Small River Flood and Streamflow Forecasting Using Deep Learning Models and Uncertainty Quantification: A Case Study of Kyll River Basin in Germany

Informatik
Hauptcampus
HOCHSCHULE TRIER

Dr. Mohammed Al-Bared
Prof. Dr. Hans-Peter Beise
Manfred Stüber

Forecasting small river flow rates is challenging due to

- The scarcity of data
- The sudden appearance of floods.

Forecasting rare floods challenges deep-learning models due to

- Heavily imbalanced data.
- The flow rate remains consistent 99% of the time
- Flood events are infrequent
- Extreme floods are unique

Deep learning models for Small Rivers Stream-Flow Forecasting and flood prediction

Deep learning models

Heavily rely on data.

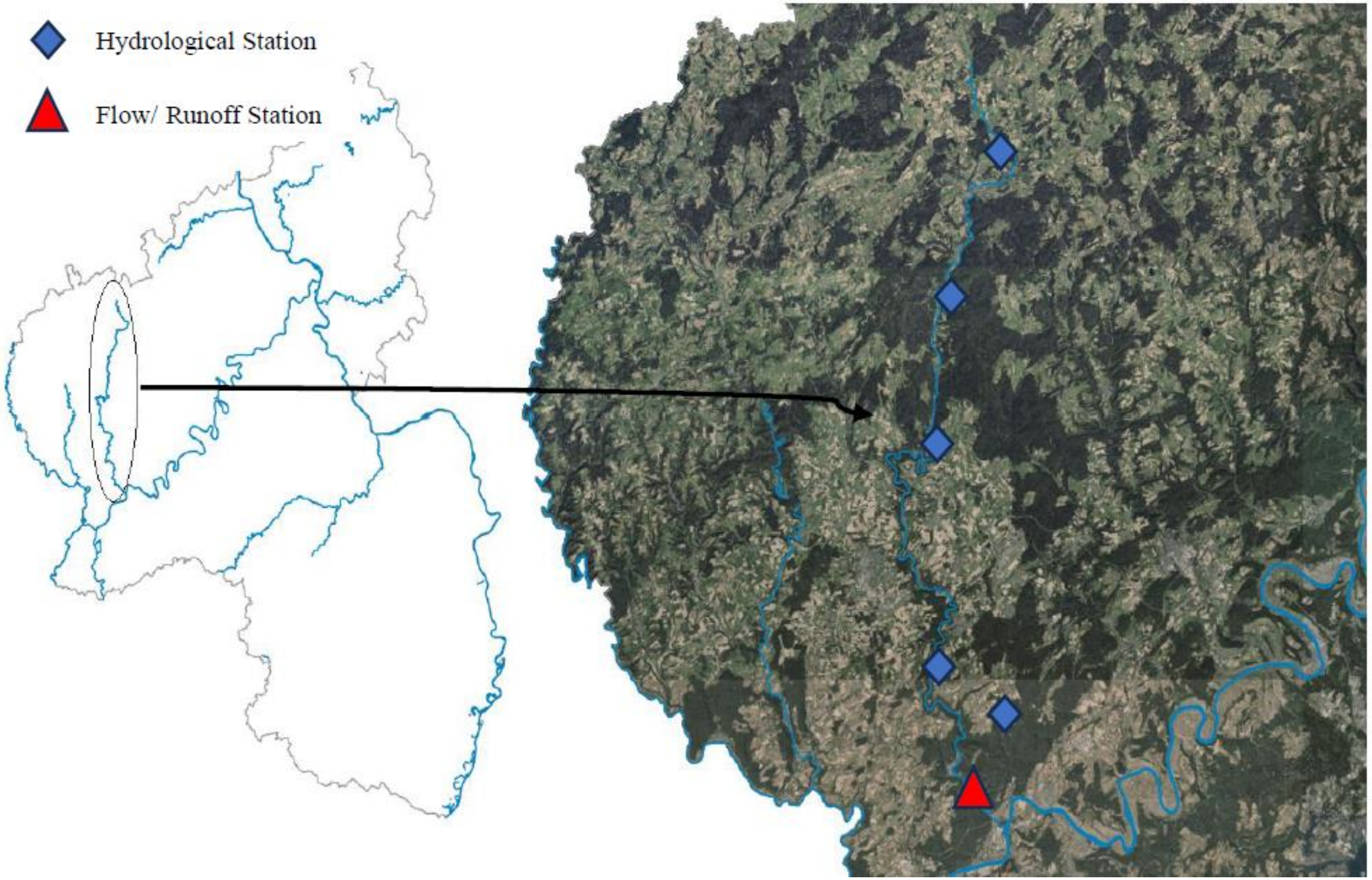
Prone to forecast inaccuracies because of extreme events.

