Advances in Water Resources 192 (2024) 104781

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Combining global precipitation data and machine learning to predict flood peaks in ungauged areas with similar climate

Zimeena Rasheed [a], Akshay Aravamudan [b], Xi Zhang [b], Georgios C. Anagnostopoulos [b], Efthymios I. Nikolopoulos [a,*]

[a] Civil and Environmental Engineering, Rutgers University, Piscataway, NJ, United States
[b] Computer Engineering and Sciences Department, Florida Institute of Technology, Melbourne, FL, United States

ARTICLE INFO

Keywords:
Global precipitation
Machine learning
Flood prediction
Ungauged basins

ABSTRACT

Increasing flood risk due to urbanization and climate change poses a significant challenge to societies at global scale. Hydrologic information that is required for understanding flood processes and for developing effective warning procedures is currently lacking in most parts of the world. Procedures that can combine global climate dataset from satellite and reanalysis with fast and low computational cost prediction systems, are attractive solutions for addressing flood predictions in ungauged areas. This work develops and tests a prediction framework that relies on two fundamental components. First, meteorological data from global datasets (IMERG and ERA5-Land) provide key input variables and second, ML models trained in the data-rich contiguous US, are applied in climatically similar regions in other parts of the world. Catchments in Australia, Brazil, Chile, Switzerland, and Great Britain were used as pseudo-ungauged regions for testing. Results indicate acceptable performance for both IMERG and ERA5-Land forced models with relative difference in flood peak prediction within 30 % and similar overall performance to locally trained ML models. Specific climate regions for which ML models have revealed good performance include Mediterranean climates like the US West Coast, subtropical areas like the Southern Atlantic Gulf, and mild temperate regions like the Mid-Atlantic Basin. This work highlights the potential of combining global precipitation dataset with pre-trained ML models in data-rich areas, for flood prediction in ungauged areas with similar climate.

1. Introduction

Rising riverine flood risk, due to increased exposure of populations (Tellmann et al., 2021; Andreadis et al., 2022) and precipitation extremes (Papalexiou and Montanari, 2019), poses a major socioeconomic threat at global scale. Climate projections highlight an increase in the frequency and intensity of climate extremes (Emmanouil et al., 2022; Jong et al., 2023; IPCC 2021; Wang et al., 2022; Winsemius et al., 2016) that will exacerbate flood risk in the future. Assessing flood risk, developing adaptation strategies, and leading communities to become more climate resilient are all extremely important tasks on the global climate task force agenda (WMO, 2022). Unfortunately, the communities most vulnerable to flood hazards, are the ones who struggle with a lack of data and an appropriate warning system that impedes even the first step of simply knowing the potential magnitude of a flood event they should expect and prepare for.

Flood peak prediction, or more generically, stream flow prediction in ungauged basins has been a major hydrological issue the community has grappled with for decades. While not an unprecedented issue, it is one however that has become more urgent to overcome as flood changes in magnitude, frequency, and timing in various locations around the world have been observed (Bertola et al., 2023; Tarouilly et al., 2021; Huang et al., 2022). The general lack of gauged data for flood prone regions globally has been established (Bloschl et al., 2013; Bertola et al., 2023). By the same token, there are locations around the globe with a long enough historical record of hydrologic data that allows for relevant research and strategic decision-making. These said locations encapsulate a wide range of hydroclimatic regions; one such location is the contiguous United States (CONUS).

Hydrologic models of varying complexity have traditionally been applied in these data-rich regions; one example being the US National Water Model (Cosgrove et al., 2024). Such models, while advanced in the ability to simulate observed and forecast streamflow along with related output, the operation of these models require vast computational

* Corresponding author.
E-mail address: efthymios.nikolopoulos@rutgers.edu (E.I. Nikolopoulos).

https://doi.org/10.1016/j.advwatres.2024.104781
Received 12 March 2024; Received in revised form 8 July 2024; Accepted 30 July 2024
Available online 3 August 2024

resources, especially at large basin scales (Solomatine et al., 2011). The operational configuration and needs of these complex models subsequently challenge transference to data-scarce regions. Strides have been made to optimize approaches for hydrologic models (Ming et al., 2020; see Zhang et al., 2014 on limitations of optimization resources), but, steadily growing as a competitive alternative to traditional hydrologic models, are the application of machine learning (ML) models or data-driven approaches that use ML techniques (Hu et al., 2019; Kratzert et al., 2019; Ma et al., 2021; Mosavi et al., 2018; Rasheed et al., 2022; Sanjay Potdar et al., 2021). *Recent advances in global precipitation datasets together with state-of-the-art flood prediction procedures relying on ML approaches, open new horizons for flood prediction at global scale.*

Studies relying solely on ML models have successfully been used for predicting daily streamflow timeseries in pseudo-ungauged basins located in the model-trained regions including United States (Kratzert et al., 2019; Sanjay Potdar et al., 2021; Rasheed et al., 2022), South Korea (Choi et al., 2022) and Brazil (Nogueira Filho et al., 2022). A data-driven sparse sampling technique (Zhang et al., 2023) that incorporates remotely-sensed data to supplement ground observations in the feature space was used to reconstruct daily streamflow timeseries in poorly-gauged locations across the contiguous U.S. *In the above cases, the trained model is typically trained and validated over the same domain or contains data related to the tested region.* Nevertheless, the ML models, if trained with adequate data (which include features that describe basin physiography, climate and precipitation), learn about general regional attributes and are able to predict in ungauged locations based on hydrological similarity. Ma et al. (2021) has used an LSTM pre-trained over data-rich CONUS and with appropriate transfer learning (TL) techniques, used the model to learn about and predict over pseudo data-rich and data-scarce regions in Great Britain, Chile and China using local meteorological and basin attribute data published for each region.

However, to the best of our knowledge, ML models trained over specific climatic regions with *global precipitation datasets* have never been applied to data-scarce regions outside of the training dataset simply based on climatic similarity. This study therefore serves as a vital assessment of the applicability of climate-specific, trained ML models towards flood peak prediction at the catchment scale and across global extents (i.e. a diversity of hydroclimatic zones) given the use of global precipitation datasets. Towards that end, in this work, we investigate (a) the performance of ML flood prediction models integrated with global hydrometeorological dataset and (b) the applicability of ML models trained in data-rich regions for flood prediction in ungauged regions. Precipitation is arguably the major driver for flood hazards and thus the performance of two global precipitation datasets (ERA5-Land and IMERG-Early Run) is investigated. By considering precipitation products at global scales, the result of our study will inform on if and where it is applicable to integrate such products with ML models for flood predictions in ungauged regions.

The remainder of the paper is organized as follows: Section 2 describes the study domains and data; Section 3 follows with the framework for training and testing ML models to predict peak flows in ungauged regions; Section 4 provides the results and discusses the same with respect to shared regional, hydroclimatic characteristics and Section 5 closes with a recapitulation of the main takeaways.

2. Study area and data

The study domains encompass catchments, ranging from 2 to 36,400 square kilometers in area, from six regions across the globe: the contiguous United States (CONUS), Australia (AUS), Brazil (BR), Chile (CL), Great Britain (UK) and Switzerland (CH). Fig. 1 and Table 1 identify the study domains, also noting the number of catchments and climate classes, per study domain. Climate classes were based on the Köppen Classification (Chen and Chen, 2013).

All data in this study may be separated according to two main sources: the Catchment Attributes and MEteorology for Large-sample Studies (CAMELS) datasets and global precipitation datasets.

2.1. The [Local] CAMELS datasets (L-CAM)

The domains of interest were selected given the availability of the CAMELS datasets which contain for each catchment: (1) time series records of locally observed (and simulated in some cases) meteorological and streamflow data, and (2) key climatic, hydrologic, soil,

Table 1
Study domains, number of catchments and climatic regions and approximate range of catchment area.

Domain [Abbrev]		Number of Catchments	Number of Climatic Regions	Area (km^2)
Australia	[AUS]	207	3	7–8650
Brazil	[BR]	648	2	11–9850
Switzerland	[CH]	325	4	2–36,400
Chile	[CL]	406	3	18–9920
United States	[CONUS]	463	10	6–14,270
Great Britain	[UK]	659	3	2–9930

Fig. 1. Study domains [10 regions represented for contiguous US (CONUS); 2 regions represented for Brazil (BR); 3 regions represented for Chile (CL); 4 regions represented for Switzerland (CH); 3 regions represented for Australia (AUS) and 3 regions represented for Great Britain (UK)].

vegetative and topographic variables (useful for supplying lumped, static attributes per catchment to models). Table 2 lists the CAMELS datasets used. The details of each compiled CAMELS dataset may be further learned from the references included in Table 2.

2.2. Global precipitation and temperature datasets

Two dynamic, meteorological variables were important for this study: precipitation and temperature. To facilitate an investigation on the sensitivity of results to precipitation input, we selected two globally available and widely-used precipitation products. The first is the IMERG Early Run V6B (Huffman et al., 2019), hereafter IMERG-ER, which is a satellite-based product with low latency (~4h). Our choice to consider the version of the "early run" instead of the "late" and "final" version of IMERG was precisely because of the low latency of this product which presents an opportunity to be used for near real-time flood predictions. The second product is the ERA5-Land Reanalysis (Muñoz-Sabater et al., 2021), hereafter ERA5-L. For temperature we used only as single input the ERA5-L 2m temperature. Both IMERG-ER and ERA5-L are gridded datasets available at 0.1deg spatial and hourly temporal scales.

To be consistent with the resolution of the local reference datasets (CAMELS), the IMERG-ER and ERA5-L datasets underwent two key preprocessing steps. First, the temporal resolution was aggregated to the daily scale and second, all grids were clipped to the boundary of each catchment in each region and thereafter a catchment-averaged value was calculated per timestep. There was an intermediate step which resampled (using the bilinear approach) the original precipitation grids to a finer scale to (a) address smaller-sized catchments, and (b) better clip to the irregular shapes of the basins. The temporal extent of the meteorological datasets considered starts on June 1, 2000, given that this is the earliest timestep available as part of the IMERG-ER product. The end date varies by region based on the end date of the respective local datasets, with the latest being December 31, 2018. Table 2 lists and summarizes the temporal and spatial resolution of the local and global datasets used.

