THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

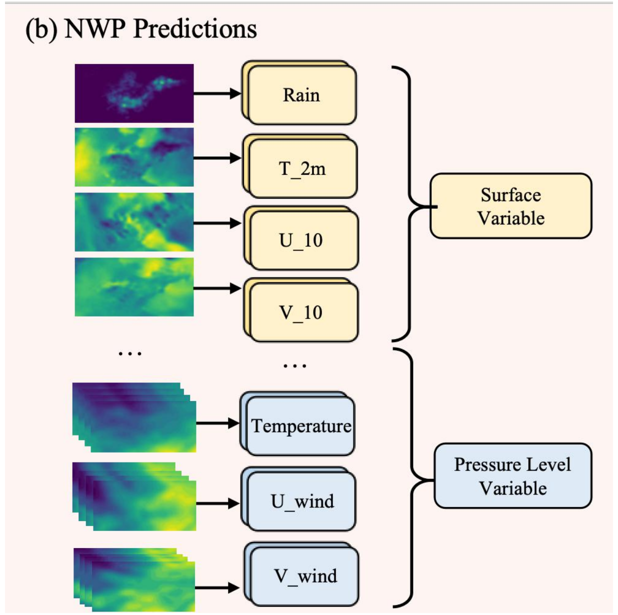
PostRainBench: A comprehensive benchmark and a new model for precipitation forecasting

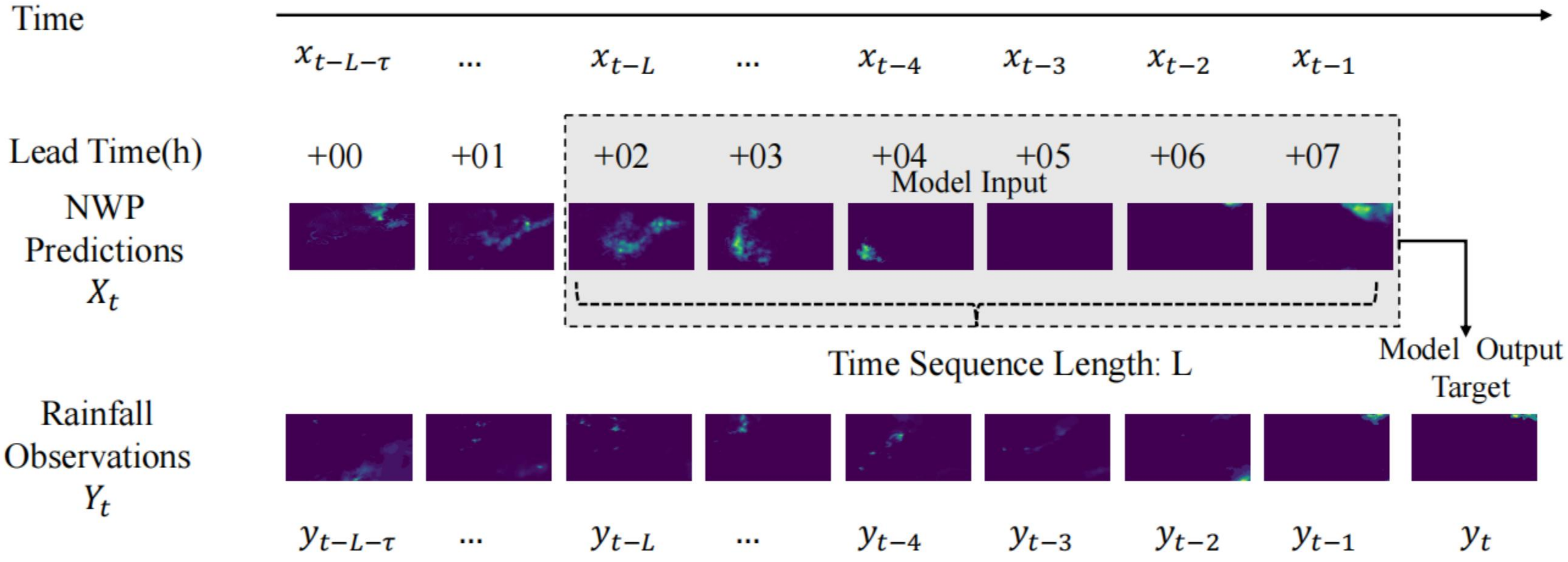
Presentation at ICLR 2024 Workshop: Tackling Climate Change with Machine Learning

Yujin Tang, Jiaming Zhou, Xiang Pan, Zeying Gong, Junwei Liang

Introduction

- Combining AI-based and NWP methods can bring about both strengths for a stronger performance.
- We focus on the Numerical Weather Prediction (NWP) post-processing based precipitation forecasting task to couple Machine Learning techniques with traditional NWP.
- For the post-processing task, NWP predictions are fed to a deep learning model which is trained to output refined precipitation forecasts, while rainfall station observations are used as ground truth.