Uncertainty Quantification (UQ)

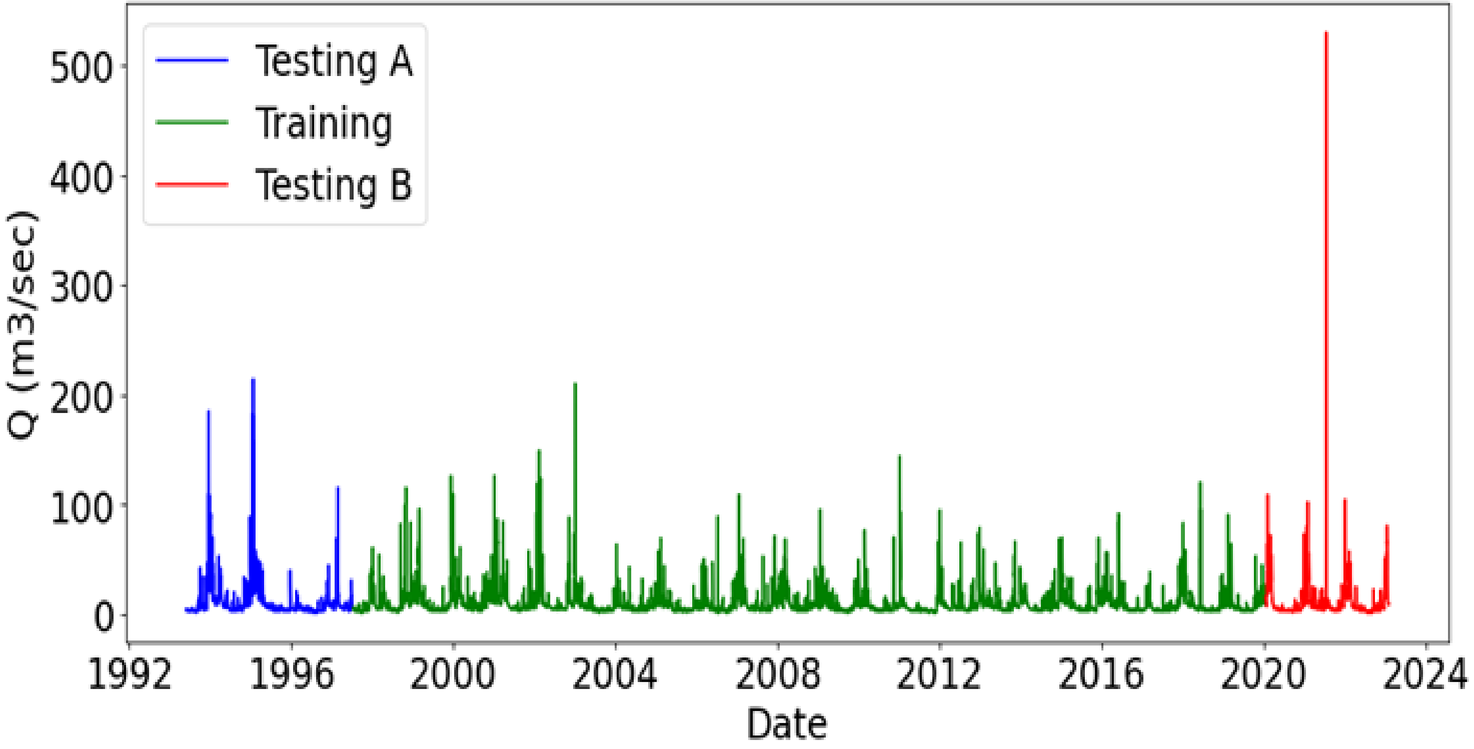
Assessing confidence and reliability in deep learning predictions

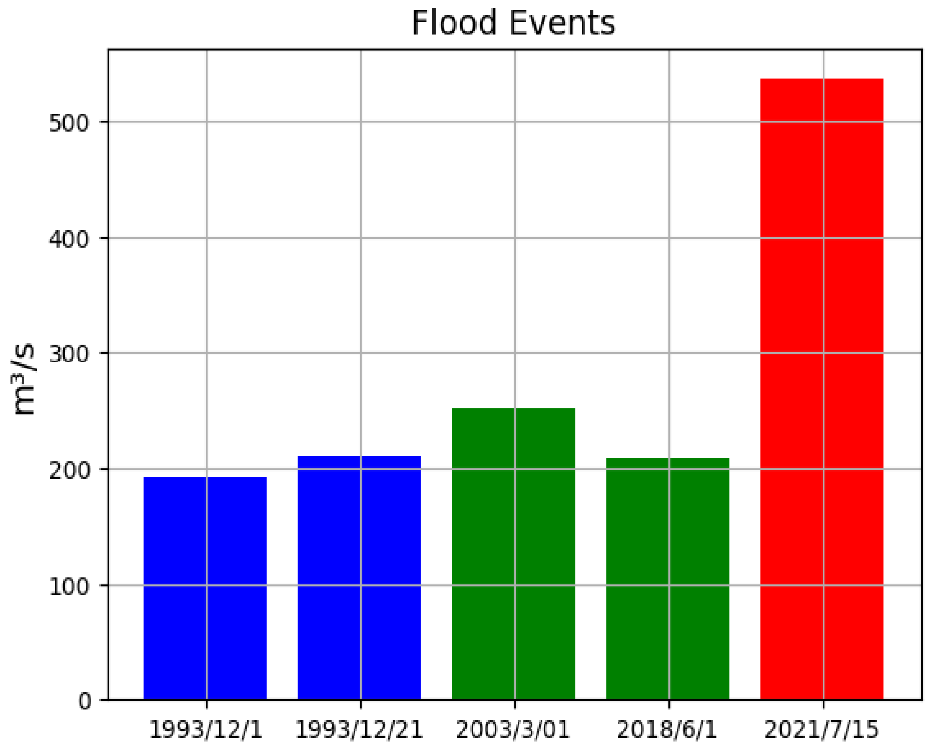
The Kyll River watershed

- Data collected from **five** weather stations
- 4 measured quantities per station
  - precipitation
  - temperature
  - humidity
  - solar radiation
- The flow rate serving as the output variable



