TC-GTN: Temporal Convolution Graph Transformer Network for Hydrological Forecasting

Ana Samac
ana.samac@ivi.ac.rs

Milan Dotlic
milan.dotlic@ivi.ac.rs

Luka Vinokic
luka.vinokic@ivi.ac.rs

Milan Stojkovic
milan.stojkovic@ivi.ac.rs

Veljko Prodanovic
veljko.prodanovic@ivi.ac.rs

The Institute for Artificial Intelligence Research and Development of Serbia
21000, Novi Sad, Serbia

Abstract

Machine learning enables accurate streamflow forecasting, vital for managing increasingly frequent flood events under climate change. However, most existing approaches do not fully exploit the inherent directional and hierarchical graph structure of hydrological systems. This paper introduces TC-GTN (Temporal Convolution Graph Transformer Network), a hybrid model designed for streamflow forecasting that integrates temporal convolution (TC) with graph transformers (GT). It uses the combination of TC for temporal pattern extraction and GT for advanced relational reasoning. It utilizes a structured graph representation of the river network with accompanying meteorological stations where the transformer's attention mechanism is critical for a better understanding of interactions between different nodes/stations and for capturing self-dependencies within each station. Experiments on the Drina–Lim River Basin dataset show that TC-GTN model outperforms baseline methods for regular flow rates, and also demonstrate improvements for high flow rates, which represent extreme hydrological events. Such performance is critical for effective flood risk mitigation and sustainable hydropower management under climate change effects. Code is available at: https://github.com/dodi007/TC-GTN-Spatio-temporal-Graph-Transormer.git.

1 Introduction and related work

Reliable streamflow forecasting and early warning systems are essential for modern water resource management, especially in the context of climate change, where extreme hydrological events such as floods pose severe risks to lives, infrastructure, and ecosystems (Martinez-Villalobos and Neelin, 2023; Rodell and Li, 2023). Beyond flood risk mitigation, streamflow forecasting supports more effective planning and real-time management of hydropower plant operations by enabling optimal water use, grid stability, and greater renewable energy integration. Thus, improved streamflow forecasting not only supports disaster preparedness but also plays a key role in the sustainable production and management of green energy (Rolnick et al., 2022).

Machine learning has transformed time series forecasting, enabling models to capture complex temporal patterns more effectively than traditional physics-based models. Common architectures

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2025.

include recurrent neural networks (RNNs) (Connor et al., 1994; Chang et al., 2014), temporal convolutional networks (TCNs) (Wan et al., 2019; Xu et al., 2021), graph neural networks (GNNs) (Farahmand et al., 2023; Jin et al., 2024), and transformers (Nie et al., 2022; Wen et al., 2022). Recent studies (Granata et al., 2024; Koya and Roy, 2024; Vinokić et al., 2025) show that TCNs outperform LSTMs and Temporal KAN models (Vinokić et al., 2025), while transformer-based architectures with attention mechanisms (Granata et al., 2024) and temporal fusion(Koya and Roy, 2024) achieve superb performance, especially for long-term predictions.

For streamflow forecasting in river basins with both temporal and spatial patterns, GNNs and their hybrids with temporal and transformers architecture have gained prominence (Farahmand et al., 2023; Ng et al., 2023; Roudbari et al., 2024; Zhang et al., 2024). Studies have shown that integrating spatial and temporal models outperforms those that do not consider both spatial and temporal attention, especially for flood forecasting (Feng et al., 2019; Ding et al., 2020; Liu et al., 2021). Combining GNNs with transformers and attention mechanisms further enhances spatio-temporal dynamics and latent causal modeling (Shi et al., 2020, 2023; Jiang et al., 2024). Representative examples, such as TFM-GCAM (Chen et al., 2024), Trafformer (Jin et al., 2023), and STGA-Former (Geng et al., 2024), capture complex spatio-temporal dependencies and achieve state-of-the-art performance.