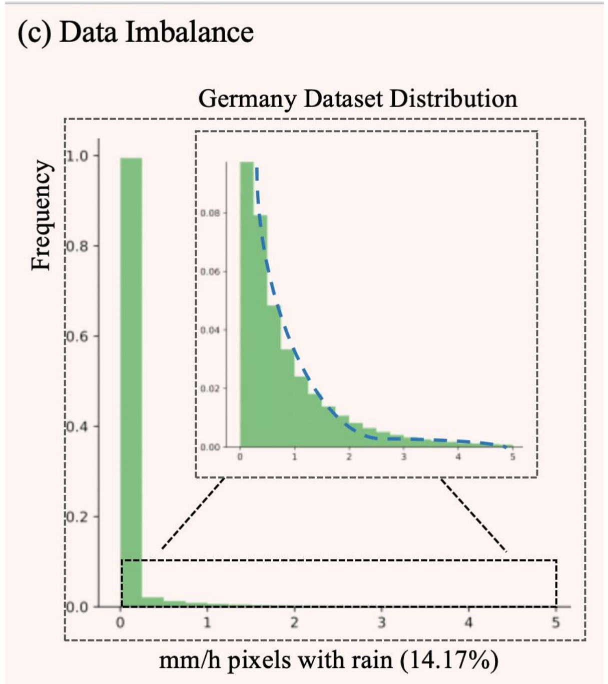
Challenges

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Numerical Weather Prediction (NWP) post-processing based precipitation forecasting poses several challenges:

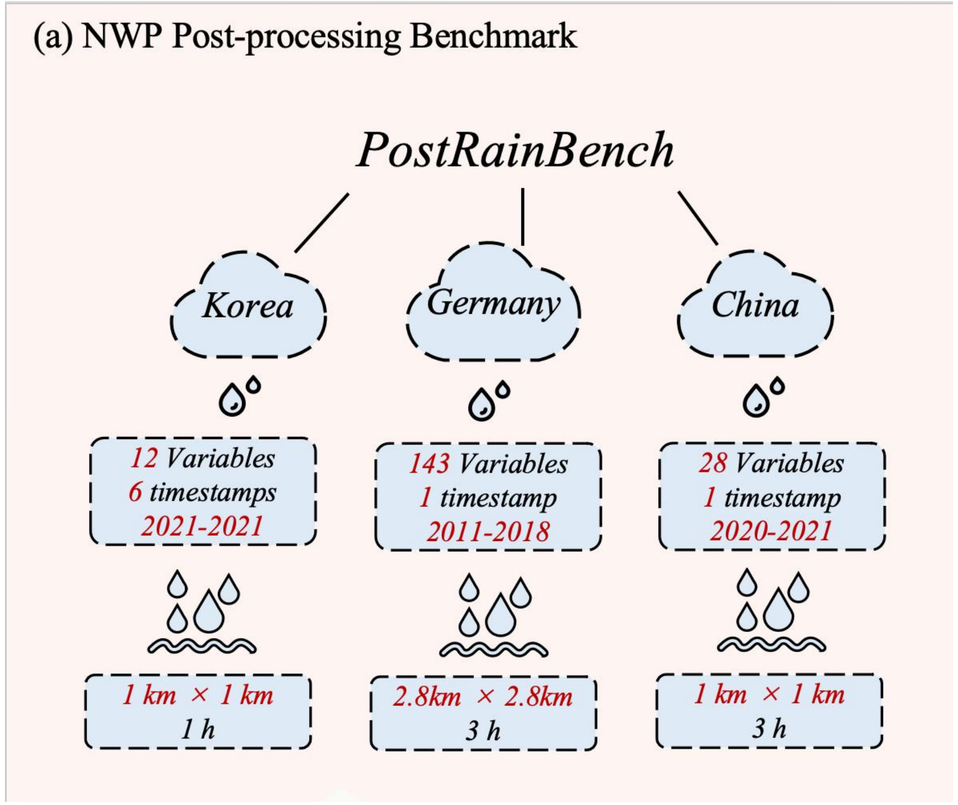
- Absence of a unified benchmark hindering cross-model evaluation.
- Uncertainty of atmospheric variable selection.
- The existing significant data imbalance.