Flood Distribution Within the Dataset





Methodology

Deep Learning Models

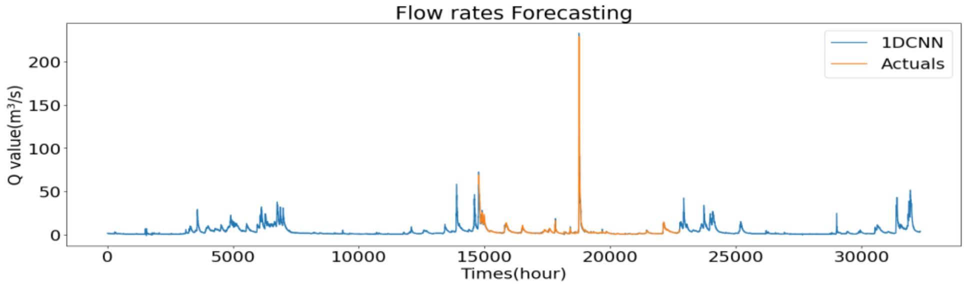
- BI-LSTM
- Conv-LSTM
- LSTM
- GRU
- 1DCNN

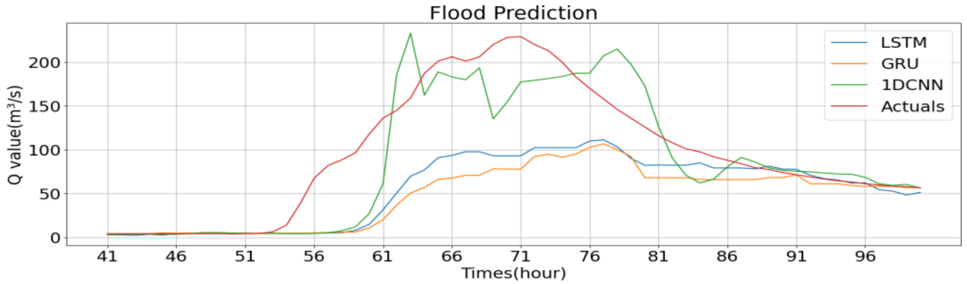
Flow rate

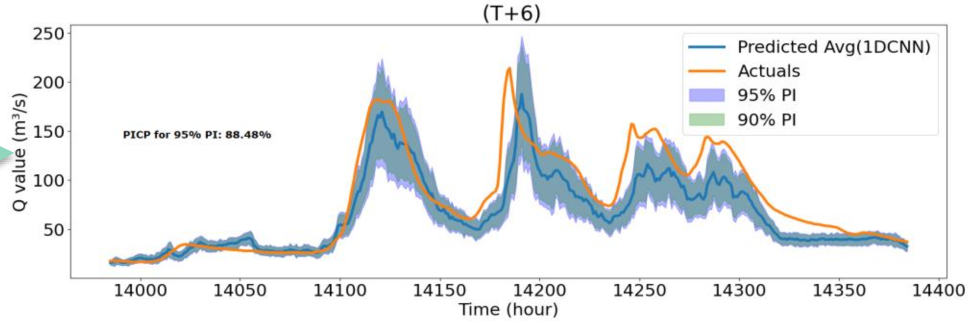
Weather data

Uncertainty Quantification

- MC Dropout
- Ensemble







The ensemble (UQ) approach combines models to predict outcomes and aggregates the average prediction from all models to estimate uncertainty.

The Monte Carlo Dropout (UQ) approach works by repeatedly applying dropout during model inference and averaging the predictions over multiple passes to estimate uncertainty.

Model Performance Metrics

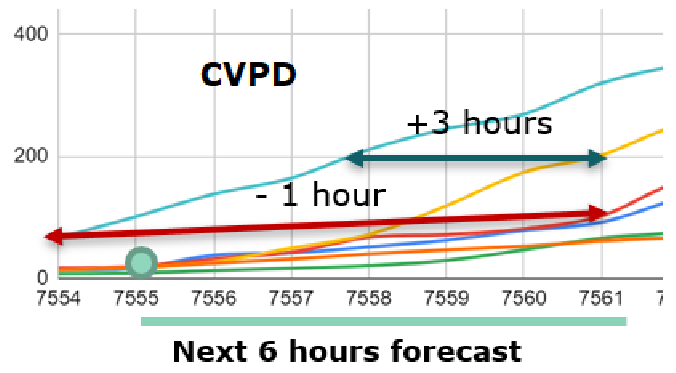
Root mean square error (RMSE) $= \sqrt{\frac{\sum_1^n (Q_i^T - Q_i^P)^2}{n}}$

Mean Absolute Error (MAE) $= \frac{1}{n}\sum_1^n \left|Q_i^T - Q_i^P\right|$

Coefficient of determination (R^2) $= 1 - \frac{\sum_1^n (Q_i^T - Q_i^p)^2}{\sum_1^n (Q_i^T - \overline{Q^T})^2}$

Flood peak value difference (FPVD)$= \max(Q_i^P) - \max(Q_i^T)$

Critical Value prediction delay (CVPD) $= h - (h_{cv}^T - h_{cv}^P)$.



