Evaluation of Surface Runoff Projections from Earth System Models in Major River Basins of the World

PUJA DAS, Northeastern University, USA

AUROOP R. GANGULY, Northeastern University, USA and Pacific Northwest National Laboratory, USA

Projections of surface runoff at large river watersheds is critical to inform water resources planner and policymakers. Although, Earth System Models (ESMs) provide runoff projections over the world, there is a large gap in projections of runoff over stakeholder-relevant spatiotemporal resolutions. In contrast, hydrological models provide runoffs at high spatial resolution but do not have the ability to incorporate feedbacks from atmospheric circulations for long term predictions. This paper assesses the credibility of the state-of-the-art Earth System Models to understand and project large river watershed runoffs. The results show that, there is wide variability due to forcing, model response, or internal variability among models at each watershed basins which amplifies the uncertainties of projections. However, despite gaps in process understanding, as well as intrinsic variability, the projected changes in runoff at regional and seasonal scales are significant enough to require re-evaluation of design curves, planning scenarios, and operations practices.

Key Words and Phrases: Earth System Model, CMIP6, Surface Runoff, Major River Basins

1 INTRODUCTION

According to World Economic Forum report, for tackling all major problems around the world, we must start focusing on water first. It is evident that world temperature is rising [2] and in IPCC Working Group I Report (Chapter 8), [24] it is stated that "Without large-scale reduction in greenhouse gas emissions, global warming is projected to cause substantial changes in the water cycle at both global and regional scales (high confidence)." It is also mentioned that, "Water cycle variability and extremes are projected to increase faster than average changes in most regions of the world and under all emissions scenarios (high confidence)". Due to this shift in hydroclimate, freshwater availability at local or global scales will be at stake [23]. Securing fresh water supply is one of the main agendas in the Sustainable Development Goals [10]. Moreover, the latest IPCC report (Working group I, chapter 11) [24] mentioned that "Significant trends in peak streamflow have been observed in some regions over the past decades (high confidence)." Thus, analyzing the effect of climate change on runoff, which is an integral part of the hydrologic cycle, is especially needed.

Projections from the climate models need to be credible for adaptation and decision-making [11, 15] and in a former paper where the loopholes in climate science were discussed it showed, how most climate models are not yet generating results at a resolution which can help in decision-making [22]. This study, provides the framework for rigorous evaluation of the performance of earth system model runoff with respect to reanalysis models as well as gridded observations at the major rivers around the world. The primary research questions that this study focuses on are: 1) How well runoff has been captured in state-of-the-art earth system models in comparison to Runoff from Reanalysis or Observed Data in Historical period at the larger river watersheds, 2) How will the hydrologically significant statistical parameters (long term mean, variance, trends) will be altered in future at these rivers, and 3) How well CMIP6 Models performs in terms of North American Rivers for annual scale projections.

The recent generation of ESMs are the Coupled Model Intercomparison Project version 6 (CMIP6) [6] which is developed by several climate institutions around the world with solid efforts of 5-6 years. The latest generation models give us the opportunity to analyze hydroclimate responses with improved modelling efforts [4]. Although climate models are doing relatively better job at predicting some variables such as temperature, but there are outstanding gaps in predictive understanding of other hydrologic variables such as runoff. This study discusses the hypothesis with case studies that how better predicted variables from ESMs, such as oceanic and atmospheric temperature, may have information content that can be leveraged with statistical and machine learning tools to improve predictions of less well predicted variables from ESMs, such as runoff and streamflow.

2 CURRENT STATE OF THE SCIENCE

With the advancement of earth system models (both in terms of resolutions and process understanding) many studies have been conducted with runoff simulations at global or continental scales [3, 16]. Extensive studies have been done with previous generations of CMIP models at different river basins of the world [1, 18]. But very few studies have been published with the performance evaluation of CMIP6 model runoff projections at an extensive scale. This study will present a global perspective of model performance of 30 river basins.

2.1 Dataset and Study Area

For this study, the largest 30 rivers are selected. The rivers are considered based on their discharge. In figure 1, all the river basins used in this study are shown. In the river watersheds map, rivers are colored according to the discharges they carry. The rivers with lower numbers and darker colors carry higher discharge. This image also shows the trends and variance in long term runoff based on observation data.