Dataset	Rain rate (mm/h)	Proportion (%)	Rainfall Level
KoMet	$[0.0, 0.1)$	87.24	No Rain
	$[0.1, 10.0)$	11.57	Rain
	$[10.0, \infty)$	1.19	Heavy Rain
Germany	$[0.0, 10^{-5})$	85.10	No Rain
	$[10^{-5}, 2.0)$	13.80	Rain
	$[2.0, \infty)$	1.10	Heavy Rain
China	$[0.0, 0.1)$	91.75	No Rain
	$[0.1, 2.0)$	3.81	Rain
	$[2.0, \infty)$	4.44	Heavy Rain



PostRainBench

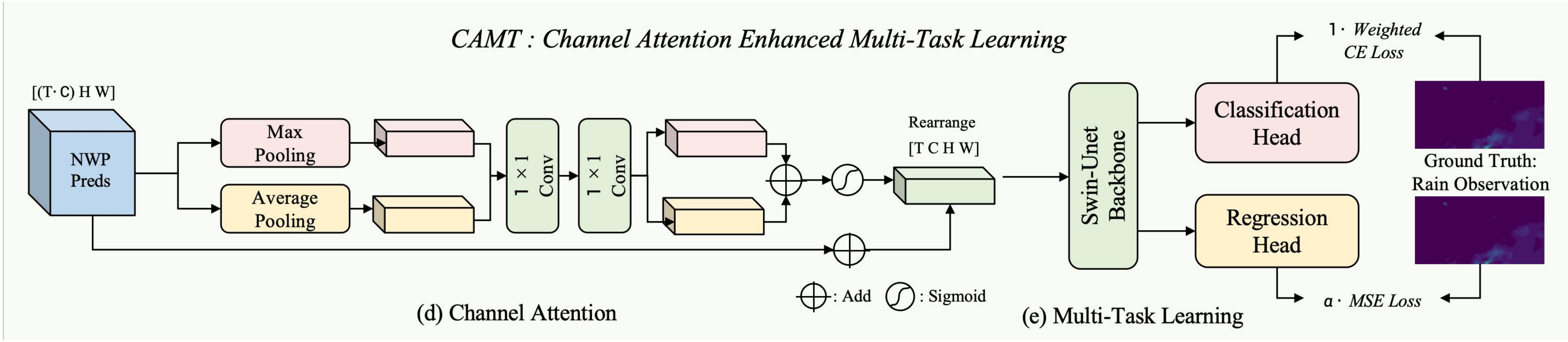
We introduce **PostRainBench**, a comprehensive multi-variable benchmark, which covers the full spectrum of scenarios with and without temporal information and various combinations of NWP input variables.



Area	Korea	Germany	China
Variable type		Pressure Level and Surface	
Variable numbers	12	143	28
Time period	2020-2021	2011-2018	2020-2021
Spatial resolution	12km × 12km	2.8km × 2.8km	1km × 1km
Temporal resolution	1h	3h	3h
Temporal Window Size	6	1	1
Data shape (T C H W)	(6, 12, 50, 65)	(1, 143, 64, 64)	(1, 28, 64, 64)
Data split [train val test]	[4920, 2624, 2542]	[15189, 2725, 2671]	[2264, 752, 760]
Data size	47.9GB	16.2GB	3.6GB

Method: CAMT Framework

We propose a new model learning framework **CAMT**, a simple yet effective Channel Attention Enhanced Multi-task Learning framework with a specially designed weighted loss function.



Experiments

Our method outperforms sota by **6.3%, 4.7%, and 26.8%** in rain CSI on respective datasets.

It's worth highlighting a significant milestone achieved by our model. In heavy rain, with improvements of **15.6%, 17.4%, and 31.8%** in CSI over NWP predictions across respective datasets. This underscores its potential to effectively mitigate substantial losses in the face of extreme weather events.