2.3. The flood peak database (FPD)

The key responding variable and therefore the predictand of this study is flood peak magnitude. Following the work of Rasheed et al. (2022), our flood prediction framework focuses on predicting flood peaks, instead of the entire streamflow hydrograph, which we consider a very important first step for developing early warning procedures. Flood peak identification for all catchments and domains are those events whose magnitudes are greater than or equal to the 90th quantile observed streamflow magnitude for each catchment. This thresholding approach permits focus on the prediction of truly high flowing events. Further selection of flood peaks from the streamflow time series was needed to ensure independence of the flood events. This was done following the approach established by Hu et al. (2020) where flood peaks separated by a minimum distance in time satisfies the condition for independence of the flood events identified in a catchment. This parameter, time separation (θ) is a function of catchment area, A (km^2):

$$\theta > 5 \text{ days} + 2.59 \times \log(A) \quad (1)$$

Dynamic Characteristics of Triggering Storms: The attribution of the storm that triggered each identified flood peak follows Rasheed et al. (2022), (see specifically Section 2.1 for details). In summary, a triggering storm event for each flood peak is the closest storm in time that (1) has a storm start date earlier than the start time of the rising limb of the flood peak's hydrograph, (2) is greater than 1 mm in precipitation magnitude, and (3) has an end time that is no longer than the time of the observed flood peak (any precipitation after the flood peak is ignored if any longer). For each identified triggering storm event, a mean and maximum triggering precipitation value is computed as two dynamic characteristics describing the corresponding flood peak. Once the start and end time of the triggering storm event is identified, an average of the maximum temperature series for this duration is calculated forming another dynamic descriptor for the flood peak called mean of maximum temperature. Additionally, an Antecedent Precipitation Index (API) for each flood peak is calculated as another dynamic variable for representing the antecedent wetness condition of the soil in a catchment just before the start of the triggering storm. The API (Kohler and Linsley, 1951; see also "retained rainfall" model by Singh, 1988) is defined in Eq. (2) and was constructed across the 30-day period prior to the start of the triggering storm (i.e. 30 antecedent days were considered).

$$API = \sum_{j=0}^{i} P_{t-j} k^j \quad (2)$$

where i = total number of antecedent days; j = lag or antecedent time of interest (days), P_t the precipitation recorded on day t and k = decay constant which ranges from 0.8 to 0.98 (Viessman and Lewis, 1995) with 0.9 used as the estimate for this study.

Constructing the FPDs: For the catchments in each domain, all flood peaks and corresponding triggering storm characteristics were compiled (to form a FPD) following the above summarized pre-processing steps for the three precipitation sources: L-CAM, IMERG-ER and for ERA5-L. Therefore, for any given domain, three FPDs were constructed (for example for CONUS there was a L-CAM FPD, an IMERG-ER FPD and an ERA5-L FPD). In each FPD, the flood peak magnitude has been normalized by the area of the catchment with resulting units of mm/day, i.e. specific discharge.

Table 2
Datasets employed noting resolution and sources.

Variables	Dataset	Resolution	Refs.
Streamflow; static catchment attributes	CAMELS-CONUS; CAMELS-BR; CAMELS-AUS; CAMELS-UK; CAMELS-CL; CAMELS-CH	Spatial: Catchment-averaged; Temporal: daily; 2000 – 2018	Addor et al. (2017); Newman et al. (2015); Chagas et al. (2020); Fowler et al. (2021); Coxon et al. (2020); Alvarez-Garreton (2018); Höge et al. (2023)
Precipitation	IMERG Early Run v6	0.1 deg 2000 – 2018, daily	Huffman et al. (2019)
	ERA5Land	0.1 deg 2000 – 2018, daily	Muñoz-Sabater (2021)
Temperature	ERA5Land	0.1 deg 2000 – 2018, daily	Muñoz-Sabater (2021)

3. Methodology

3.1. General framework

Flood peak prediction using ML models and based on hydrometeorological data and catchment attributes have demonstrated good performance when trained and tested over catchments within a basin sharing hydroclimatic features (in essence regional models; Rasheed et al., 2022; Kratzert et al., 2019; Razavi and Coulibaly, 2013). Fig. 2 illustrates the general methodology for flood peak prediction in ungauged catchments using global hydrometeorological input. This work connects the use of machine-learning for flood peak prediction in ungauged basins with global precipitation datasets. Models are trained in a data-rich region, CONUS, and applied to pseudo-designated ungauged regions across the globe, namely in Australia, Brazil, Chile, Switzerland, and Great Britain.

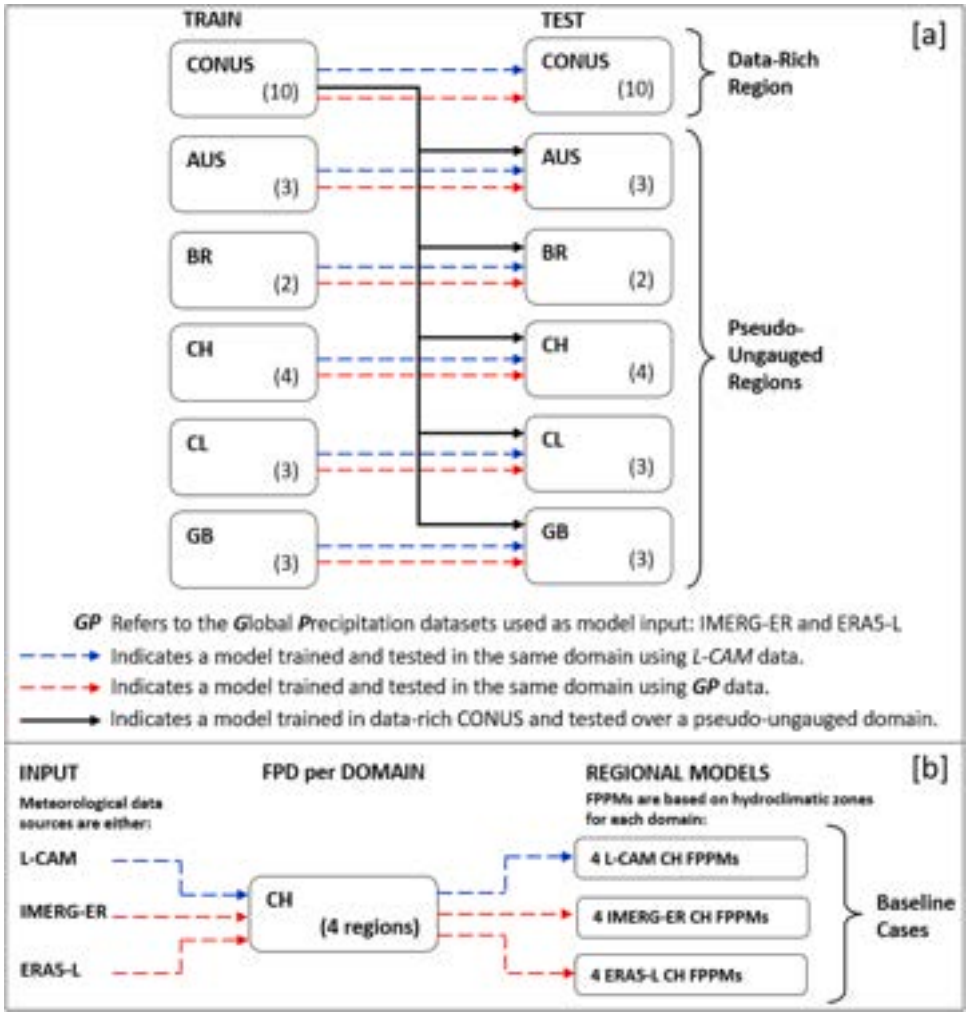

Fig. 2. Methodology for flood peak prediction in ungauged catchments using global precipitation (GP) input; [a] Schematic of the train-test scenarios for both the data-rich CONUS domain and pseudo-ungauged domains. The number of unique regions for which FPPMs are developed per precipitation dataset is shown in parentheses of the box for each study domain. The bold black line indicates "applied only scenarios" from CONUS to any other study domain. The dashed blue line is the baseline case for Experiment #1. The dashed red line represents the baseline cases for Experiments #2 and #3 that use global precipitation input. [b] Resultant baseline cases for the three experiments for example domain, Switzerland.

3.2. Input variables and hydroclimatic zoning

Enhancement of the flood peak prediction models (to be deliberated in Section 3.3) shows a list of seven predictors. These are the: (1) maximum precipitation, (2) mean precipitation, (3) mean of the maximum daily temperatures - for the triggering storm period, (4) antecedent precipitation index over a 30-day period before the start of the triggering storm period, (5) fraction of forest cover as a land cover predictor, (6) mean annual potential evapotranspiration and (7) aridity. The last three variables are static attributes provided as a catchment-averaged value from respective L-CAM datasets. Expectedly, the first four variables are dynamic predictors that will change not only per flood peak identified in each catchment, but also, for each hydrometeorological dataset sourced (IMERG-ER, ERA5-L and L-CAM). Consequently, the relative importance of each predictor varies given the latitude and longitudinal spread of the domains under study; in which case there are likely (static) predictors not included in our models that may be more important – or by the same token – a predictor to be excluded (for more on predictor importance of flows see Massari et al., 2023 and Rasheed et al., 2022). However, notwithstanding the sound rationale for these variables from a hydrologic perspective, we strived to maintain simplicity in data dimensionality in our models. Additional intentions of doing so were to (1) maintain easy reproducibility of model structure across domains and (2) ensure that these static variables (just as their dynamic counterparts) could be sourced from remotely-sensed datasets if one's objective is to apply to a "non-CAMELS" documented domain (in other words a true external or ungauged region).