Prediction Interval Coverage Probability (PICP) is defined as the ratio of samples that fall within their respective Prediction Intervals(PI).

Evaluation over the entire time course

Dataset B (The Entire Time Course)						
	NEXT 3 HOURS			NEXT 6 HOURS		
Model	**RSMA**	**MAE**	**R2**	**RSMA**	**MAE**	**R2**
LSTM	5.46	0.97	0.88	6.84	1.2	0.82
GRU	5.57	1.08	0.88	7.06	1.28	0.81
BiLSTM	8.05	1.58	0.75	8.07	1.93	0.75
1DCNN	2.51	0.57	0.98	3.65	0.89	0.95
ConvLSTM	10.23	3.49	0.59	10.68	3.71	0.57

	Dataset A (The Entire Time Course)					
	NEXT 3 HOURS			NEXT 6 HOURS		
MODELS	**RSMA**	**MAE**	**R2**	**RSMA**	**MAE**	**R2**
LSTM	3.04	1.06	0.96	3.65	1.26	0.95
GRU	3.58	1.22	0.95	4.03	1.3	0.94
BiLSTM	4.07	1.73	0.94	3.99	1.93	0.94
1DCNN	1.68	0.62	0.99	2.62	0.98	0.97
ConvLSTM	7.45	3.53	0.75	7.9	4.08	0.76

Evaluation of Forecasting During Flood Periods (Accuracy evaluation)

- Our study reveals ineffectiveness of LSTM, BI-LSTM, ConvLSTM, and GRU models in predicting flow rates during the extreme 2021 flood.

- 1DCNN shows a good level of performance across all prediction horizons.

During The 2021 Extreme Flood Period						
	NEXT 3 HOURS			NEXT 6 HOURS		
MODELS	RSMA	MAE	R2	RSMA	MAE	R2
LSTM	120.78	85.83	0.37	151.85	113.34	0.09
GRU	121.49	88.85	0.41	155.45	116.5	0.05
BiLSTM	179.45	128.94	-0.25	175	128.59	-0.19
1DCNN	52.82	34.59	0.89	75.17	51.83	0.78
ConvLSTM	205.21	155.55	-0.66	217.62	165.72	-0.84

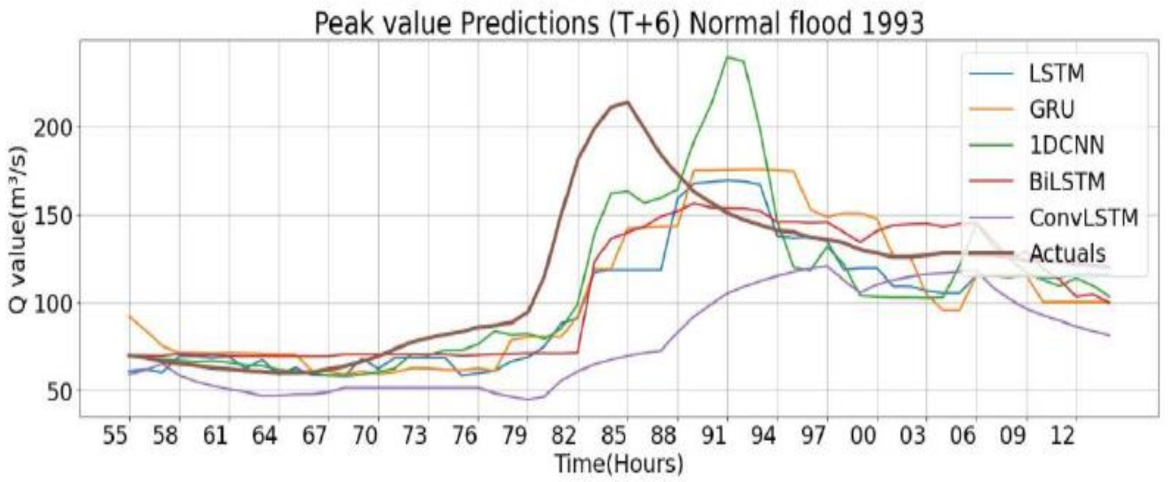
During The 1993 Flood Period						
	NEXT 3 HOURS			NEXT 6 HOURS		
MODELS	RSMA	MAE	R2	RSMA	MAE	R2
LSTM	29.34	20.53	0.5	37.24	24.83	0.23
GRU	35.07	22.75	0.3	38.57	27.02	0.16
BiLSTM	34.67	22.83	0.33	30.14	16.87	0.5
1DCNN	21.32	12.09	0.75	30.8	19.13	0.48
ConvLSTM	57.68	45.48	-1.08	60.93	46.01	-1.06