		Rain				Heavy Rain			
		CSI↑		HSS↑		CSI↑		HSS↑	
		Mean(Std)	Best	Mean(Std)	Best	Mean(Std)	Best	Mean(Std)	Best
Korea	NWP	0.263(±0.000)		*		0.045(±0.000)		*	
	U-Net	0.300 (±0.025)	0.322	**0.384**(±0.025)	**0.408**	0.006(±0.005)	0.010	0.011(±0.009)	0.018
	ConvLSTM	0.302 (±0.009)	0.312	**0.384**(±0.009)	0.395	0.009(±0.007)	0.015	0.016(±0.012)	0.026
	MetNet	0.298 (±0.012)	0.307	0.375(±0.014)	0.384	0.005(±0.007)	0.012	0.009(±0.012)	0.023
	Ours	**0.321** (±0.005)	**0.326**	**0.384**(±0.007)	0.389	**0.052**(±0.010)	**0.058**	**0.089**(±0.017)	**0.097**
	Ours Δ	+6.3%		+0%		+15.6%		+456.3%	
Germany	NWP	0.338(±0.000)		0.252(±0.000)		0.178(±0.000)		0.173(±0.000)	
	U-Net	0.491 (±0.007)	0.495	0.601(±0.006)	0.605	0.082(±0.028)	0.107	0.148(±0.048)	0.189
	ConvLSTM	0.477 (±0.026)	0.478	0.587(±0.004)	0.590	0.091(±0.041)	0.121	0.162(±0.068)	0.212
	MetNet	0.485 (±0.002)	0.487	0.595(±0.005)	0.599	0.027(±0.016)	0.094	0.147(±0.027)	0.168
	Ours	**0.514** (±0.003)	**0.518**	**0.609**(±0.006)	**0.616**	**0.209**(±0.014)	**0.224**	**0.339**(±0.020)	**0.359**
	Ours Δ	+4.7%		+1.3%		+17.4%		+96.0%	
China	NWP	0.164(±0.000)		0.123(±0.000)		0.110 (±0.000)		0.089(±0.000)	
	U-Net	0.065 (±0.007)	0.073	0.093(±0.009)	0.103	0.058(±0.014)	0.070	0.089(±0.024)	0.110
	ConvLSTM	0.054 (±0.011)	0.066	0.079(±0.009)	0.088	0.065(±0.003)	0.068	0.104(±0.010)	0.114
	MetNet	0.064 (±0.019)	0.078	0.061(±0.047)	0.106	0.057(±0.017)	0.076	0.069(±0.057)	0.118
	Ours	**0.208** (±0.007)	**0.216**	**0.274**(±0.014)	**0.289**	**0.145**(±0.015)	**0.163**	**0.225**(±0.019)	**0.246**
	Ours Δ	+26.8%		+122.8%		+31.8%		+116.3%	

Ablation Study

- We conduct an ablation study by systematically disabling certain components of our CAMT.
- Weighted loss and multi-task learning are effective in improving simultaneous forecasting under the unbalanced distribution of light rain and heavy rain, while CAM provides comprehensive improvements.

Table 2: Ablation study on Germany dataset (Rojas-Campos et al., 2022). We disable components of the framework in each experiment and report rain and heavy rain CSI as the evaluation metric.

	Weighted Loss	Multi-Task Learning	CAM	Rain CSI↑	Rain HSS↑	Heavy Rain CSI↑	Heavy Rain HSS↑
(a)	✓	✓	✓	0.514	0.609	0.209	0.339
(b)	✗	✓	✓	0.517 (+0.6%)	0.625 (+2.6%)	0.042 (−97.6%)	0.008 (−11.1%)
(c)	✓	✗	✓	0.495 (−3.7%)	0.588 (−3.4%)	0.192 (−8.1%)	0.317 (−6.5%)
(d)	✓	✓	✗	0.505 (−1.8%)	0.602 (−1.1%)	0.183 (−11.1%)	0.305 (−11.1%)
(e)	✗	✗	✗	0.521	0.628	0.000	0.000
(f)	✓	✗	✗	0.490 (−6.0%)	0.580 (−7.6%)	0.188 ↑↑↑	0.307 ↑↑↑
(g)	✗	✓	✗	0.516 (−0.1%)	0.629 (+0.2%)	0.067 ↑	0.007 ↑
(h)	✗	✗	✓	0.513 (−1.5%)	0.624 (−0.6%)	0.115 ↑↑	0.204 ↑↑

Conclusion

- We introduce a comprehensive benchmark, **PostRainBench** and a new method, **CAMT** framework.
- Our approach demonstrates outstanding performance improvements compared to baseline models and NWP method.
- Our research provides novel insights into the challenging domain of highly imbalanced precipitation forecasting tasks.
- We believe our benchmark could help advance the model development of the research community.

Thank you!
Code and Data Available!