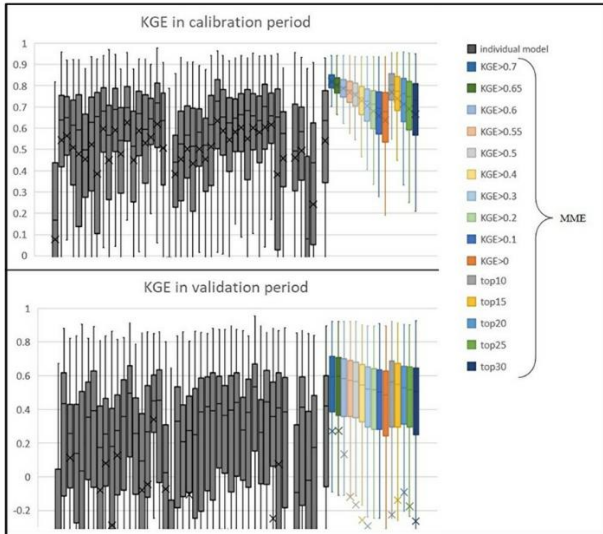


図3: (灰色の箱) 米国582流域におけるそれぞれのモデルのKGE。(色付きの箱) 米国582流域におけるマルチモデルアンサンブルのKGE。色の違いはマルチモデルアンサンブルに参入するモデルの基準 (例えばキャリブレーション期間においてKGE>0.7であったもののみを使ってマルチモデルアンサンブルを取る、あるいは上位10モデルのみをとる等)。上がキャリブレーション期間、下がバリデーション期間を表す。

本研究の主たる目的は気候変動により流域の気候レジームが大きく変化しても頑健な流出予測を実現することであった。上記で示したマルチモデルアンサンブルの汎化性能・頑健性は流域の気候や流況が大きく変化しても保持されているだろうか。私たちは流量の上位1パーセンタイルがキャリブレーション期間とバリデーション期間でどのくらい異なっているかを流域の流況変化とみなすことで582流域を疑似的に「気候変化の小さい流域」「気候変化が若干ある流域」「気候変化の大きい流域」の三類型に分類して、それぞれの類型におけるバリデーション期間のKGEを解析した。図4に結果を示す。既往研究から予測されるように気候変化が大きいほど、流況変化が大きいほど流出解析の性能は大きく低下する。図4に顕著なのは気候変化が大きい流域ほど、赤で示したマルチモデルアンサンブルのKGEの中央値と、各モデルのKGEの中央値の差は大きくなっていくという点である。これは気候変化が大きいほどマルチモデルアンサンブルが予測性能の低下を防ぐ効果が顕著になることを示している。マルチモデルアンサンブルによる系統誤差の考慮が流出解析の気候変動に対する頑健性を高めることを示唆する結果となっている。

図5ではブートストラップ法により、これまで議論してきたマルチモデルアンサンブルによるKGEの上昇が、より少ないモデル数、より少ない流域数でも安定して得られるかを分析した補助的な解析の結果である。詳細については出版済みの論文に譲るが、図5からは流域数やモデル数が少ないと、安定し