Within each study domain (CONUS, CH, BR, CL, BR, AUS, GB) displayed in Fig. 1, we separated the catchments into regions distinguished as larger basins, each having unique climatology and hydrologic characteristics. In most cases, this is already provided as an attribute in the CAMELS datasets. Specifically, in CONUS, separation is based on the HUC-02 level water regions stipulated by the United States Geological Service. For the AUS, BR, CH and CL domains, the respective CAMELS datasets similarly provided as a feature the drainage basin name or monitored water region to which each catchment belonged. For the GB catchments, we took a hydroclimate approach and grouped the catchments according to localized terrain and precipitation regimes. Fig. 4 may also be referenced for a spatial attribution of the unique regions in each domain used for this study. The ML models were then developed per unique climatic region. Table 1 and Fig. 2 outline the resultant number of unique regions in each study domain.

3.3. ML flood peak prediction model (FPPM) experiments

Building upon the previous work of Rasheed et al. (2022), we carry

out a series of experiments to investigate the potential applicability of ML models, trained within CONUS, for flood prediction in other climatologically similar regions of the world.

While similar performance between the two datasets is not anticipated for flood peak prediction from one hydroclimate to another, this investigation highlighted regions across latitudes where performance was best for each product and conversely, where each product had the most difficulty being applied. The training, validation (for hyperparameter tuning) and (withheld) test ratios for dataset splitting are 50 %, 30 % and 20 %, respectively.

3.3.1. The HGBR model

Histogram Gradient Boosting Regressor (HGBR; Friedman, 2001; Ke et al., 2017) is a type of gradient boosting learning algorithm that learns incrementally as it improves on the previous weak learner. It continuously improves since each subsequent model's prediction is weighted so that the ensemble's prediction is rendered more accurate. The addition of models to the ensemble will have predictions that are maximally correlated to the negative gradient of the ensemble's overall loss function. Though there is great difficulty with rule extraction from this type of model (see Cormen et al., 2001), these ensemble-based learning models allow for some sense of interpretability through permutation feature importance estimates. These estimates contributed to our interpretation of the trained models insofar as understanding the interaction of peak flow event drivers and the relative significance of the said drivers across various hydroclimatic regions. From flood alert systems for forecasting flood stages (Jarajapu et al., 2022) or streamflow (Ni et al., 2020) to flood susceptibility assessments (Mirzaei et al., 2021) and design flood estimates (Xu et al., 2023), Extreme Gradient Boosting models have been proven useful in the hydrological field. Moreover, a hybrid approach that uses a Light Gradient Boosting Model in tandem with a traditional hydrologic model to detect urban floods (Sanders et al., 2022) serves as a good example for alleviating the computational strain of using only traditional hydrologic models.

The loss function selected was the Mean Square Error (MSE); an applicable objective function to compare the performance of models in each region given that all input variables were continuous. The hyperparameters of the HGBR model were tuned using Optuna, a python library for hyper-parameter optimization. More specifically, we used the Tree-structured Parzen Estimator (Bergstra et al., 2011) which is a Bayesian optimization method that explores the search space via a tree structured approach. We used the HGBR implementation provided by scikit-learn (Pedregosa et al., 2011), a Python library for machine learning tasks. Hyperparameters were tuned for each regional model and the standard method for the splitting criteria which opts for the split that minimizes the residual sum of squares was used. The hyperparameters tuned included the number of estimators (ranged from 10 to 300), the maximum number of leaves per learner (ranged from 3 to 60), the $\ell 2$ regularization parameter (ranged from 0 to 2000) for the learned weights, and the learning rate (ranged from $1e^{-5}$ to $5e^{-1}$).

3.3.2. Trained models – "Data-Rich" region, CONUS and pseudo-ungauged regions

CONUS is considered our reference or data-rich domain, for which the ML models are trained. There are 10 climatic regions and therefore 10 trained models for each of the three given precipitation data input (see Fig. 2). The contiguous US boasts, in fact, 18 unique hydroclimatic zones (HUC-02; Kratzert et al., 2019; Rasheed et al., 2022) albeit for this paper the results of only 10 are presented. ML models were established for all 18 regions but with the goal of maintaining model simplicity in both structure and predictors (flood peak drivers), the performance in 8 regions did not measure up to our range of acceptable results when using the global precipitation data as input (IMERG-ER and ERA5-L). Notably, these 8 regions constitute the central US. While it is difficult to pinpoint exactly the weight or reasoning for the poor performance in these 8 regions, a few common observations follow. First, these poor-performing regions were mostly water-limited, specifically having arid-steppe, [cold] semi-arid and desert climatology. Building a large enough flood peak dataset was a challenge and the reduced number of gauges for which data was available with a long enough record made this more difficult.

Similar to the unsatisfactory performance with the L-CAM dataset, the 8 excluded regions from data-rich CONUS performed poorly when trained also with the IMERG-ER and ERA5-L data as input. This observation was also noticed for any models trained in the external domains that were climatically classed as arid-steppe or desert-like regions. For example, the northern regions of Chile, or central and southern regions of Australia. Since models trained and tested in the same region could not perform, i.e. as baseline test cases, these arid-steppe or desert-like regions were altogether excluded from the study. It would not be wise to consider applying CONUS-trained models to external domains that could not perform even in CONUS. Fig. 3 depicts the 10 climatic regions in "data-rich" CONUS retained for this study. Fig. 3 depicts the 10 climatic regions in "data-rich" CONUS retained for this study.

For evaluating performance of the CONUS FPPMs, there were three key experiments. In the first experiment (#1), the FPPMs were trained and tested in CONUS using L-CAM (locally observed meteorological input from the CAMELS dataset) as input. The result was 10 CONUS, L-CAM FPPMs that set the threshold for accepting the performance of FPPMs over CONUS when trained instead with the global precipitation input (see Section 2.2). This brings us to the second and third experiments (#2 and #3) which entailed models trained and tested over CONUS, but with IMERG-ER and ERA5-L global precipitation data. Respectively, this results in 10 CONUS, IMERG-ER FPPMs and 10 CONUS ERA5-L FPPMs. If the results of experiments #2 and #3 are comparable to the results of experiment #1, then we will be able to confirm the first objective of this study: *global precipitation data products are viable inputs for predicting flood peaks with a trained ML model in a given region.*

We repeat the three experiments for all other study domains listed in Table 1. The L-CAM datasets used for the respective domains are listed in Table 2. In Fig. 2b, the dash blue line is the baseline case for Experiment #1. The dash red line represents the baseline cases for Experiments #2 and #3 that use global precipitation (GP) input; IMERG-ER and ERA5-L. Resultantly, for example for the pseudo-ungauged Switzerland region, there are CH L-CAM, CH IMERG-ER and CH ERA5-L models. Schematically, these reference scenarios of training and validating in the same region are reflected in Fig. 2a. Fig. 2b shows the resulting models (from the three experiments) for the baseline cases, for Switzerland.

Importantly, the static variables required by models during training when supplying the global meteorological input use the same static variables as provided in the CAMELS datasets which are also used during training of the LOCAL-based (L-CAM) models.

3.3.3. Applying the "Data-Rich" CONUS models to "pseudo-ungauged" regions

We go through a climatic zone matching exercise to determine the trained CONUS, IMERG-ER and ERA5-L FPPMs, to apply in each of the pseudo-ungauged regions per study domain. We mainly utilize the Köppen Climate Classification (according to Chen and Chen 2013) as a guide.

Detailed hydrological or geomorphological characteristics local to each pseudo-ungauged region that are not expressed at the coarse-climate designations of the Köppen Classification system constitute the other part for selecting the appropriate (i.e., regionally similar) CONUS IMERG-ER and ERA5-L FPPMs to apply. Fig. 4 illustrates each CONUS FPPM being applied to regions within each pseudo-ungauged study domain. The details that led up to the matched regions in Fig. 4 follows.

The Mediterranean-like US West Coast FPPM: The western coast of the US has a Mediterranean climate, with very hot summers where most of the precipitation (convective storms) over the region occurs in the winter-early spring seasons. The **Pacific Northwest** has more of a cool

Fig. 3. The ten CONUS regions for which trained FPPMs are available for application to similar climatic regions. The CONUS FPPMs were trained with, and prepared to test input from, either ERA5Land or IMERG hydrometeorological data.