However, most existing approaches do not fully exploit the inherent directional and hierarchical graph structure of hydrological systems that combine meteorological stations, which measures precipitation and temperature with hydrological stations and form distinct types of nodes and edges. Further, usually proposed spatio-temporal forecasting models combine either GNN with temporal machine learning model or GNN with transformer. This paper proposes a novel TC-GTN (Temporal Convolution Graph Transformer Network) architecture, a unified spatio-temporal model that addresses whole hydrological system with meteorological measurements for streamflow forecasting and combines graph transformer with additional temporal convolution to further improve the extraction of temporal patterns. Key contributions are: 1) A structured graph representation explicitly modeling three hydrological relationships via a static adjacency matrix for domain-aware message passing. 2) A hybrid architecture combining the temporal convolution with relational reasoning and positional encoding strengths of graph transformers and with residual connections to jointly learn temporal and spatial patterns. 3) A quantile loss targeting high-flow events critical for flood prediction.

2 Methodology

2.1 Graph structure

The study area is represented as graph $G = (V, E)$, where nodes V are meteorological and hydrological stations, and edges E encode spatial and directional relationships. The spatial connectivity between nodes is encoded in a static adjacency matrix $A \in \{0,1\}^{V \times V}$, where $A_{ij} = 1$ indicates a link from node i to node j. Three different types of edges are defined to represent spatial and directional relationships in the network. **Meteo-Meteo** undirected edges connect meteorological stations to capture spatial correlations in weather patterns. **Meteo-Hydro** directed edges link meteorological stations to hydrological stations, representing the influence of precipitation and temperature on discharge. **Hydro-Hydro** directed edges connect hydrological stations according to the direction of river flow, capture the downstream propagation of discharge. This graph structure enables explicit modeling of spatial dependencies.

2.2 TC-GTN model architecture

The architecture of TC-GTN model consists of three main components: a temporal encoder based on 1D convolution, a graph transformer module based on the graph transformer operator, and a feedforward decoder.

Temporal encoder processes the input tensor. Each node's time series is processed independently using a sequence of 1D convolutional layers followed by a max pooling operation and a linear projection that transforms the output of the temporal encoder into a shape suitable for input to the graph transformer.

Graph transformer models spatial dependencies using a stack of Graph Transformer layers (Shi et al., 2020). For each node and its neighbors, the model learns to represent their features as queries and keys through trainable transformations. These representations enable the computation of attention scores that measure the importance of neighbors relative to the node in question. The model calculates attention weights that quantify how much influence each neighboring node should have when updating a node's representation. Instead of using a single attention mechanism, multiple "heads" are used in parallel to capture different types of relationship or aspects of the data. The multi-head attention score between nodes for each edge of every head is calculated and normalized with softmax function. The contributions from these heads are then combined by averaging to maintain consistent feature dimensions throughout the layers.

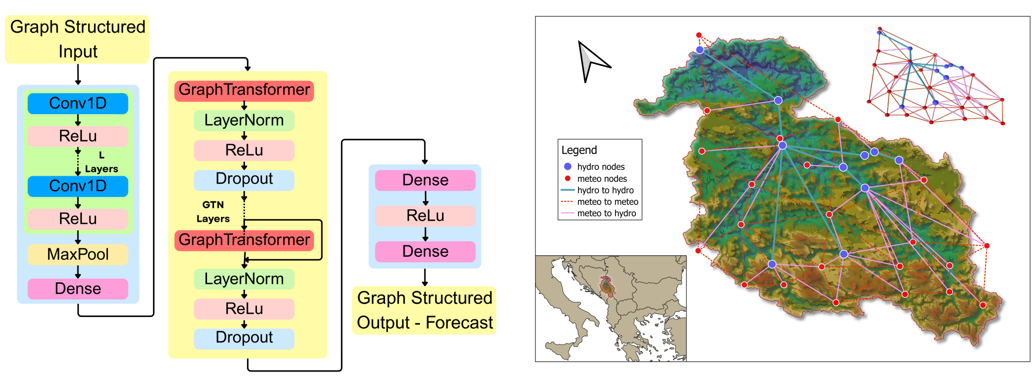
Decoder maps the learned spatio-temporal node representations into the target prediction space, i.e., forecasting the future values for each node. It consists of a two-layer fully connected feedforward neural network with a ReLU activation between. This design allows the model to learn a non-linear mapping from the high-dimensional node embeddings to the desired forecast horizon.