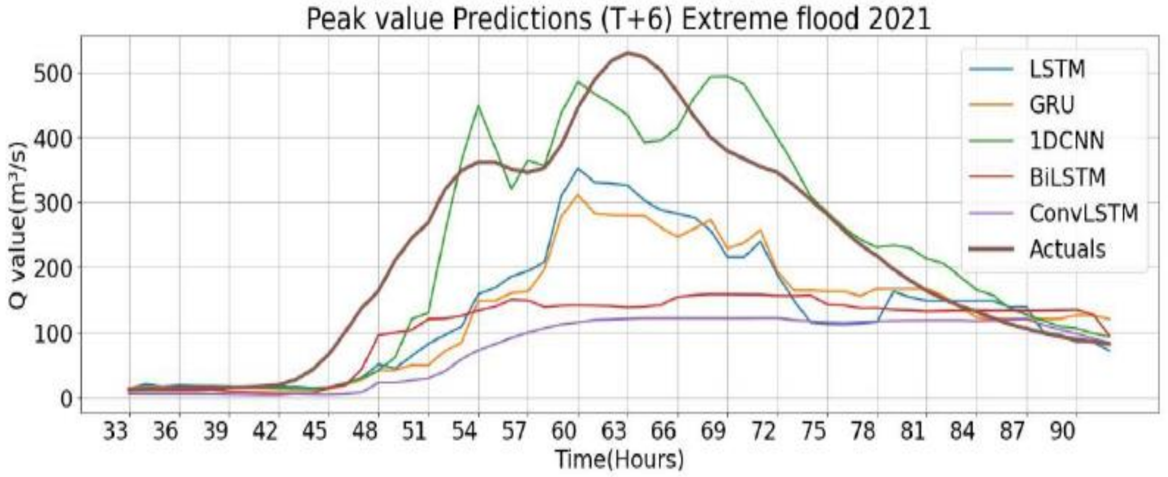
Evaluation of Forecasting During Flood Periods (Timing evaluation)

- Late prediction (negative values)
- Early prediction (positive values)

During the 1993 flood period (6 Hours)			
Model	**V150 m³/s**	**V250 m³/s**	**V350 m³/s**
LSTM	1.12±(0.96)	**-4.13±(4.81)**	**-∞**
GRU	1.1±(1.8)	**-4.7±(5.5)**	**-∞**
BiLSTM	2.8±(0.2)	0±(1.3)	**-3.9±(3.6)**
1DCNN	3.37±(0.83)	0.67±(2.17)	**-2.96±(3.69)**
ConvLSTM	**-∞**	**-∞**	**-∞**

During the 2021 flood period (6 Hours)				
Model	**V150 m³/s**	**V250 m³/s**	**V350 m³/s**	**V500 m³/s**
LSTM	**-1.9±(1.76)**	**-5.11±(3.3)**	**-8.33±(4.8)**	**-∞**
GRU	**-3±(2.8)**	**-6.9±(5.3)**	**-∞**	**-∞**
BiLSTM	**-6.6±(0.9)**	**-∞**	**-∞**	**-∞**
1DCNN	3.37±(0.39)	4.42±(1)	5.46±(1.64)	7.02±(2.61)
ConvLSTM	**-∞**	**-∞**	**-∞**	**-∞**





Analysis of false positive errors

- Only one False Positive case appears for 1DCNN in each forecast horizon (**black plot**)