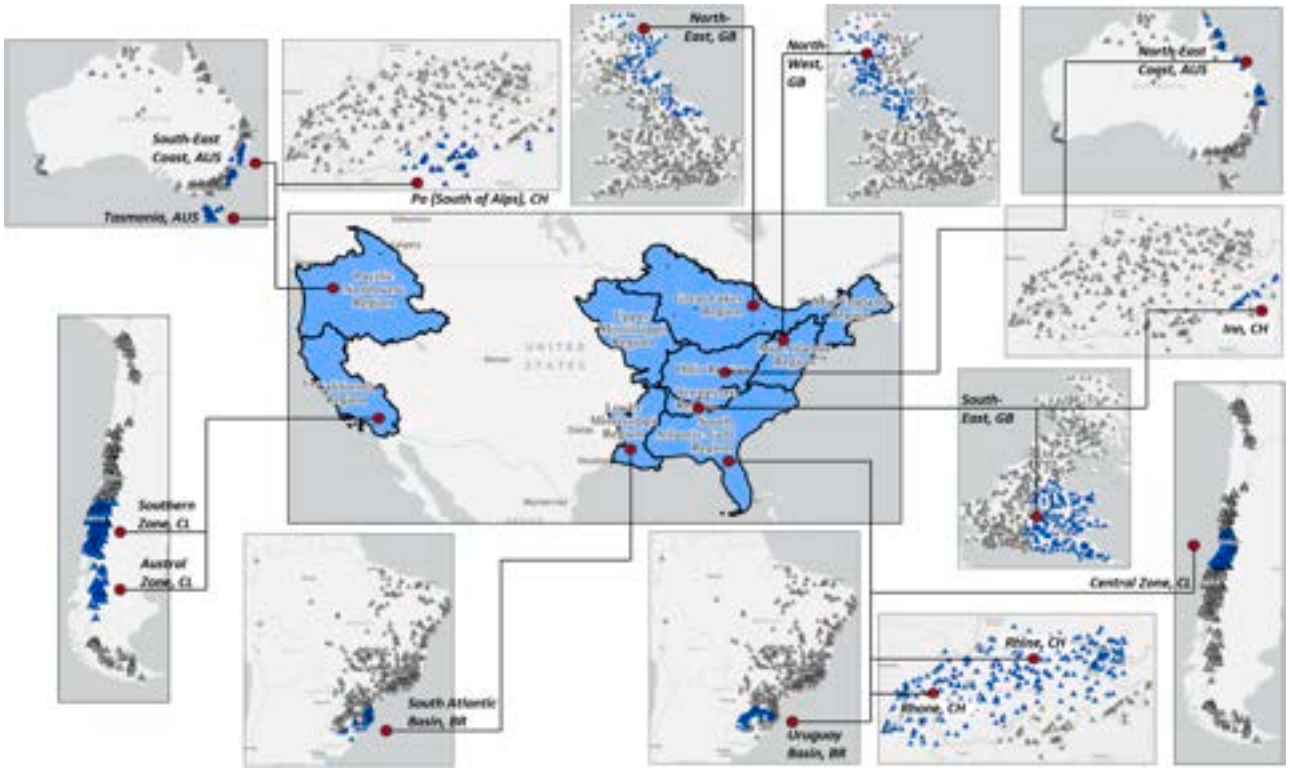

Fig. 4. CONUS FPPMs (shaded blue basins in US Map, center) being applied to external, pseudo-ungauged regions (triangles identify with all catchments of each external domain and the blue triangles indicate the specific test catchments used in the application of the line-matched CONUS FPPM). For example, the CONUS California FPPM was applied to the Southern and Austral Zones of Chile. CONUS FPPMs applied are the same for models trained with ERA5-L and IMERG-ER global precipitation datasets.

and wet climate compared to the dryer **California region**. To the south, south-east of the Pacific Northwest region and east of the California region is the mountainous Great Basin region that contribute to orographic precipitation events occurring over the two mentioned US West regions. Catchments *south of the Swiss Alpine Mountains* share a similar geography as the Pacific Northwest region and the said catchments together with the *South-East Coast and Tasmania regions of Australia* are impacted by a similar Mediterranean-like climate. The *Southern and Austral Zones of Chile* are yet another location that experiences weather and climate like the US West. While the latitudinal coverage of these Chilean Zones theoretically implies colder temperatures, in actuality, there is a general cool temperature throughout the region all year round, with precipitation occurrences much like the California region (south of the US West Coast).

The Mild-Temperate, Mid-Atlantic FPPM: A region with cold winters and warm summers and varying physical geography from coastal plains to the east, the piedmont at the foothills of the Appalachian range follows inwards then a mix of ranges and valleys and rolling hills and plains to the northwest of the region. Land cover for the region is mostly forests and predominantly agricultural activities ensue. The varying

elevation with the presence of mountains, the general geography, the influence of the Atlantic Ocean, and the latitudinal location of the region are all shared by the *North-East Highlands of the United Kingdom*.

The Subtropical, South-Eastern Regions: Three regions (and therefore three FPPMs) within CONUS, share a general subtropical climate type: the South Atlantic Gulf, Tennessee and Lower Mississippi regions (see Fig. 3). Being fully humid regions, precipitation does not vary significantly from one season to the next. Therefore, flood triggering events (using precipitation as the key driver and all other variables held constant) are possible all year round. In terms of temperatures, from south to north, the summer season varies from hot to warm, respectively.

The South Atlantic Gulf FPPM: Bordered and affected by the weather patterns of the Atlantic to the east and the Gulf of Mexico to the south and south-west, this region experiences hot and humid summers where short-duration, high intensity storms are common, often leading to flash floods. The South-Atlantic region is also prone to tropical cyclones which often propagate to flood events with devastating impacts. On the leeward side of the Swiss Alps (i.e., *north, and north-west Switzerland*) we find a transition zone between oceanic and continental climates along the Swiss Plateau that experience distinct seasons with temperatures that are not too extreme given the influence of the Atlantic Ocean. This creates a similar climate like the upper latitudes of the South Atlantic Gulf region.

More easily analogous in climate as seen in the South Atlantic Gulf region is the *Central Chilean Zone* and the *Uruguayan basin* segment located in Brazil. These two regions are located within the same latitude as the South Atlantic Gulf region and share the fully humid, mild temperate climate with hot summers as seen along the general south-east quarter of the contiguous US. It is during the summers that the flood generating storm events are most prevalent in both regions. From a climatic perspective even the FPPMs for the Tennessee and Lower Mississippi regions would be applicable for application to the Central Chile Zone and Southern Brazil (the Uruguay Basin).

The Appalachian-Containing, Tennessee FPPM: Much like the Inn Basin, that lie in the higher elevations along the Swiss Alps, the drivers of flood events in the region behave similar to those within the subtropical Tennessee basin. From the foregoing paragraph, we note the climatology of the Switzerland area, which therefore makes the attribution of the Tennessee FPPMs to the Inn Basin, reasonable.

The Agricultural-Intensive, Lower Mississippi FPPM: Sharing a similar subtropical and humid climate as the South Atlantic Gulf basin in contiguous US, the Lower Mississippi basin contains the Mississippi River which is the largest river in North America - it discharges into the Gulf of Mexico runoff from not only the Upper Mississippi but also Missouri and Ohio Basins. The alluvial plains of the Lower Mississippi are noted for supporting a vast diversity of agriculture in the US given the conducive temperatures and abundance of precipitation and streamflow network throughout the latitudinal extents of the basin. As such we apply this model to the South Atlantic basin of Brazil; a fellow subtropical basin, and a key developed water region supporting an abundance of agricultural activity in Brazil.

The Ohio FPPM: The great majority of Ohio to the south is subtropical climate and to the north, a snow climate region as we head into the Upper Mississippi basin and further north, the Great Lakes region. As a mid-latitude basin with the absence of mountains both north and south of the basin, there are varying extremes that catchments within the basin are exposed to. In the winter, cold air masses from the north (Arctic) and warm air masses from the south (Gulf of Mexico) during the summer season. Irrespective of the season, floods are a worrisome hazard (Frankson et al., 2022; Austin et al., 1998) for residents with floods induced during the winter season from snow or ice storms, and by convective rainstorms (and thunderstorms) during the late spring and summer seasons. *Australia's north-east coast* similarly lacks any sort of mountains or elevated land masses and shares this sub-tropical, fully humid climate, year-round. The east coast of Australia is generally wetter than the west coast, and from a strict climate match, the Tennessee and Lower Mississippi FPPMs are also good candidates.

3.3.4. Metrics for evaluation

The root mean square error (RMSE) is one of two metrics used to evaluate the FPPMs. RMSE is a commonly used metric for assessing regression-based ML models and follows as a good choice given that the loss function of the FPPMs is MSE. In addition to an absolute error metric, we also employ a relative error metric, which is the median value of the percent relative difference (PRD) as a second metric for inter-regional comparisons. PRD also provides information on the direction (overestimation or underestimation) in magnitude of flood peak predictions.

$$RMSE = \sqrt{E\left\{\left(X_{prediction} - X_{observation}\right)^2\right\}} \tag{3}$$

$$PRD\ (\%) = median\left(\frac{\left[X_{prediction} - X_{observation}\right]}{E\left\{X_{observation}\right\}}\right) x\ 100 \tag{4}$$

where X is the series of flood peak magnitudes predicted from FPPMs ($X_{prediction}$) or observed ($X_{observation}$) as reported in the CAMELS dataset and $E\{\}$ corresponds to the arithmetic mean operator.

4. Results and discussion

4.1. Comparison of precipitation input in "Data-rich" CONUS

In Fig. 5, we demonstrate the performance of the CONUS models for precipitation input from three sources: L-CAM, IMERG-ER and ERA5-L. We considered any CONUS-based, IMERG-ER and ERA5-L model acceptable for predicting flood peaks if PRD performance is within 30 % error and therefore measures up to the respective, baseline L-CAM model trained and validated using the CAMELS-US dataset. Comparative performance is noted between the IMERG-ER and ERA5-L models and the baseline L-CAM models in terms of RMSE (Fig. 5a). Relative differences within 30 % are considered acceptable performance per regional model. Fig. 5b illustrates that this holds for all regions to a large degree, with some regions (e.g. New England and Mid-Atlantic) exhibiting higher variability in performance; a finding consistent for all three precipitation datasets.

Non-parametric, Wilcoxon Rank tests were conducted for ERA5-L and IMERG-ER separately against the CONUS L-CAM dataset with the null hypothesis that they have the same median and are therefore drawn from the same distribution (i.e. ERA5-L and IMERG-ER FPPMs have comparative performance with the CONUS L-CAM dataset). For all ten CONUS regions, and for both ERA5-L and IMERG-ER, the results indicated a failure to reject the null hypothesis at the 5 % significance level. Overall, another noticeable feature from these results is that from a PRD perspective, the CONUS FPPMs have a greater tendency to overestimate peak flows than underestimate (i.e. boxplots are predominantly placed above zero value in Fig. 5).