The complete architecture of the proposed model is shown in Figure 1a.

3 Experimental setup

3.1 Dataset: Drina River basin

The data set consists of daily time series data collected from 10 hydrological stations and 22 meteorological stations located in the Drina River basin in Southeastern Europe. Each meteorological station records the mean daily temperature $[°C]$ and the total daily precipitation $[mm]$, while each hydrological station measures the daily flow of the river $[m^3/s]$. This results in a total of $V = 22 \times 2 + 10 = 54$ nodes, accounting for two variables per meteorological station. The location and layout of the Drina River basin, along with its corresponding graph structure, are shown in Figure 1b. The temporal input sequence has a length of 7 days. For the hydro nodes, it consists of the previous 7 days and for the meteo nodes that sequence has values for the 2 previous days and 5 future day values, resulting in a total sequence length of 7 days. The dataset spans the period from 1968 to 2018 and is divided into training (1968–2015 for a total of 17525 samples), validation (2016 for a total of 366 samples), and test sets (2017-2018 for a total of 725 samples). All input data are normalized using a Min-Max scaler to avoid extreme values.



(a) TC-GTN architecture

(b) Drina River basin

Figure 1: a) Architecture of the TC-GTN model and b) Graph representation of the Drina River Basin: blue nodes denote hydrological stations, red nodes represent meteorological stations, and edges (varying by type) indicate connections.

3.2 Baselines

To evaluate the effectiveness of the proposed model architecture, we conducted a comparative analysis against two baseline models: (1) a Graph Transformer model without any temporal encoder (GTN model) and (2) a Graph Transformer combined with an LSTM temporal encoder (LS-GTN model). In all experiments, identical settings for the Graph Transformer component are maintained. This ensured a fair comparison by isolating the impact of different temporal encoders.

To comprehensively assess the performance of our TC-GTN forecasting model, we employed four commonly used regression evaluation metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared Score (R^2), and Mean Absolute Percentage Error (MAPE). Given the nature of our application, forecasting river discharge levels with a particular emphasis on flood risk, we also compute these metrics on values belonging to the fourth quantile (values above the 75th percentile). This filtering ensures that the evaluation focuses on the most critical events where accurate predictions are essential for timely and effective flood management.

All of the models are trained for up to 500 epochs using the Adam optimizer with early stopping applied. For the loss function we used the quantile loss which proved effective in targeting flood events, often treated as outliers by traditional loss functions like MSE.

4 Results & discussion

Our proposed model, TC-GTN, consistently outperforms the baseline methods (GTN and LS-GTN) across most stations and prediction intervals. Even in scenarios where LS-GTN shows marginally better performance on the first or second prediction day, TC-GTN excels on longer horizons, highlighting the benefit of adding temporal convolutional layers to enhance feature extraction beyond GTN's graph transformer mechanisms. This supports recent findings (Zeng et al., 2023) that transformers alone are less effective for time series forecasting and shows that integrating time series–specific modules significantly boosts performance. On average, TC-GTN achieves an error reduction of approximately 11.63% in RMSE, 16% in MAE, 40.35% in MAPE, and an improvement of 21.02% in R^2 compared to the best performing baseline (See Table 1 in the Appendix for full comparison). This trend is particularly evident in low-flow stations where the baselines struggle. These stations (3,4 and 5) have flow levels around 100 m^3/s, compared to up to 3500 m^3/s at other stations, and pose challenges for GNNs alone due to the relatively small error margin, but not so much for TC-GTN. This aligns with findings that extracting distinct temporal features improves forecasting in complex time series and outperforms convolution-only or transformer-only approaches (Liu et al., 2022).