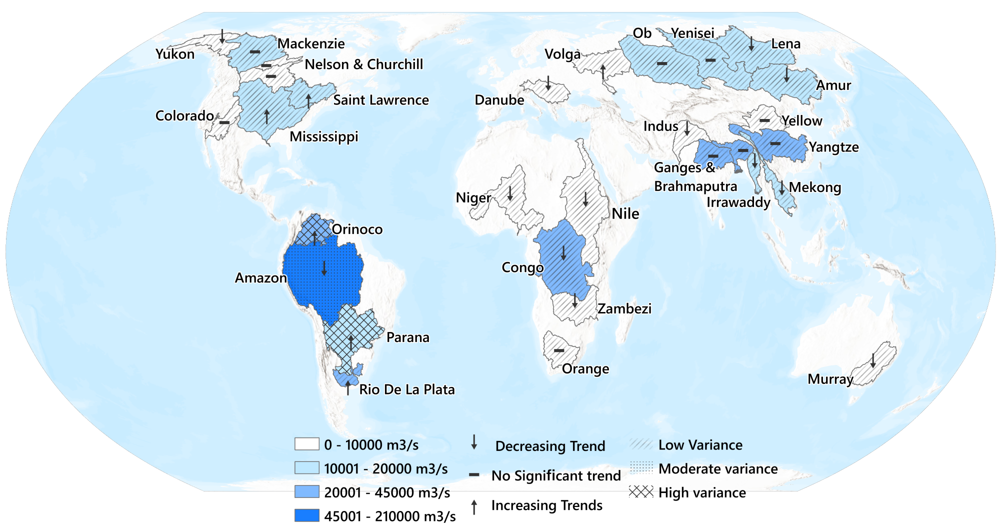

Fig. 1. Major River Basins of the World with Discharge (m^3/s).

Monthly Runoff data are collected from CMIP6 models, Reanalysis data and observed dataset. All available models with runoff projections were used in this study, discarding the models with missing data. For historical projections, 25 models were available, whereas for future projections, 21 models were found. In Table 1, the list of models used for each case and the name of their modelling group and resolutions are listed. CMIP6 models use shared socioeconomic

Table 1. IPCC CMIP6 Earth System Models with name of modeling center (of group) and horizontal grid resolution used in this study

No	Modeling center (or group)	Model name	Grid size (lat x lon)	Hist-orical	SSP 370
1	Australian Research Council Centre of Excellence for Climate System Science	ACCESS-ESM1-5	145 ×192	✓	✓
2	Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research	AWI-ESM-1-1-LR	96 x 192	✓	
3	Beijing Climate Center	BCC-CSM2-MR	160 x 320	✓	✓
4	Chinese Academy of Meteorological Sciences	CAMS-CSM1-0	160 x 320	✓	✓
5	Canadian Centre for Climate Modelling and Analysis	CanESM5	64 x 128	✓	✓
6	Chinese Academy of Sciences	CAS	128 x 256		✓
7	National Center for Atmospheric Research	CESM2	192 x 288	✓	✓
8	Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici	CMCC-CM2-SR5	192 x 288	✓	✓
9	Centre National de Recherches Meteorologiques	CNRM-CM6-1	128 x 256	✓	
10	E3SM-Project	E3SM-1-0	180 x 360	✓	
11	EC-Earth-Consortium	EC-Earth3	256 x 512	✓	✓
12	Chinese Academy of Sciences	FGOALS-g3	80 x 180	✓	✓
13	National Oceanic and Atmospheric Administration	GFDL-ESM4	180 x 288	✓	✓
14	Goddard Institute for Space Studies	GISS-E2-1-G	90 x 144	✓	✓
15	Met Office Hadley Centre	HadGEM3-GC31-LL	144 x 192	✓	
16	Institute for Numerical Mathematics	INM-CM5-0	120 x 180	✓	
17	Institut Pierre Simon Laplace	IPSL-CM6A-LR	143 x 144	✓	
18	Korea Institute of Ocean Science and Technology,	KIOST-ESM	96 x 192	✓	
19	National Institute of Meteorological Sciences	KACE-1-0-G	144 x 192		✓
20	Japan Agency for Marine-Earth Science and Technology	MIROC-ES2L	128 x 256	✓	✓
21	Max Planck Institute for Meteorology	MPI-ESM1-2-LR	96 x 192	✓	✓
22	Meteorological Research Institute	MRI-ESM2-0	160 x 320	✓	✓
23	NorESM Climate modeling Consortium	NorESM2-LM	96 x 144	✓	✓
24	Seoul National University	SAM0-UNICON	192 x 288	✓	
25	Research Center for Environmental Changes	TaiESM1	192 x 288	✓	✓
26	Met Office Hadley Centre	UKESM1-0-LL	144 x 192	✓	✓
27	Department of Geosciences, University of Arizona	MCM-UA-1-0	80 x 96	✓	✓