4.2. Application of trained regional models to "similar" external regions

The CONUS FPPMs trained on ERA5-L and IMERG-ER datasets were applied to the external, pseudo-ungauged regions as summarized in Fig. 4. The FPPMs were applied following the approach detailed in the methodology (Section 3.3.3) for matching climatic characteristics. For a spatial understanding of the CONUS ERA5-L and IMERG-ER FPPMs applied to external regions, refer to Fig. 4. The South Atlantic Gulf FPPMs seems to be the most "versatile" of the list given acceptable test performance in Switzerland, Brazil, and Chile. The specific basins in these countries for which the CONUS, South Atlantic Gulf models are applied do share similar climatology as the CONUS region (i.e. subtropical), and we surmise that a key contributing factor for performance is the similar type of flash floods the regions face during the

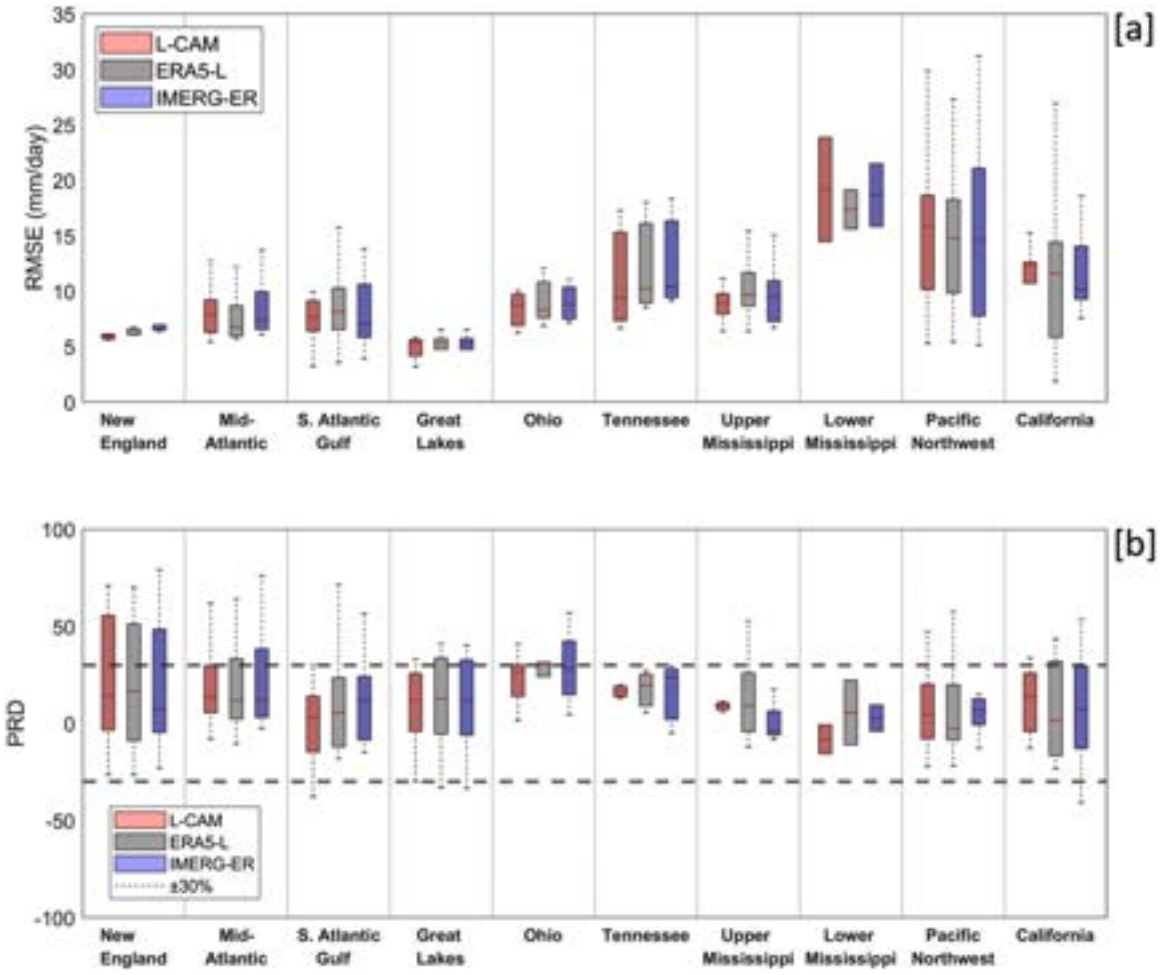

Fig. 5. Comparison of model performance [in terms of RMSE [a] and PRD [b]] for the three precipitation datasets for the CONUS domain (considered the "Data Rich" region).

summer seasons triggered by local, extreme precipitation events.

The median RMSE and PRD performance of each CONUS FPPM when applied to the pseudo-ungauged regions depicted in Fig. 4, is demonstrated in Figs. 6 and 7, respectively. Each row reflects the results for a specific climate region matched in a single external pseudo-ungauged domain. Evaluation of each applied CONUS FPPM is essential from three perspectives, shown in Columns B, C and D. A comparison between ERA5-L and IMERG-ER FPPMs when trained and tested in each CONUS region is shown in Column B (of Figs. 6 and 7). We observe similar performance between the two FPPMs for the global precipitation products except for the California (Panel B4, Fig. 6) and Pacific Northwest (Panel B1, Fig. 6) models where the CONUS, ERA5-L FPPMs perform slightly better than the IMERG-ER FPPMs. The CONUS, California IMERG-ER FPPM had a median PRD of +25 %; one of the highest regional overestimations of flood peaks noted across all FPPMs (see Panel B4, Fig. 7). Instead, the CONUS, California ERA5-L FPPM reported the lowest median PRD at just 3 % overestimation in flood peaks across the region. As accuracy in the global precipitation products have significant geographic dependencies (Derin et al., 2019; Pradhan et al., 2022), we surmise that the ERA5-L may be better at representing flood inducing storm events over the CONUS West Coast.

Column C of Figs. 6 and 7 report the performance of FPPMs trained and tested within the pseudo-ungauged regions. The results in Column C are meant to provide a baseline to compare the results in Column D where the CONUS FPPMs were applied to the external regions. As highlighted in Column C, the FPPMs driven by global precipitation data in the regions performed similarly and, in some cases, slightly better than the FPPM driven by the L-CAM data for the region (e.g. in Panel C3, the CH, Inn Basin IMERG-ER FPPM outperforms the said region's L-CAM FPPM). This is rather counter-intuitive, given that one would expect the local data from CAMELS to be associated with best performance. However, it is worth noting that when considering catchment-averaged precipitation (which we use as a predictor), it is possible to have cases of large catchments and/or complex terrain where IMERG-ER and ERA5-L outperforms station-based averages. This is a plausible speculation but verifying it is beyond the scope of this work.

All three FPPMs trained and tested over the Southern Zone of Chile (Panel C4) and the Uruguay Basin in Brazil (Panel C2) showed decreased performance with respect to the RMSE (Fig. 6). Intermittent streamflow records, and lack of some catchment attributes dwindled not only the number of catchments available to train and test but also the number of flood peaks acquired for the databases in these regions. A closer look at relative performance, shows acceptable PRD within ±10 % for the Uruguay Basin (Panel C2, Fig. 7), but the model that employs the L-CAM, Chile dataset (Panel C4) as input, indicates one of the highest errors (across all experiments) for both the RMSE and PRD metrics. While PRD results are within acceptable ranges, there is an overall tendency for models to overestimate flood peaks with no distinct trend towards

ERA5-L or IMERG-ER data input in the external regions. Geography may also play a role and we note that not only the global precipitation datasets incur difficulty with reflecting weather conditions but also the locally-based meteorological data for the country as is seen in the L-CAM, Chile dataset (Panel C4, Figs. 6 and 7).

With reference to the baseline performance shown in Columns B and

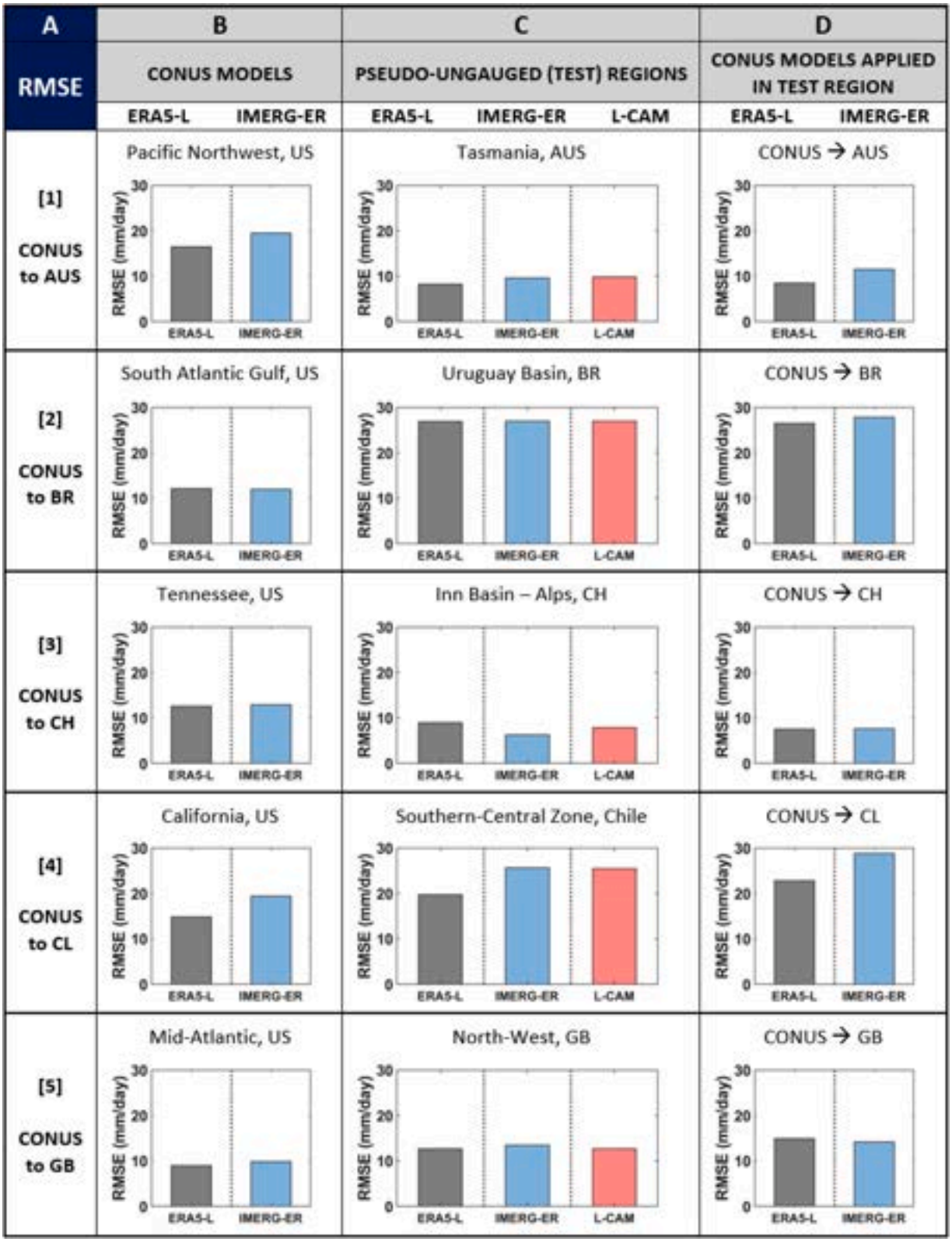

Fig. 6. RMSE performance of CONUS FPPMs when applied to pseudo-ungauged regions. Each row reflects the results for a specific climate region in a single external pseudo-ungauged domain. Column A tells the region where the CONUS FPPM is applied to. Column B gives the median RMSE for the CONUS FPPM when tested in its own region, for both the IMERG-ER and ERA5-L models. Column C gives the performance of FPPMs trained and tested in the designated climate region within the pseudo-ungauged domain, for the three precipitation datasets (IMERG-ER, ERA5-L and L-CAM). Column D reports the performance when the CONUS, IMERG-ER and ERA5-L FPPMs in Column B, are applied to the climatic region mentioned in Column C.