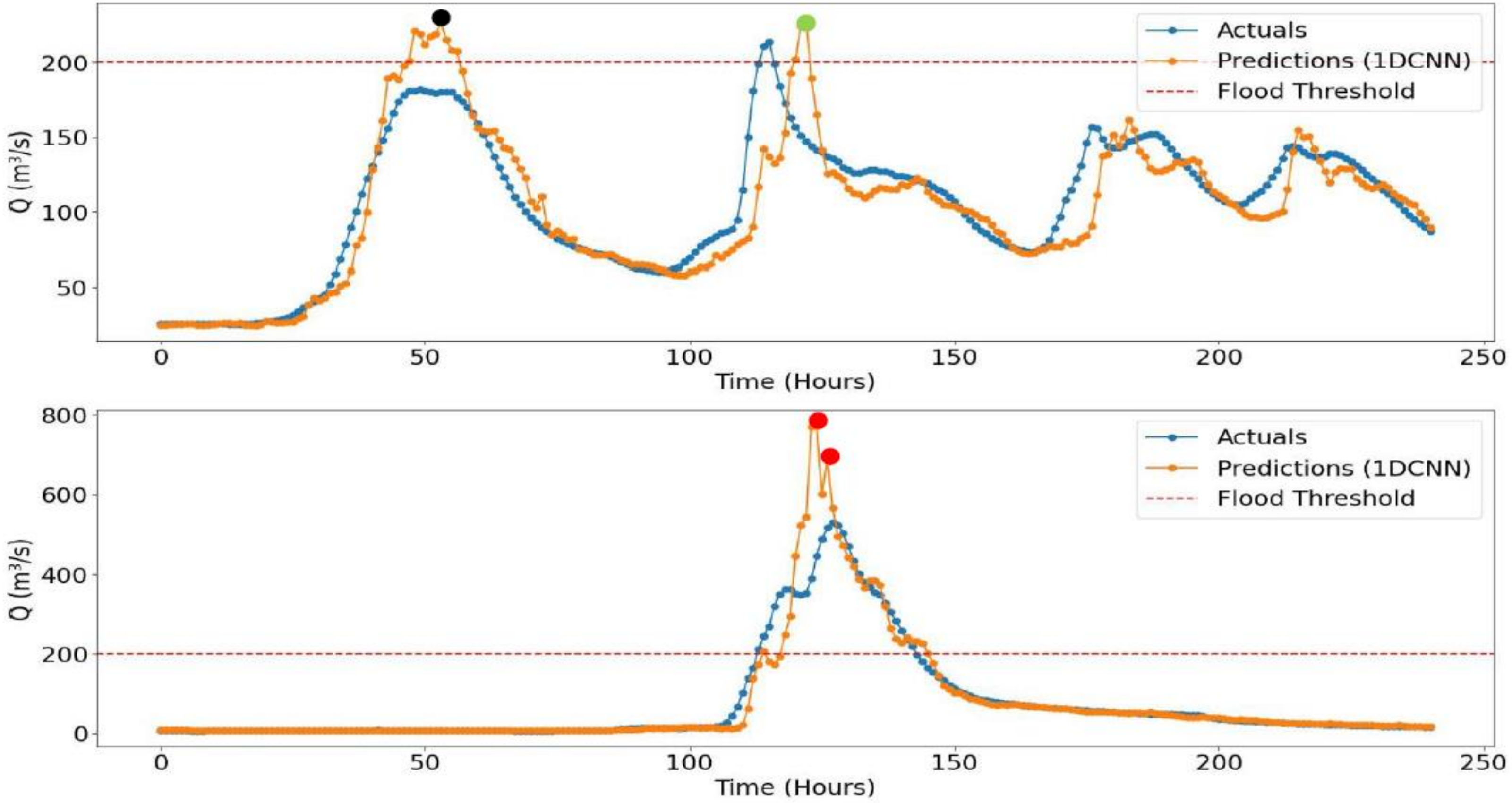


Figure 11: Analysis of false positive patterns of the 1DCNN model with 6-hour Forecast Horizon on both test sets

Evaluation of Uncertainty quantification (Ensemble)

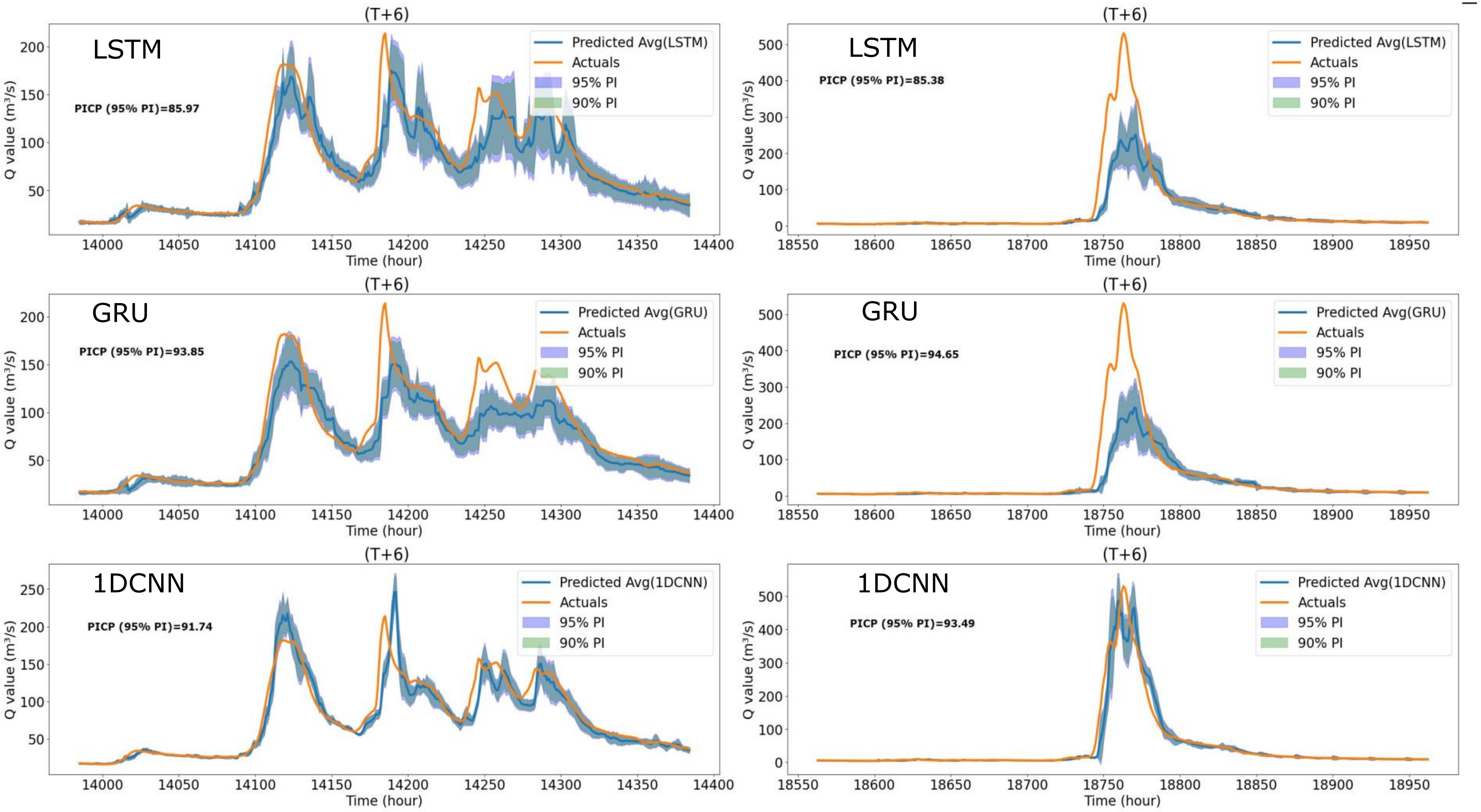
Ensemble + LSTM, GRU:

UQ for OOD points seems not trustworthy

Ensemble method:

- Effectively captures predictive uncertainty
- Informative uncertainty bounds

Normal Flood 1993 | Extreme Flood 2021



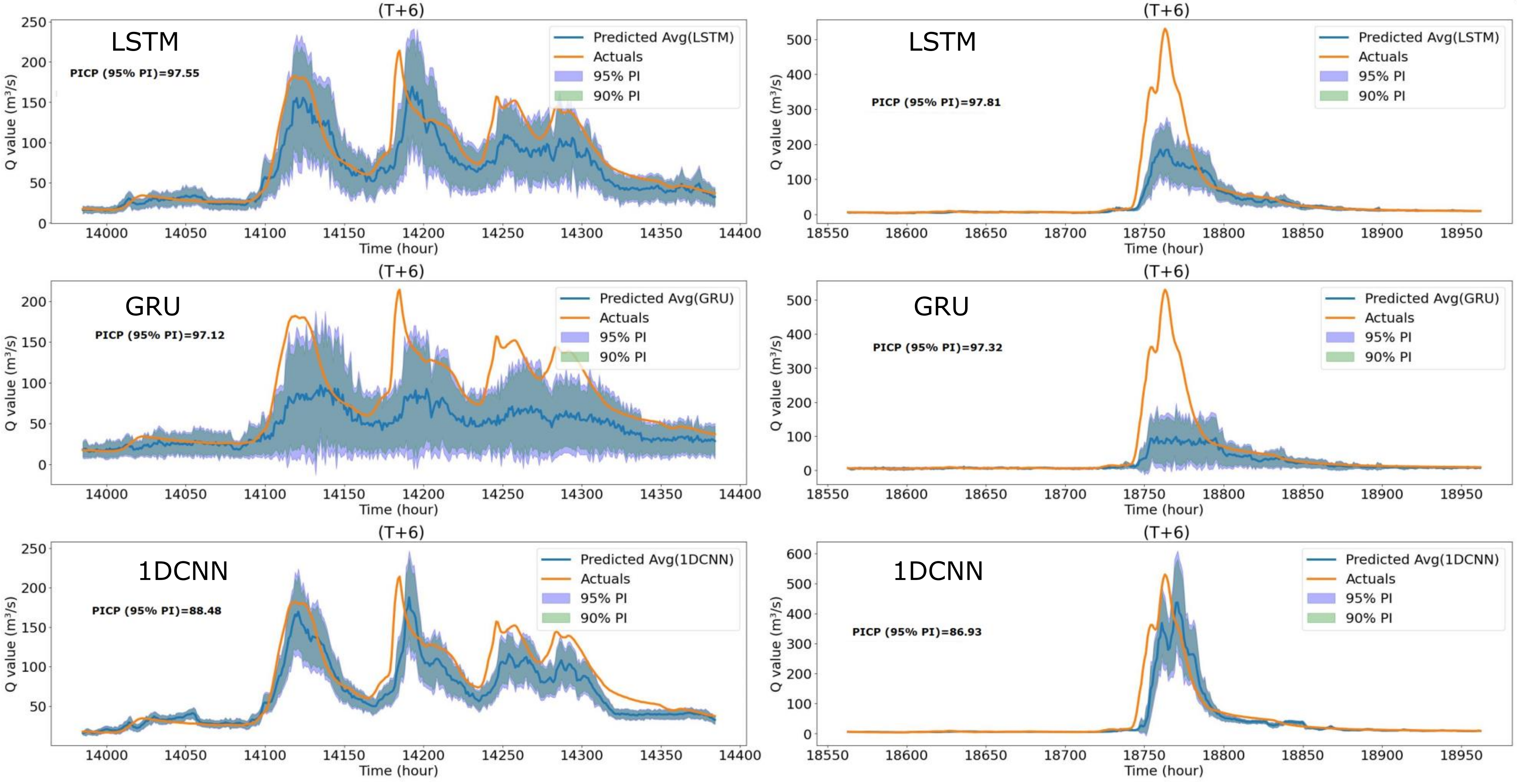
PICP is defined as the ratio of samples that fall within their respective PIs.

Evaluation of uncertainty quantification (Monte Carlo Dropout)

MCD:

Although the PICP values are high, MCD is uninformative as its **uncertainty bounds** are **too wide.**

Normal Flood 1993 | Extreme Flood 2021



PICP is defined as the ratio of samples that fall within their respective PIs.

Conclusion

This study focus:

- Analysing deep learning methods for small river streamflow forecasting.
- Special focus on extreme flood prediction.
- Investigating uncertainty estimation methods with deep learning models.

Aspects Analyzed:

- Prediction accuracy
- Timely prediction
- Quality and robustness of prediction

Overall Findings:

- All models give adequate forecasts **on average** over the entire time course.
- 1DCNN model outperformed other models.
- Predicting extreme floods is a challenges for most models.

Future Work

- Models show errors (accuracy and time errors) in peak flood forecast, *which cannot be ignored*.