Furthermore, we also examined the performance specifically for flow rates above the 75th percentile, focusing on high-flow scenarios that increase flood risks. This results also show that TC-GTN achieves average error reduction of approximately 6.87% in RMSE, 8.83% in MAE, 10.39% in MAPE, and an improvement of 6.05% in R^2 compared to the best performing baseline (for full comparison see Table 2 in the Appendix). Figure 2 confirms that TC-GTN consistently outperforms baselines, especially during high-flow (95–99 percentile) events. With lower median APE, less variability, and fewer extreme errors, TC-GTN delivers more accurate and stable predictions.

Limitations of the proposed approach are the temporal resolution of the data. Although we used daily data, higher-frequency data (e.g., hourly) would likely improve the model's efficiency in capturing rapid streamflow changes. Another limitation is the need for expert knowledge in hydrology to model the graph structure and define connections, as the node linking process is not automated. Automating this process based on topological features would be a valuable next step, as each new case study currently requires a manually designed graph.

5 Conclusion

In this paper, we propose a novel TC-GTN model for streamflow forecasting that combines temporal convolutions, graph transformers, attention mechanism and residual connections, to jointly model spatial and temporal dependencies and improve extreme event prediction. Experiments show notable error reductions over baselines and strong performance on high-flow events, enabling reliable flood forecasting, efficient hydropower operation, and sustainable water management, capabilities that are increasingly vital as climate change intensifies extreme hydrological events.

6 Acknowledgments

This research received a support from the European Union's Horizon Europe project ARTIFACT, under Grant Agreement 101159480.

References

Chang, F.J., Chen, P.A., Lu, Y.R., Huang, E., Chang, K.Y., 2014. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. Journal of Hydrology 517, 836–846.

Chen, J., Zheng, L., Hu, Y., Wang, W., Zhang, H., Hu, X., 2024. Traffic flow matrix-based graph neural network with attention mechanism for traffic flow prediction. Information Fusion 104, 102146.

Connor, J.T., Martin, R.D., Atlas, L.E., 1994. Recurrent neural networks and robust time series prediction. IEEE transactions on neural networks 5, 240–254.

Ding, Y., Zhu, Y., Feng, J., Zhang, P., Cheng, Z., 2020. Interpretable spatio-temporal attention lstm model for flood forecasting. Neurocomputing 403, 348–359.

Farahmand, H., Xu, Y., Mostafavi, A., 2023. A spatial–temporal graph deep learning model for urban flood nowcasting leveraging heterogeneous community features. Scientific Reports 13, 6768.

Feng, J., Yan, L., Hang, T., 2019. Stream-flow forecasting based on dynamic spatio-temporal attention. IEEE Access 7, 134754–134762.

Geng, Z., Xu, J., Wu, R., Zhao, C., Wang, J., Li, Y., Zhang, C., 2024. Stgaformer: Spatial–temporal gated attention transformer based graph neural network for traffic flow forecasting. Information Fusion 105, 102228.

Granata, F., Zhu, S., Di Nunno, F., 2024. Advanced streamflow forecasting for central european rivers: the cutting-edge kolmogorov-arnold networks compared to transformers. Journal of Hydrology 645, 132175.

Jiang, M., Weng, B., Chen, J., Huang, T., Ye, F., You, L., 2024. Transformer-enhanced spatiotemporal neural network for post-processing of precipitation forecasts. Journal of Hydrology 630, 130720.

Jin, D., Shi, J., Wang, R., Li, Y., Huang, Y., Yang, Y.B., 2023. Trafformer: Unify time and space in traffic prediction, in: Proceedings of the AAAI conference on artificial intelligence, pp. 8114–8122.