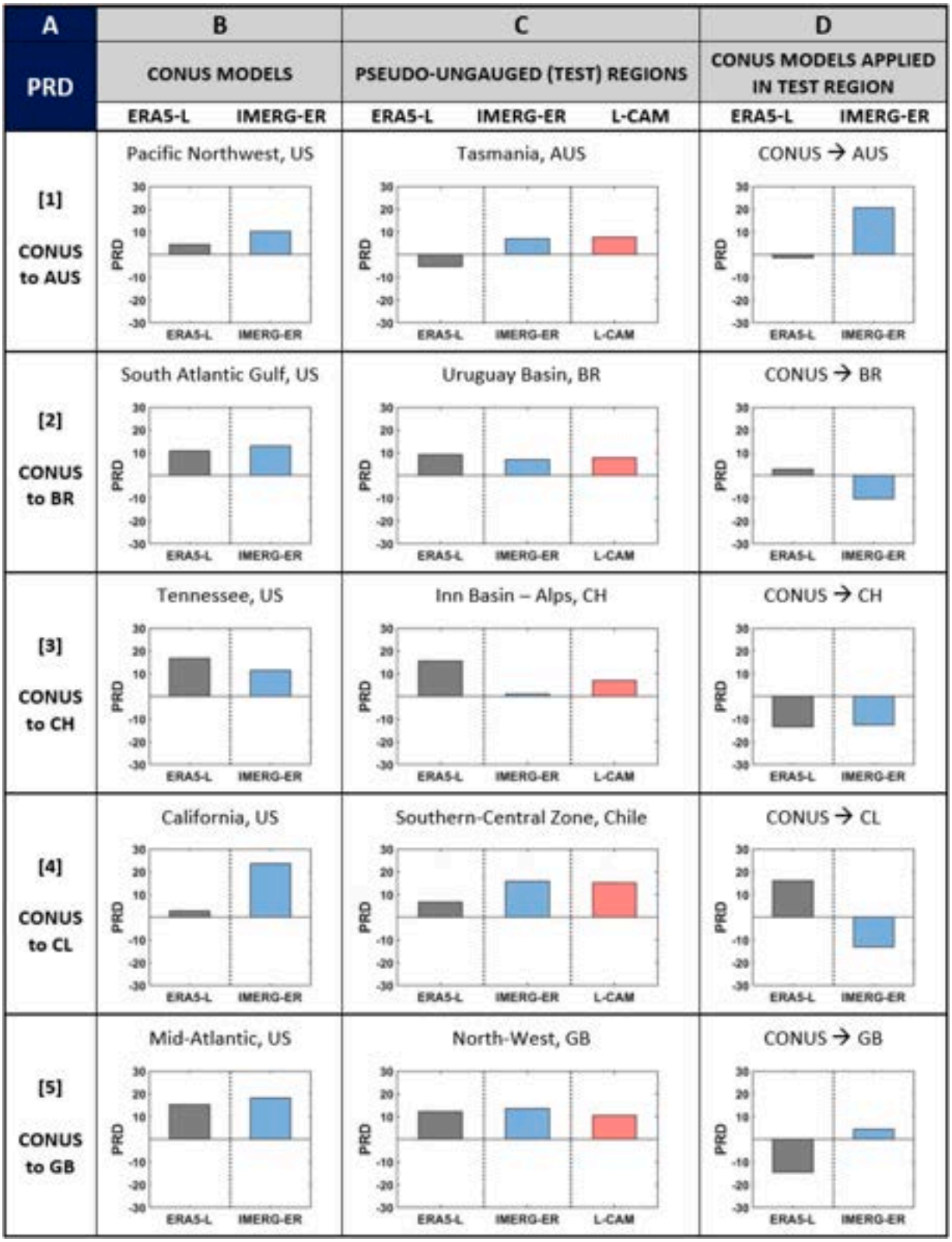

Fig. 7. Same as Fig. 6, but for median PRD performance.

C, *the results (for all applied cases demonstrated in Column D) validated the application of the trained CONUS, ERA5-L and IMERG-ER FPPMs to external regions based on climatic similarity*. Fig. 7 (and Supplementary Figure S2) attest that median regional peak predictions across all regions are within ~20 % relative difference to the observed median flood peak magnitude (Column D). Overall, the CONUS-trained, ERA5-L FPPMs reveal a higher likelihood of overestimating flood peaks than do the IMERG-ER equivalent models. Consistent underestimation of flood peaks is evident for all applied cases of the CONUS-trained, South Atlantic Gulf, IMERG-ER FPPM (Panel D2, Fig. 7). This includes Chile's Southern (Panel D4, Fig. 7) and Central Zones (Supplementary Fig. S2) and the Rhine and Rhone Basins situated North and North-west of the

Alps in Switzerland (Fig. S2). However, it is worth noting that in most cases PRD is well within 20 %.

It is worthy to note some other trends where the trained CONUS FPPMs applied to external regions show improved performance over the FPPMs trained and tested in the external regions, and vice versa. Interesting are the cases where the CONUS FPPM applied to the pseudo-ungauged region performs better than (1) the FPPMs trained and tested in the external region itself (i.e. Column D outperforms Column C), and (2) the performance on the test catchments in the very CONUS region for which the FPPM was trained (i.e. Column D also outperforms Column B). One such example, the CONUS Mid-Atlantic FPPM based on IMERG-ER applied to the North-West region of Great Britain (Panel D5, Fig. 7). We can infer that the data quantity and quality over data-rich CONUS afforded a well-trained CONUS Mid-Atlantic FPPM (Panel B5, Fig. 7) that aided predictive performance when applied over the North-west, GB region. This performance is better than the GB, IMERG-ER FPPM tested in its own region (Panel C5, Fig. 7). Similar cases are also noted for the GB South-west and North-east regions (Supplementary Fig. S2). Albeit with less significant differences, the trend is also evident for the ERA5-L FPPMs when applied to the Uruguay Basin in Brazil (Panel D2, Fig. 7) and the Inn Basin in Switzerland (Panel D3, Fig. 7).

There is no conclusive choice for one global precipitation FPPM being superior to the other. Although, for the regions where either the ERA5-L or IMERG-ER FPPMs result in PRD values > 30 %, the equivalent L-CAM models also indicate similar magnitude errors. For example, in Supplementary Fig. S2, the North-east Coast of Australia, the North-east and South-west regions of Great Britain and the Ohio Basin in CONUS, all posit this finding. The CONUS-trained, IMERG-ER FPPMs have a slight tendency to underestimate flood peak magnitudes compared to the ERA5-L FPPMs, particularly for the applied cases of the South Atlantic Gulf FPPM.

Further analyses of the magnitude of flood peaks predicted when the CONUS FPPMs were applied to the external regions indicated a general *dependence of prediction performance with drainage area*. RMSE and PRD of the normalized peak flows have opposite trends because of the inherent scale dependence of these variables (i.e. lower values for higher drainage areas) and its impact to each metric (e.g. low values in denominator of PRD equation results in higher PRD values and the opposite for RMSE). Beyond that, we expect that global precipitation dataset will have increased accuracy of catchment-average precipitation for larger drainage areas (e.g. due to reduction of random error) but it is not straightforward to attribute such scale dependence (shown in Fig. 8) to