pathways (SSPs), which are realistic representations of socioeconomic global changes of future world [21]. In this study, SSP 370 situation has been considered as this signifies a forcing level familiar to several unmitigated SSP baselines [13] and this corresponds to a 7 w/m^2 radiative forcing [6] during the end of the century. All the CMIP6 models had data from 1860 to 2014. For reanalysis dataset, runoff datasets are extracted from the National Oceanic and Atmospheric Administration (NOAA) [14] and European Union's Earth Observation Program (ERA5) [19]. Reanalysis datasets are created from sparsely available observation data combined with data from climate models or remote sensing. Both reanalysis datasets are gridded and the grid size (latitude x longitude) for NOAA and ERA5 Runoff model is 94 x 192 and 1800 x 3600 respectively. Runoff from NOAA and ERA5 climate reanalysis datasets were available from 1948 to 2022 and 1950 to 2021 respectively. For grid-based observations of monthly runoff data, GRUN [9] dataset has been used in this study, which is available from 1902 to 2014 with a grid size of 360 x 720. These datasets were gridded using optimal interpolation. Preprocessing was performed for aligning the coordinates of all models and datasets. To maintain corresponding time frame, 1960-2010 was considered as the historical study period. For future runoff projections, runoff data for the period of 2017-2098 was selected. Also, all datasets were converted into mm/day unit for ease of comparison. CMIP6 models datasets are available at World Climate Research Program Website hosted by Lawrence Berkeley National Laboratory (https://esgf-node.llnl.gov/search/cmip6/). NOAA reanalysis datasets are available at Physical Science Laboratory Website (https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.derived.surfaceflux.html). ERA5 and GRUN runoff data can be obtained from ECMWF Website (https://climate.copernicus.eu/climate-reanalysis) and here (https://doi.org/10.6084/m9.figshare.9228176) respectively. The spatial information of the rivers was extracted from the Global Runoff Data Centre (GRDC) which can be downloaded from their website (https://www.bafg.de/GRDC/EN/02_srvcs/22_gslrs/gislayers_node.html).

2.2 Model Performance Evaluation

In this study, multi model ensemble (MME) statistical performance has been assessed as it encapsulates a collection of probable future scenarios, and the central tendency of measurement (mean or median) is the best possible representation of future in this case [15]. Long term (50 years; 1960-2010) mean, variance and trends of runoffs in the river basins are estimated from the MME of CMIP6 models, and they were compared with observations and reanalysis datasets. In figure **??**, the difference of MME with observations and reanalysis datasets are shown with the estimation of uncertainties in each river basin.

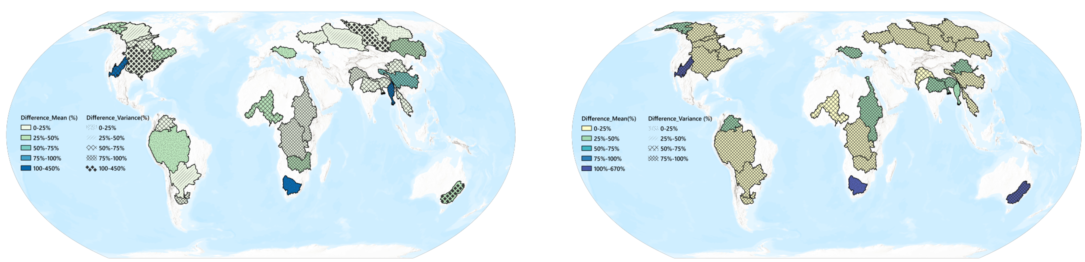

Fig. 2. Difference in long term mean and variances of CMIP6 MME with observation-based (left) and reanalysis-based (right) runoffs.