た結果は得られないことを示しており、本研究のような結論に至るためには十分な大きさのデータが必要であることがわかる。

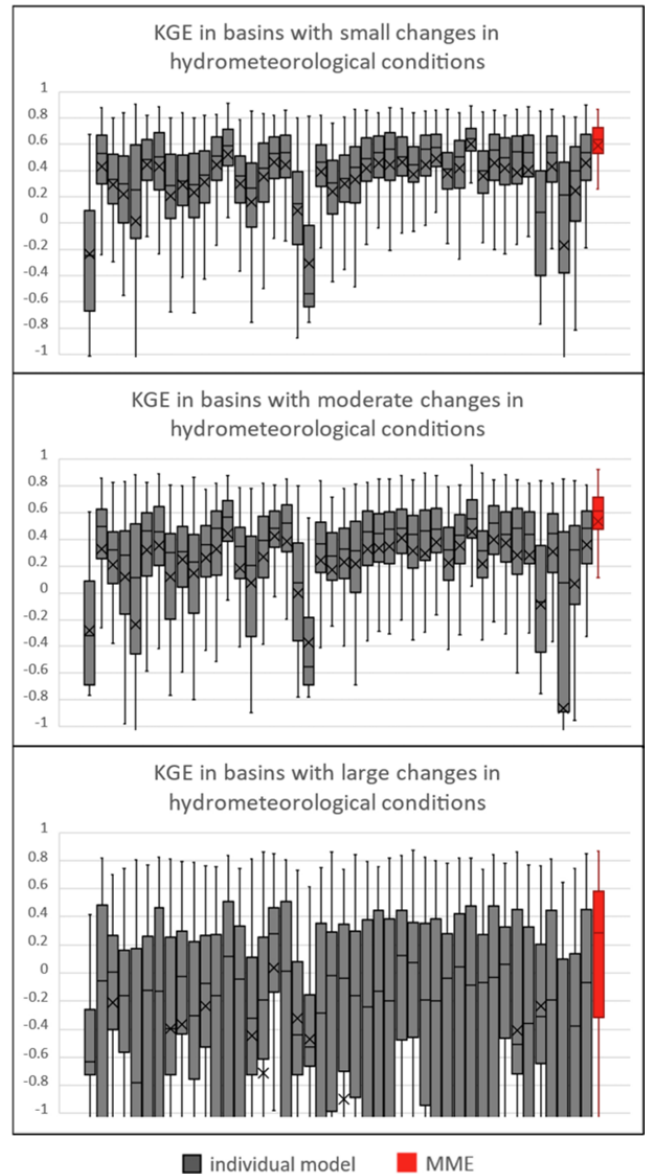


図4: 図3と同じくバリデーション期間のKGEを示しているが、582流域を上位1パーセンタイルの流量を利用して3つに分類した。(上) 上位1パーセンタイルの流量変化がキャリブレーション期間の流量の20%未満で「気候変化が小さい」流域。(中) 変化が20-40%で「気候変化が若干ある」流域。(下) 変化が40%以上で「気候変化が大きい」流域。灰色が各モデル、赤がマルチモデルアンサンブル。