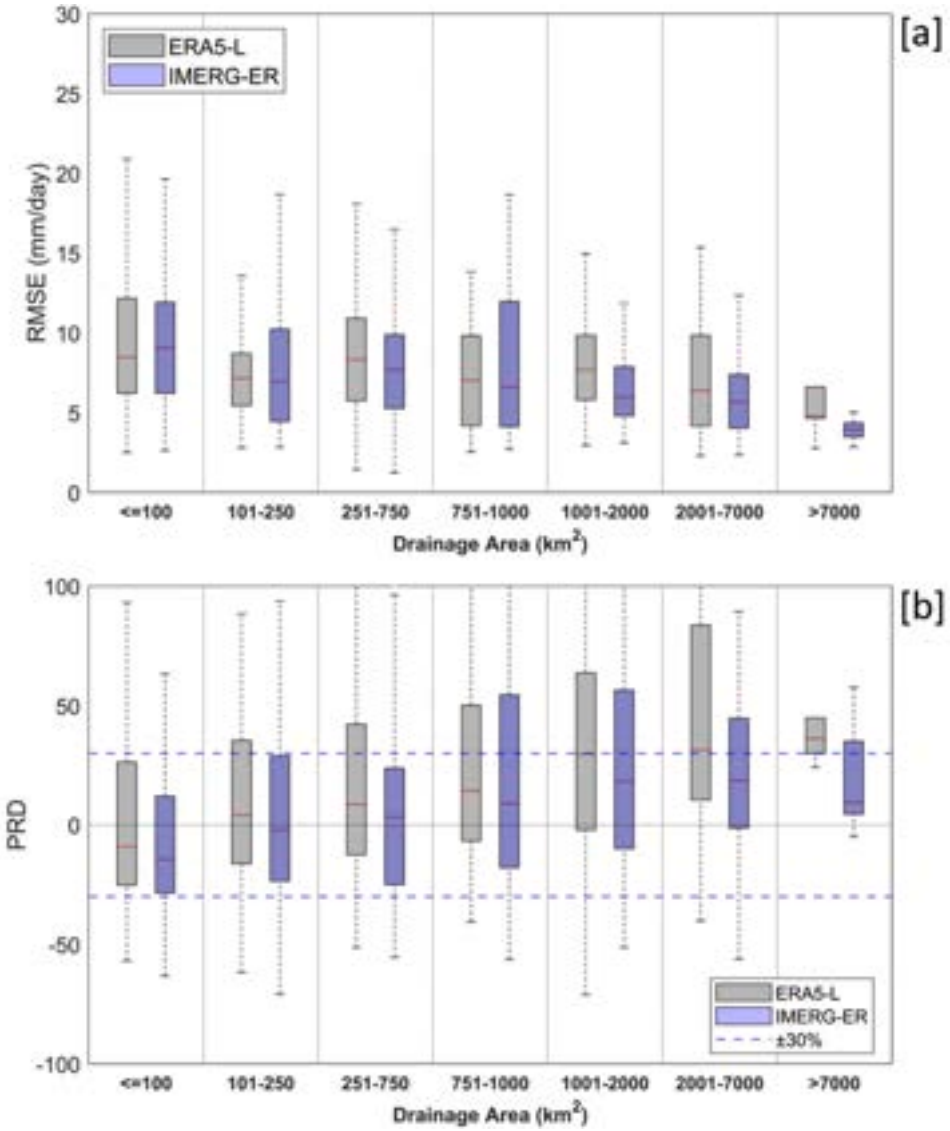

Fig. 8. Skill in flood peak prediction across drainage areas (binned) using RMSE [a] and PRD [b] as proxies. The error in prediction when the trained CONUS, ERA5-L and IMERG-ER FPPMs are applied to all the relevant external regions is shown (see Fig. 4 for all applied FPPMs).

this expectation, simply because there is an interplay with the model's performance (irrespective of input) across scales. This potentially highlights that training and testing ML models for certain ranges of drainage areas may be an important consideration, but it requires a sufficiently high number of catchments for each range of drainage areas, which was not the case here.

5. Conclusion

The main objective of this work was to investigate the potential of combining global hydrometeorological datasets with ML models, which have been trained over data-rich regions, for predicting flood peaks in ungauged regions. The key components of this investigation included the a) ML models trained over data-rich CONUS, b) the use of global precipitation datasets, IMERG-ER and ERA5-L and c) identifying and matching climatic characteristics between the data-rich region used to train the ML model and the ungauged region to which it was applied. Within CONUS we were able to assemble and validate flood peak prediction models for 10 unique climatic zones. Given the data limitations, we were only able to successfully apply 8 CONUS FPPMs to a total of 15 pseudo-ungauged regions in 5 countries: Australia, Brazil, Chile, Great Britain, and Switzerland. Performance with the applied FPPMs for both IMERG-ER and ERA5-L, yielded results within 30 % PRD in flood peak magnitude. We consider this an acceptable performance especially if one considers that the target is to obtain flood peak estimates in purely ungauged regions where there is no other source of information.

Results across regions and precipitation datasets exhibited considerable variability, without though highlighting clearly which precipitation dataset should be considered as the one associated with best overall performance. Something to be highlighted though is the fact that IMERG-ER is a near real-time precipitation product which opens opportunities for operational early warning procedures, while ERA5-L can be considered more suitable for long-term climatological studies. The dependencies of prediction performance across climatic regions and drainage scales provided in this work, offers an initial guide for implementation of such approaches, and furthermore highlights areas for improvement. Future work needs to focus on more climatic regions, which were not covered in this work, such as the typically arid, desert or steppe-like climates, which are largely ungauged and very challenging for predicting their hydrologic response. The framework presented in this work relied, as a first step, only on climate similarity for matching trained models with ungauged areas. Future work can potentially adopt more rigorous regional classification methods that are based on catchment-related predictors (Massari et al., 2023) to refine the method by which ML models trained in data-rich regions are developed and applied to ungauged regions

This work highlights the potential of combining global precipitation dataset with pre-trained ML models in data-rich areas, for flood prediction in ungauged areas with similar climate. It provides a starting point from which more and improved ML procedures and precipitation datasets can be integrated to potentially address the PUB (Prediction in Ungauged Basins) problem at the global scale.

CRediT authorship contribution statement

Zimeena Rasheed: Writing – original draft, Visualization, Methodology, Formal analysis. **Akshay Aravamudan:** Writing – review & editing, Methodology. **Xi Zhang:** Writing – review & editing, Methodology. **Georgios C. Anagnostopoulos:** Writing – review & editing, Supervision. **Efthymios I. Nikolopoulos:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the link to code at the attach file step.

Acknowledgments

This work was supported by the National Science Foundation, United States, under Grant No. (2243809). We would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.advwatres.2024.104781.

References

Addor, N., Newman, A.J., Mizukami, N., Clark, M.P., 2017. The CAMELS data set: catchment attributes and meteorology for large-sample studies. Hydrol. Earth Syst. Sci. 21 (10), 5293–5313. https://doi.org/10.5194/hess-21-5293-2017.

Alvarez-Garreton, C., Mendoza, P.A., Boisier, J.P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., Ayala, A., 2018. The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset. Hydrol. Earth Syst. Sci. 22, 5817–5846. https://doi.org/10.5194/hess-22-5817-2018.

Andreadis, K.M., Wing, O.E., Colven, E., Gleason, C.J., Bates, P.D., Brown, C.M., 2022. Urbanizing the floodplain: global changes of imperviousness in flood-prone areas. Environ. Res. Lett. 17 (10), 104024 p.

Austin, G., Rizzo K., Matte A., and Finnerty B. (1998). Service assessment: *Ohio River Valley Flood of March 1997*. National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD, 35 pp. Date Accessed: February 15, 2024. https://repository.library.noaa.gov/view/noaa/6398.

Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B., 2011. *Algorithms for hyper-parameter optimization.* In: Vol 24 of Neural Information Processing Systems Foundation.

Bertola, M., Blöschl, G., Bohac, M., et al., 2023. Megafloods in Europe can be anticipated from observations in hydrologically similar catchments. Nat. Geosci. 16, 982–988. https://doi.org/10.1038/s41561-023-01300-5.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., Savenije, H.H, 2013. Runof Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales. Cambridge Univ. Press.

Chagas, V.B.P., Chaffe, P.L.B., Addor, N., Fan, F.M., Fleischmann, A.S., Paiva, R.C.D., Siqueira, V.A., 2020. CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil. Earth Syst. Sci. Data 12, 2075–2096. https://doi.org/10.5194/essd-12-2075-2020.

Chen, D., Chen, H.W., 2013. Using the Köppen classification to quantify climate variation and change: an example for 1901–2010. Environ. Dev. 6, 69–79. https://doi.org/10.1016/j.envdev.2013.03.007.

Choi, J., Lee, J., Kim, S., 2022. Utilization of the Long Short-Term Memory network for predicting streamflow in ungauged basins in Korea. Ecol. Eng. 182, 106699. https://doi.org/10.1016/j.ecoleng.2022.106699.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., 2001. Introduction to algorithms, second edition. The Knuth-Morris-Pratt Algorithm.

Cosgrove, B., Gochis, D., Flowers, T., Dugger, A., Ogden, F., Graziano, T., Clark, E., Cabell, R., Casiday, N., Cui, Z., Eicher, K., Fall, G., Feng, X., Fitzgerald, K., Frazier, N., George, C., Gibbs, R., Hernandez, L., Johnson, D., Zhang, Y., 2024. NOAA's National Water Model: Advancing operational hydrology through continental-scale modeling. J. Am. Water Resour. Assoc. 60 (2), 247–272. https://doi.org/10.1111/1752-1688.13184.

Coxon, G., Addor, N., Bloomfield, J.P., Freer, J., Fry, M., Hannaford, J., Howden, N.J.K., Lane, R., Lewis, M., Robinson, E.L., Wagener, T., Woods, R., 2020. CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. Earth Syst. Sci. Data 12, 2459–2483. https://doi.org/10.5194/essd-12-2459-2020.

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C. H., Chen, H., Delrieu, G., Hsu, Y.C., Lavado-Casimiro, W., 2019. Evaluation of GPM-era global satellite precipitation products over multiple complex terrain regions. Remote Sens. 11 (24), 2936 p.

Emmanouil, S., Langousis, A., Nikolopoulos, E.I., Anagnostou, E.N., 2022. The spatiotemporal evolution of rainfall extremes in a changing climate: a CONUS-wide assessment based on multifractal scaling arguments. Earths Future 10 (3), e2021EF002539 p.