In 43% of the rivers, the difference between long term mean of CMIP6 MME and observations were less than 10%. For 83% of the rivers, this difference was less than 30%. Whereas in terms of reanalysis datasets, 60% of the rivers were showing less than 30% difference with CMIP6 MME. Reanalysis datasets (especially reanalysis data from NOAA) are showing high difference with observed datasets as well in high latitudes. For future studies, reanalysis datasets with higher spatial resolutions can be used in the study for better comparison. In terms of variance, in the majority of the rivers, CMIP6 MME has less than 25difference with observation and reanalysis based runoffs. Also, CMIP6 MME agrees more with reanalysis based runoffs in terms of variance than observation based runoffs.

Furthermore, uncertainties have been estimated here in terms of two aspects. Firstly, in fig 3, variability in model projections are shown which has been distinguished with colors. This is widely used for ESM uncertainty measurements. Secondly, variability of MME is shown with different shading options. From this analysis, it can be seen that there are high uncertainties in runoff projections at rivers with higher discharges in terms of variability in model projections. For future work, uncertainties can be measured at a more data driven perspective, such as with KL distance or Bhattacharya distance.

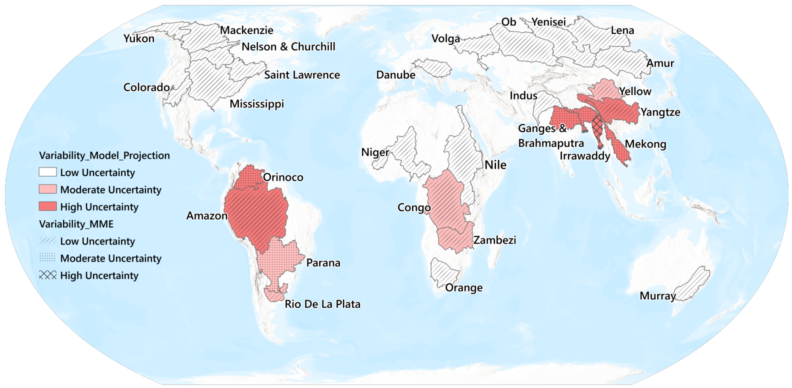

Fig. 3. Uncertainty in CMIP6 Runoff Projections of Major Rivers in terms of Variability of CMIP6 MME and Variability in Model Projections. Variability of MME lower than 0.002 mm/day,0.005 mm/day and 0.02 mm/day are considered as low, moderate and high uncertainty respectively. Variability in Model projection lower than 1 mm/day,2.5 mm/day and 4.5 mm/day are considered as low, moderate and high uncertainty respectively.

In terms of trend analysis, in 43% of the rivers, CMIP6 MME agree with observed and reanalysis-based runoffs. While assessing future projections of runoffs from CMIP6 MME, it was observed that around 75% of the river basins show an increase in their long term mean in future years. In figure 4, trends of river basin runoffs in and future time frames have been compared. For future runoffs, trends in CMIP6 MMEs have been shown. Here, a higher slope of increasing trend is observed in 60% of the river basins in the future.

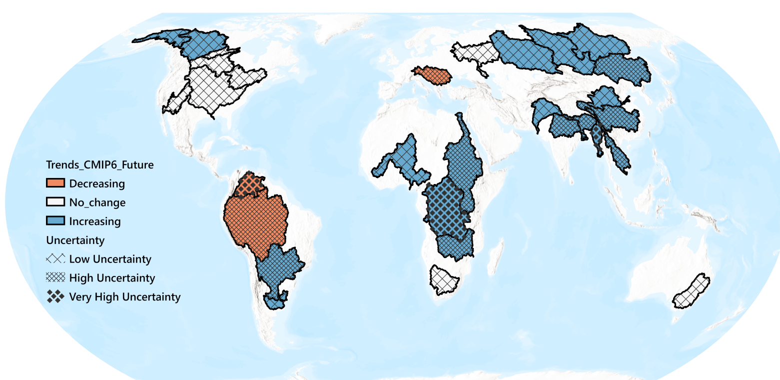

Fig. 4. Long term Trends in Runoff of Major River basin in the Future. Trends higher and lower than ±0.0005 mm/day are considered increasing or deceasing trends. Uncertainty is measured depending on how CMIP6 MME performs with respect to observed and reanalysis datasets in historical timescale.