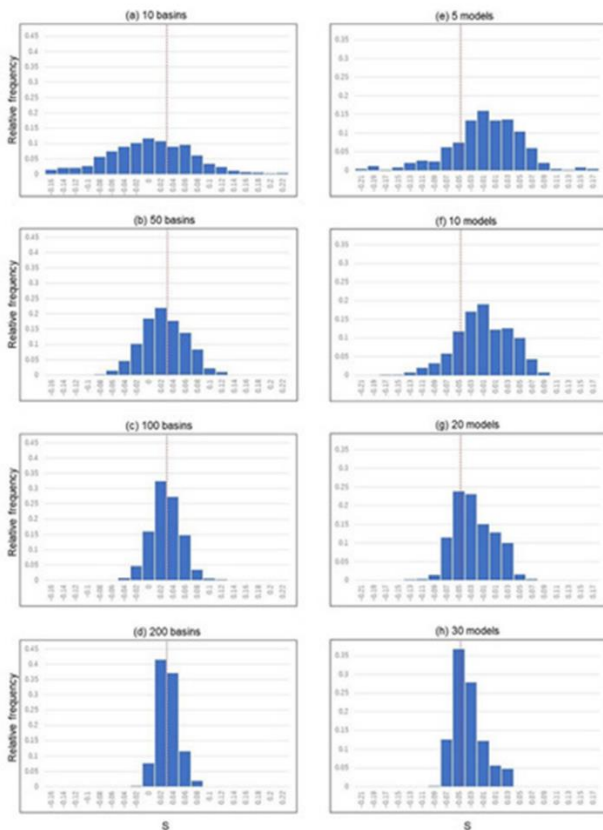


図5: 流域数とモデル数をブートストラップすることによって得られた頻度分布。横軸は各モデルとマルチモデルアンサンブルのKGEの差を指標化したもので値が大きいほどマルチモデルアンサンブルが各モデルよりも卓越していることを示す（詳しくはKimizuka and Sawada 2022）。(a-d) 流域数を10, 50, 100, 200流域として全流域からランダムに選択して解析を1000回繰り返した結果。(e-h) モデル数を5, 10, 20, 30として全モデルからランダムに選択して解析を1000回繰り返した結果。赤線はもともとの全流域、全モデルを使用した結果。

4. 結論

本研究では降雨一流出シミュレーションの主要な不確実性である系統誤差を考慮するために、44個もの概念モデルを統合したマルチモデルアンサンブル手法を提案した。マルチモデルアンサンブルは現在気候に対するモデルパラメータの過学習を抑制し、洪水流出解析の頑健性を高めることを十分な量のデータで実証した。特に流域の気候・流況が大きく変化したときにマルチモデルアンサンブルの有効性は際立つ。このことから本研究の知見が気候変動に対して頑健な洪水流出解析に貢献できることが分かった。

5. その他の研究活動

2-4章で述べたアプローチを日本域の135流域においても適用した。降雨一流出モデルの開発者が自らのモデルの性能を客観的に測るベンチマークとして使用することを企図して、44モデルによる135流域の流出解析結果をその検証結果と共

にオープンソースのデータとして公表した⁸⁾。

本研究のさらなる発展形として、パラメータ推定においても不確実性を考慮することのできるソフトウェア開発を行った。モデルの定式化自体に含まれる系統誤差のみならず、パラメータ最適化の過程で生じるパラメータ誤差も考慮することで、現在気候への過学習をさらに有効に避けることができるかを検証していくことが今後の研究課題である。

謝辞：本研究は河川情報センター研究助成の支援を受けた。

参考文献

- 1) Vora, A., & Singh, R. (2022). Improving rainfall-runoff model reliability under nonstationarity of model parameters: A hypothesis testing based framework. *Water Resources Research*, 58, e2022WR032273. <https://doi.org/10.1029/2022WR032273>
- 2) Singh, R., Wagener, T., Van Werkhoven, K., Mann, M. E., & Crane, R. (2011). A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate-accounting for changing watershed behavior. *Hydrology and Earth System Sciences*, 15(11), 3591-3603. <https://doi.org/10.5194/hess-15-3591-2011>
- 3) Duethmann, D., Blöschl, G., & Parajka, J. (2020). Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change? *Hydrology and Earth System Sciences*, 24(7), 3493-3511. <https://doi.org/10.5194/hess-24-3493-2020>
- 4) Kimizuka, T. & Y. Sawada (2022), How robust is a multi-model ensemble mean of conceptual hydrological models to climate change?, *Water*, 14, 2852, <https://doi.org/10.3390/w14182852>
- 5) Knoben, W. J. M., Freer, J. E., Fowler, K. J. A., Peel, M. C., & Woods, R. A. (2019a). Modular Assessment of Rainfall-Runoff Models Toolbox (MARRMoT) v1.2: An open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations. *Geoscientific Model Development*, 12(6), 2463-2480. <https://doi.org/10.5194/gmd-12-2463-2019>
- 6) Addor, N., Newman, A., Mizukami, M. & Clark, M.P. (2017). Catchment Attributes for Large-Sample Studies; UCAR/NCAR: Boulder, CO,

USA

7) Gupta, H.V., Kling, H., Yilmaz, K.K., & Martinez, G.F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377, 80-91.

<https://doi.org/10.1016/j.jhydrol.2009.08.003>

8) Sawada, Y., S. Okugawa & T. Kimizuka (2022), Multi-model ensemble benchmark data for hydrological modeling in Japanese river basins, *Hydrological Research Letters*, 16, 73-79, <https://doi.org/10.3178/hrl.16.73>