Fowler, K.J.A., Acharya, S.C., Addor, N., Chou, C., Peel, M.C., 2021. CAMELS-AUS: hydrometeorological time series and landscape attributes for 222 catchments in

Australia. Earth Syst. Sci. Data 13, 3847–3867. https://doi.org/10.5194/essd-13-3847-2021.
Frankson, R., Kunkel K.E., Champion S.M., and Easterling D.R., 2022. State Climate Summaries 2022: *Ohio State climate summary 2022*. 5 pp. NOAA Technical Report NESDIS 150-OH. NOAA NESDIS, Silver Spring, MD.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232. http://www.jstor.org/stable/2699986.
Höge, M., Kauzlaric, M., Siber, R., Schönenberger, U., Horton, P., Schwanbeck, J., Floriancic, M.G., Viviroli, D., Wilhelm, S., Sikorska-Senoner, A.E., Addor, N., Brunner, M., Pool, S., Zappa, M., Fenicia, F., 2023. CAMELS-CH: hydro-meteorological time series and landscape attributes for 331 catchments in hydrologic Switzerland. Earth Syst. Sci. Data Discuss. https://doi.org/10.5194/essd-2023-127 [preprint]in review.
Hu, R., Fang, F., Pain, C.C., Navon, I.M., 2019. Rapid spatiotemporal flood prediction and uncertainty quantification using a deep learning method. J. Hydrol. 575, 911–920. https://doi.org/10.1016/j.jhydrol.2019.05.087.
Hu, L., Nikolopoulos, E.I., Marra, F., Anagnostou, E.N., 2020. Sensitivity of flood frequency analysis to data record, statistical model, and parameter estimation methods: An evaluation over the contiguous United States. J. Flood Risk Manag. 13 (1). https://doi.org/10.1111/jfr3.12580.
Huang, H., Fischella, M.R., Liu, Y., Ban, Z., Fayne, J.V., Li, D., Cavanaugh, K.C., Lettenmaier, D.P., 2022. Changes in mechanisms and characteristics of Western U.S. floods over the last sixty years. Geophys. Res. Lett. (3), 49. https://doi.org/10.1029/2021gl097022.
Huffman, G.J., Stocker, E.F., Bolvin, D.T., Nelkin, E.J., Tan, J., 2019. GPM IMERG Early Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06. Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD.
IPCC, 2021. Climate Change 2021: the Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel On Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 2391. https://doi.org/10.1017/9781009157896 [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger].
Jarajapu, D.C., Rathinasamy, M., Agarwal, A., Bronstert, A., 2022. Design flood estimation using extreme Gradient Boosting-based on Bayesian optimization. J. Hydrol. 613, 128341 https://doi.org/10.1016/j.jhydrol.2022.128341.
Jong, B.T., Delworth, T.L., Cooke, W.F., et al., 2023. Increases in extreme precipitation over the Northeast United States using high-resolution climate model simulations. npj Clim. Atmos. Sci. 6, 18. https://doi.org/10.1038/s41612-023-00347-w.
Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. Lightgbm. a highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 30, 3146–3154. https://www.academia.edu/download/59228936/LightGBM_f20190512-55031-xxtk42.pdf.
Kohler, M.A., Linsley, R.K., 1951. Predicting the runoff from storm rainfall, 30. US Department of Commerce, Weather Bureau, Washington, USA. https://www.nrc.gov/docs/ML0819/ML081900279.pdf.
Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., Nearing, G., 2019. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrol. Earth Syst. Sci. 23 (12), 5089–5110. https://doi.org/10.5194/hess-23-5089-2019.
Ma, K., Feng, D., Lawson, K., Tsai, W.-P., Liang, C., Huang, X., et al., 2021. Transferring hydrologic data across continents – leveraging data-rich regions to improve hydrologic prediction in data-sparse regions. Water Resour. Res. 57 https://doi.org/10.1029/2020WR028600 e2020WR028600.
Massari, C., Pellet, V., Tramblay, Y., Crow, W.T., Gründemann, G.J., Hascoetf, T., Penna, D., Modanesi, S., Brocca, L., Camici, S., Marra, F., 2023. On the relation between antecedent basin conditions and runoff coefficient for European floods. Journal of Hydrology (Amsterdam) 625, 130012. https://doi.org/10.1016/j.jhydrol.2023.130012.
Ming, X., Liang, Q., Xia, X., Li, D., Fowler, H.J., 2020. Real-Time Flood Forecasting Based on a High-Performance 2-D Hydrodynamic Model and Numerical Weather Predictions. Water Resour. Res. 56 (7). https://doi.org/10.1029/2019WR025583.
Mirzaei, S., Vafakhah, M., Pradhan, B., Alavi, S.J., 2021. Flood susceptibility assessmentusing extreme gradient boosting (EGB), Iran. Earth Sci. Inf. 14 (1), 51–67. https://doi.org/10.1007/s12145-020-00530-0.
Mosavi, A., Pinar, O., Kwok-wing, C., 2018. Flood prediction using machine learning models: literature review. Water 10, 1536. https://doi.org/10.3390/w10111536.
Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.
Muñoz-Sabater, J. (2021). *ERA5-Land hourly data from 1950 to 1980*, Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Date Accessed: June 1, 2022, https://doi.org/10.24381/cds.e2161bac.
Newman, A.J., Clark, M.P., Sampson, K., Wood, A., Hay, L.E., Bock, A., Viger, R.J., Blodgett, D., Brekke, L., Arnold, J.R., Hopson, T., Duan, Q., 2015. Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance. Hydrol. Earth Syst. Sci. 19 (1), 209–223. https://doi.org/10.5194/hess-19-209-2015.
Ni, L., Wang, D., Wu, J., Wang, Y., Tao, Y., Zhang, J., Liu, J., 2020. Streamflowforecasting using extreme gradient boosting model coupled with Gaussian mixturemodel. J. Hydrol. 586, 124901. https://doi.org/10.1016/j.jhydrol.2020.124901.
Nogueira Filho, F. J. M., Souza Filho, F. de A., Porto, V. C., Vieira Rocha, R., Sousa Estácio, Á. B., & Martins, E. S. P. R. (2022). Deep Learning for Streamflow Regionalization for Ungauged Basins: Application of Long-Short-Term-Memory Cells in Semiarid Regions. Water (Basel), 14(9), 1318. https://doi.org/10.3390/w14091318.
Papalexiou, S.M., Montanari, A., 2019. Global and regional increase of precipitation extremes under global warming. Water Resour. Res. 55 (6), 4901–4914.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. https://doi.org/10.5555/1953048.2078195.
Pradhan, R.K., Markonis, Y., Godoy, M.R.V., Villalba-Pradas, A., Andreadis, K.M., Nikolopoulos, E.I., Papalexiou, S.M., Rahim, A., Tapiador, F.J., Hanel, M., 2022. Review of GPM IMERG performance: a global perspective. Remote Sens. Environ. 268, 112754.
Rasheed, Z., Aravamudan, A., Sefidmazgi, A.G., Anagnostopoulos, G.C., Nikolopoulos, E. I., 2022. Advancing flood warning procedures in ungauged basins with machine learning. J. Hydrol. 609, 127736. https://doi.org/10.1016/j.jhydrol.2022.127736.
Razavi, T., Coulibaly, P., 2013. Streamflow prediction in ungauged basins: review of regionalization methods. J. Hydrol. Eng. 18 (8), 958–975.
Sanders, W., Li, D., Li, W., Fang, Z.N., 2022. Data-driven Flood Alert System (FAS) using extreme gradient boosting (XGBoost) to forecast flood stages. Water 14 (5), 747. https://doi.org/10.3390/w14050747.
Sanjay Potdar, A., Kirstetter, P.-E., Woods, D., Saharia, M., 2021. Towards Predicting Flood Event Peak Discharge in Ungauged Basins by Learning Universal Hydrological Behaviors with Machine Learning. Journal of Hydrometeorology. https://doi.org/10.1175/JHM-D-20-0302.1.
Singh, V.P., 1988. Hydrologic Systems: Watershed modeling. Prentice Hall. https://play.google.com/store/books/details?id=QbUPAQAAIAAJ.
Solomatine, D.P., Wagener, T., 2011. Hydrological modeling. In: Treatise on Water Science, 2. Elsevier, pp. 435–457. https://doi.org/10.1016/B978-0-444-53199-5.00044-0.
Tarouilly, E., Li, D., Lettenmaier, D.P., 2021. Western U.S. superfloods in the recent instrumental record. Water Resour. Res. 57, e2020. https://doi.org/10.1029/2020WR029287. WR029287.
Tellman, B., Sullivan, J.A., Kuhn, C., Kettner, A.J., Doyle, C.S., Brakenridge, G.R., Erickson, T.A., Slayback, D.A., 2021. Satellite imaging reveals increased proportion of population exposed to floods. Nature 596 (7870), 80–86.
Viessman, W., Lewis, G.L., 1995. *Introduction to hydrology*, 4th ed. HarperCollins.
Wang, B., Lee, M.Y., Xie, Z., et al., 2022. A new Asian/North American teleconnection linking clustered extreme precipitation from Indian to Canada. npj Clim. Atmos. Sci. 5, 90. https://doi.org/10.1038/s41612-022-00318-7.
Winsemius, H.C., Aerts, J.C.J.H., Van Beek, L.P.H., Bierkens, M.F.P., Bouwman, A., Jongman, B., Kwadijk, J., Ligtvoet, W., Lucas, P.L., Van Vuuren, D.P., Ward, P.J., 2016. Global drivers of future river flood risk. Nat. Clim. Change 6, 381–385. https://doi.org/10.1038/NCLIMATE2893. Advance online publication.
WMO, 2022. Report on Early Warnings for All: *The UN Global Early Warning Initiative for the Implementation of Climate Adaptation* Executive Action Plan for 2023-2027. World Meteorological Organization.
Xu, K., Han, Z., Xu, H., Bin, L., 2023. Rapid prediction model for urban floods based on a light gradient boosting machine approach and hydrological–hydraulic model. Int. J. Disaster Risk Sci. 14 (1), 79–97. https://doi.org/10.1007/s13753-023-00465-2.
Zhang, K., Luhar, M., Brunner, M.I., Parolari, A.J., 2023. Streamflow prediction in poorly gauged watersheds in the United States through data-driven sparse sensing. Water Resour. Res. 59, e2022. https://doi.org/10.1029/2022WR034092. WR034092.
Zhang, Y., Vaze, J., Chiew, F.H.S., Teng, J., Li, M., 2014. Predicting hydrological signatures in ungauged catchments using spatial interpolation, index model, and rainfall–runoff modelling. Journal of Hydrology (Amsterdam) 517, 936–948. https://doi.org/10.1016/j.jhydrol.2014.06.032.