However, for historical projections as well, trends in runoffs from 40% of the river basin were showing higher trends in CMIP6 models than observation-based runoffs. Here, uncertainty is measured depending on how CMIP6 MME performs with respect to observed and reanalysis datasets in historical timescale. So, this factor should be considered before jumping into conclusions that, in the future, runoff will increase. For further investigations, major rivers from North America continents are selected. In figure 5, mean annual runoff (1960-2010) and inter-annular variability in CMIP6 MME of 7 major rivers (Mississippi, Colorado, Yukon, Saint Lawrence, Mackenzie, Nelson and Churchill) have been presented with observation and reanalysis-based runoffs.

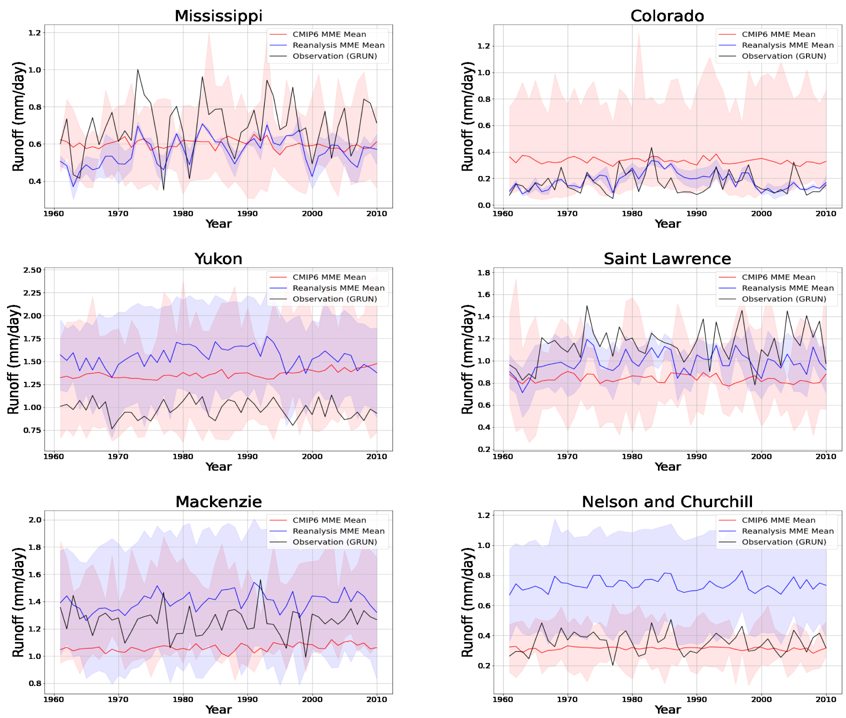

Fig. 5. Mean Annual Runoff from CMIP6 MME, Observations and Reanalysis based Runoffs in Major River Basins in North America. Red and blue shaded zones are showing spread of models from CMIP6 and reanalysis datasets.

For all the rivers, it was observed that the inter-annular variability was not clearly articulated in the CMIP6 MME. Another thing to note here is that both CMIP6 models and Reanalysis models are showing a wide spread for the river basins in higher latitudes. Lastly, results from different performance metrics such as mean squared error, Nash-Sutcliffe Efficiency and Kling-Gupta Efficiency showed that, runoff from CMIP6 MMEs are not credible enough to use for hydrologic and water resources management purposes at annual or shorter time scales.

3 FUTURE WORKS INTEGRATING HYDROCLIMATE SCIENCE WITH MACHINE LEARNING

From this study it was evident that although CMIP6 MME performs well in terms of long term statistics but for short term projections at annual or seasonal timescale, it does not have a satisfactory performance yet at individual river basins. Again, there is wide variability among models, which further increases the uncertainties of projections. This solidifies the hypothesis that climate models are not yet better at predictive understanding of hydrologic variables.