- Rooms for future work :
  - More investigation in Uncertainty Quantification
  - Causal effect analysis
  - Designing balanced datasets
  - Cross-domain for flood prediction

References

- Cao, Q., Zhang, H., Zhu, F., Hao, Z., & Yuan, F. (2022). Multi-step-ahead flood forecasting using an improved BiLSTM-S2S model. Journal of Flood Risk Management, 15(4). https://doi.org/10.1111/jfr3.12827
- Chaudhary, P., Leitão, J. P., Donauer, T., D'Aronco, S., Perraudin, N., Obozinski, G., Perez-Cruz, F., Schindler, K., Wegner, J. D., & Russo, S. (2022). Flood Uncertainty Estimation Using Deep Ensembles. Water (Switzerland), 14(19). https://doi.org/10.3390/w14192980
- Donnelly, P. J., & Junkins, O. (2022). Short-Term River Forecasting with a Stacked Ensemble of Tributary Models. 2022 7th International Conference on Frontiers of Signal Processing (ICFSP), 189–193.
- Feng, J., Yan, L., & Hang, T. (2019). Stream-flow forecasting based on dynamic spatio-temporal attention. IEEE Access, 7, 134754–134762. https://doi.org/10.1109/ACCESS.2019.2941799
- Ghobadi, F., & Kang, D. (2022). Improving long-term streamflow prediction in a poorly gauged basin using geo-spatiotemporal mesoscale data and attention-based deep learning: A comparative study. Journal of Hydrology, 615, 128608.
- Gude, V., Corns, S., & Long, S. (2020). Flood Prediction and Uncertainty Estimation Using Deep Learning. Water (Switzerland), 12(3). https://doi.org/10.3390/w12030884
- Hu, Y., Yan, L., Hang, T., & Feng, J. (2020). Stream-Flow Forecasting of Small Rivers Based on LSTM. http://arxiv.org/abs/2001.05681
- Ilhan, A. (2022). Forecasting of river water flow rate with machine learning. Neural Computing and Applications, 34(22), 20341–20363. https://doi.org/10.1007/s00521-022-07576-9
- Kim, D., Lee, J., Kim, J., Lee, M., Wang, W., & Kim, H. S. (2022). Comparative analysis of long short-term memory and storage function model for flood water level forecasting of Bokha stream in NamHan River, Korea. Journal of Hydrology, 606, 127415.

References

- Klotz, D., Kratzert, F., Gauch, M., Keefe Sampson, A., Brandstetter, J., Klambauer, G., Hochreiter, S., & Nearing, G. (2022). Uncertainty estimation with deep learning for rainfall-runoff modeling. Hydrology and Earth System Sciences, 26(6), 1673–1693. https://doi.org/10.5194/hess-26-1673-2022
- Le, X. H., Nguyen, D. H., Jung, S., Yeon, M., & Lee, G. (2021). Comparison of Deep Learning Techniques for River Streamflow Forecasting. IEEE Access, 9, 71805–71820. https://doi.org/10.1109/ACCESS.2021.3077703
- Li, X., Xu, W., Ren, M., Jiang, Y., & Fu, G. (2022). Hybrid CNN-LSTM models for river flow prediction. Water Supply, 22(5), 4902–4920. https://doi.org/10.2166/ws.2022.170
- Liu, S., Lu, D., Painter, S. L., Griffiths, N. A., & Pierce, E. M. (2023). Uncertainty quantification of machine learning models to improve streamflow prediction under changing climate and environmental conditions. Frontiers in Water, 5. https://doi.org/10.3389/frwa.2023.1150126
- Moazam, H. M. Z. H., Dehghani, A., Mortazavizadeh, F., Ranjbar, V., Mirzaei, M., Hin, L. S., Ng, J. L., & Dehghani, A. (2023). Comparative evaluation of LSTM, CNN, and ConvLSTM for hourly short-term streamflow forecasting using deep learning approaches. Ecological Informatics, 102119.
- Rathinasamy, M., Adamowski, J., & Khosa, R. (2013). Multiscale streamflow forecasting using a new Bayesian Model Average based ensemble multi-wavelet Volterra nonlinear method. Journal of Hydrology, 507, 186–200.
- Wegayehu, E. B., & Muluneh, F. B. (2022). Short-Term Daily Univariate Streamflow Forecasting Using Deep Learning Models. Advances in Meteorology, 2022. https://doi.org/10.1155/2022/1860460
- Wu, J., Wang, Z., Hu, Y., Tao, S., & Dong, J. (2023). Runoff Forecasting using Convolutional Neural Networks and optimized Bi-directional Long Short-term Memory. Water Resources Management, 1–17.
- Xu, W., Jiang, Y., Zhang, X., Li, Y., Zhang, R., & Fu, G. (2020). Using long short-term memory networks for river flow prediction. Hydrology Research, 51(6), 1358–1376. https://doi.org/10.2166/nh.2020.026

References

- Yan, L., Feng, J., & Hang, T. (2019). Small watershed stream-flow forecasting based on LSTM. Advances in Intelligent Systems and Computing, 935, 1006–1014. https://doi.org/10.1007/978-3-030-19063-7_79
- Zhu, Y., Feng, J., Yan, L., Guo, T., & Li, X. (2020). Flood Prediction Using Rainfall-Flow Pattern in Data-Sparse Watersheds. IEEE Access, 8, 39713–39724. https://doi.org/10.1109/ACCESS.2020.2971264
- Zou, Y., Wang, J., Lei, P., & Li, Y. (2023). A novel multi-step ahead forecasting model for flood based on time residual LSTM. Journal of Hydrology, 620, 129521.

Thanks for listening.
Any questions?