In a recent white paper [7] it was mentioned that it is highly challenging to incorporate hydrologic and hydroclimate processes and their internal interactions in earth system models because they operate in highly heterogeneous environments. However, with data driven models these problems can be tackled. Studies show many variables from CMIP6 models performs well in terms of projections. Using such variables, other hydrologically important variables such as projection of runoff can be improved with help of machine learning and artificial intelligence. Although this hypothesis is examined very often, in most cases it is not explicitly stated. For example, several studies showed how predictive knowledge can be generated using climate data with theory guided data science [8, 12]. Another widely used example of this hypothesis is downscaling of climate variables. This method tries to obtain a high resolution mapping of global scale earth system model simulation at scales important for stakeholders. Studies also showed how physics based approaches can be used to either constraint uncertainties [20] or project climate variables.

In spite of being the buzzing hypothesis in the climate community, very few studies used climate variables either from observation or earth system models to examined this above-mentioned hypothesis. A recent study, [17], showed climate variable that are well predicted from ESMs, such as sea surface or atmospheric temperature, has information content that can be integrated with statistical and machine learning tools to increase predictive understanding of less well predicted variables from ESMs, such as runoff and streamflow. They explored how using information outside the typical ENSO region integrated with AI helps in prediction of hydrology. Another study [5] used observation of precipitation extreme for estimating the dependence of extremes on covariates which helps in identifying the causal drivers and ultimately inform predictive modeling. So for future works, along with understanding the physics and biogeochemistry for the earth system models, we can quantify informed risk and improve our predictive understanding by integrating Artificial intelligence with earth system model projections [7].

4 CONCLUSION

Surface runoff is an integral component of hydrologic cycle and projections of runoff in large water basins are extremely important for ensuring food and water security as well as water resources management. In this study, we compared the performance of surface runoff from earth system models with observation and reanalysis-based runoffs in major river basins of the world. Although, multimodel ensemble performed well in terms of predicting long term means, projections of variations or trends were not satisfactory. So for better future runoff projections, along with incorporating complex hydrological parameters in ESMs, data driven approached also can be implemented.

Acknowledgements:

This project was partially funded by the U.S. Department of Defense Strategic Environmental Research and Development Program project number RC20-1183 titled Networked Infrastructures under Compound Extremes. This material is also based upon work supported by the U.S. Department of Homeland Security under Grant Award Number 22STESE00001-01-00. The views and conclusions in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

REFERENCES

[1] R Alkama, L Marchand, A Ribes, and B Decharme. 2013. Detection of global runoff changes: results from observations and CMIP5 experiments. *Hydrology and Earth System Sciences* 17, 7 (2013), 2967–2979.

[2] Richard P Allan, Ed Hawkins, Nicolas Bellouin, and Bill Collins. 2021. IPCC, 2021: Summary for Policymakers. (2021).

[3] Marcia L Branstetter and David J Erickson III. 2003. Continental runoff dynamics in the Community Climate System Model 2 (CCSM2) control simulation. *Journal of Geophysical Research: Atmospheres* 108, D17 (2003).

[4] Martyn P Clark, Ying Fan, David M Lawrence, Jennifer C Adam, Diogo Bolster, David J Gochis, Richard P Hooper, Mukesh Kumar, L Ruby Leung, D Scott Mackay, et al. 2015. Improving the representation of hydrologic processes in Earth System Models. *Water Resources Research* 51, 8 (2015), 5929–5956.

[5] Debasish Das, Auroop R Ganguly, and Zoran Obradovic. 2015. A Bayesian sparse generalized linear model with an application to multiscale covariate discovery for observed rainfall extremes over the United States. *IEEE Transactions on Geoscience and Remote Sensing* 53, 12 (2015), 6689–6702.

[6] Veronika Eyring, Sandrine Bony, Gerald A Meehl, Catherine A Senior, Bjorn Stevens, Ronald J Stouffer, and Karl E Taylor. 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* 9, 5 (2016), 1937–1958.

[7] Auroop Ganguly. 2021. *Science-integrated Artificial-intelligence for Flooding and precipitation Extremes (SAFE).* Technical Report. AI4ESP.

[8] AR Ganguly, EA Kodra, Ankit Agrawal, A Banerjee, S Boriah, Sn Chatterjee, So Chatterjee, A Choudhary, D Das, J Faghmous, et al. 2014. Toward enhanced understanding and projections of climate extremes using physics-guided data mining techniques. *Nonlinear Processes in Geophysics* 21, 4 (2014), 777–795.

[9] Gionata Ghiggi, Vincent Humphrey, Sonia I Seneviratne, and Lukas Gudmundsson. 2019. GRUN: an observation-based global gridded runoff dataset from 1902 to 2014. *Earth System Science Data* 11, 4 (2019), 1655–1674.

[10] David Griggs, Mark Stafford-Smith, Owen Gaffney, Johan Rockström, Marcus C Öhman, Priya Shyamsundar, Will Steffen, Gisbert Glaser, Norichika Kanie, and Ian Noble. 2013. Sustainable development goals for people and planet. *Nature* 495, 7441 (2013), 305–307.

[11] Mike Hulme, Roger Pielke, and Suraje Dessai. 2009. Keeping prediction in perspective. *Nature Climate Change* 1, 911 (2009), 126–127.

[12] Anuj Karpatne, Gowtham Atluri, James H Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar. 2017. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on knowledge and data engineering* 29, 10 (2017), 2318–2331.

[13] David P Keller, Andrew Lenton, Vivian Scott, Naomi E Vaughan, Nico Bauer, Duoying Ji, Chris D Jones, Ben Kravitz, Helene Muri, and Kirsten Zickfeld. 2018. The carbon dioxide removal model intercomparison project (CDRMIP): Rationale and experimental protocol for CMIP6. *Geoscientific Model Development* 11, 3 (2018), 1133–1160.

[14] Robert Kistler, Eugenia Kalnay, William Collins, Suranjana Saha, Glenn White, John Woollen, Muthuvel Chelliah, Wesley Ebisuzaki, Masao Kanamitsu, Vernon Kousky, et al. 2001. The NCEP–NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bulletin of the American Meteorological society* 82, 2 (2001), 247–268.

[15] Devashish Kumar, Evan Kodra, and Auroop R Ganguly. 2014. Regional and seasonal intercomparison of CMIP3 and CMIP5 climate model ensembles for temperature and precipitation. *Climate dynamics* 43, 9 (2014), 2491–2518.

[16] ZW Kundzewicz, B Merz, Sergiy Vorogushyn, H Hartmann, D Duethmann, M Wortmann, Sh Huang, B Su, T Jiang, and Valentina Krysanova. 2015. Analysis of changes in climate and river discharge with focus on seasonal runoff predictability in the Aksu River Basin. *Environmental Earth Sciences* 73, 2 (2015), 501–516.

[17] Yumin Liu, Kate Duffy, Jennifer G Dy, and Auroop R Ganguly. 2022. Explainable deep learning for insights in El Nino and river flows. *arXiv preprint arXiv:2201.02596* (2022).

[18] Paul CD Milly, Kathryn A Dunne, and Aldo V Vecchia. 2005. Global pattern of trends in streamflow and water availability in a changing climate. *Nature* 438, 7066 (2005), 347–350.

[19] Joaquín Muñoz-Sabater, Emanuel Dutra, Anna Agustí-Panareda, Clément Albergel, Gabriele Arduini, Gianpaolo Balsamo, Souhail Boussetta, Margarita Choulga, Shaun Harrigan, Hans Hersbach, et al. 2021. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data* 13, 9 (2021), 4349–4383.

[20] Paul A O'Gorman and John G Dwyer. 2018. Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. *Journal of Advances in Modeling Earth Systems* 10, 10 (2018), 2548–2563.

[21] Brian C O'Neill, Claudia Tebaldi, Detlef P Van Vuuren, Veronika Eyring, Pierre Friedlingstein, George Hurtt, Reto Knutti, Elmar Kriegler, Jean-Francois Lamarque, Jason Lowe, et al. 2016. The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development* 9, 9 (2016), 3461–3482.

[22] Quirin Schiermeier. 2010. The real holes in climate science: like any other field, research on climate change has some fundamental gaps, although not the ones typically claimed by sceptics. Quirin Schiermeier takes a hard look at some of the biggest problem areas. *Nature* 463, 7279 (2010), 284–288.

[23] Charles J Vorosmarty, Pamela Green, Joseph Salisbury, and Richard B Lammers. 2000. Global water resources: vulnerability from climate change and population growth. *science* 289, 5477 (2000), 284–288.

[24] Zhu Zhongming, Lu Linong, Yao Xiaona, Zhang Wangqiang, Liu Wei, et al. 2021. AR6 Climate Change 2021: The Physical Science Basis. (2021).