Jin, M., Koh, H.Y., Wen, Q., Zambon, D., Alippi, C., Webb, G.I., King, I., Pan, S., 2024. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. IEEE Transactions on Pattern Analysis and Machine Intelligence .

Koya, S.R., Roy, T., 2024. Temporal fusion transformers for streamflow prediction: Value of combining attention with recurrence. Journal of Hydrology 637, 131301.

Liu, M., Zeng, A., Chen, M., Xu, Z., Lai, Q., Ma, L., Xu, Q., 2022. Scinet: Time series modeling and forecasting with sample convolution and interaction. Advances in Neural Information Processing Systems 35, 5816–5828.

Liu, Y., Zhang, T., Kang, A., Li, J., Lei, X., 2021. Research on runoff simulations using deep-learning methods. Sustainability 13, 1336.

Martinez-Villalobos, C., Neelin, J.D., 2023. Regionally high risk increase for precipitation extreme events under global warming. Scientific Reports 13, 5579.

Ng, K., Huang, Y., Koo, C., Chong, K., El-Shafie, A., Ahmed, A.N., 2023. A review of hybrid deep learning applications for streamflow forecasting. Journal of Hydrology 625, 130141.

Nie, Y., Nguyen, N.H., Sinthong, P., Kalagnanam, J., 2022. A time series is worth 64 words: Long-term forecasting with transformers. arXiv preprint arXiv:2211.14730 .

Rodell, M., Li, B., 2023. Changing intensity of hydroclimatic extreme events revealed by grace and grace-fo. Nature Water 1, 241–248.

Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A.S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A.S., Maharaj, T., Sherwin, E.D., Mukkavilli, S.K., Kording, K.P., Gomes, C.P., Ng, A.Y., Hassabis, D., Platt, J.C., Creutzig, F., Chayes, J., Bengio, Y., 2022. Tackling climate change with machine learning. ACM Comput. Surv. 55. URL: `https://doi.org/10.1145/3485128`, doi:`10.1145/3485128`.

Roudbari, N.S., Punekar, S.R., Patterson, Z., Eicker, U., Poullis, C., 2024. From data to action in flood forecasting leveraging graph neural networks and digital twin visualization. Scientific reports 14, 18571.

Shi, J., Stebliankin, V., Wang, Z., Wang, S., Narasimhan, G., 2023. Graph transformer network for flood forecasting with heterogeneous covariates. arXiv preprint arXiv:2310.07631 .

Shi, Y., Huang, Z., Feng, S., Zhong, H., Wang, W., Sun, Y., 2020. Masked label prediction: Unified message passing model for semi-supervised classification. arXiv preprint arXiv:2009.03509 .

Vinokić, L., Dotlić, M., Prodanović, V., Kolaković, S., Simonovic, S.P., Stojković, M., 2025. Effectiveness of three machine learning models for prediction of daily streamflow and uncertainty assessment. Water Research X 27, 100297.

Wan, R., Mei, S., Wang, J., Liu, M., Yang, F., 2019. Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting. Electronics 8, 876.

Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., Sun, L., 2022. Transformers in time series: A survey. arXiv preprint arXiv:2202.07125 .

Xu, Y., Hu, C., Wu, Q., Li, Z., Jian, S., Chen, Y., 2021. Application of temporal convolutional network for flood forecasting. Hydrology Research 52, 1455–1468.

Zeng, A., Chen, M., Zhang, L., Xu, Q., 2023. Are transformers effective for time series forecasting?, in: Proceedings of the AAAI conference on artificial intelligence, pp. 11121–11128.

Zhang, Z., Tian, W., Lu, C., Liao, Z., Yuan, Z., 2024. Graph neural network-based surrogate modelling for real-time hydraulic prediction of urban drainage networks. Water Research 263, 122142.

A Technical Appendices and Supplementary Material

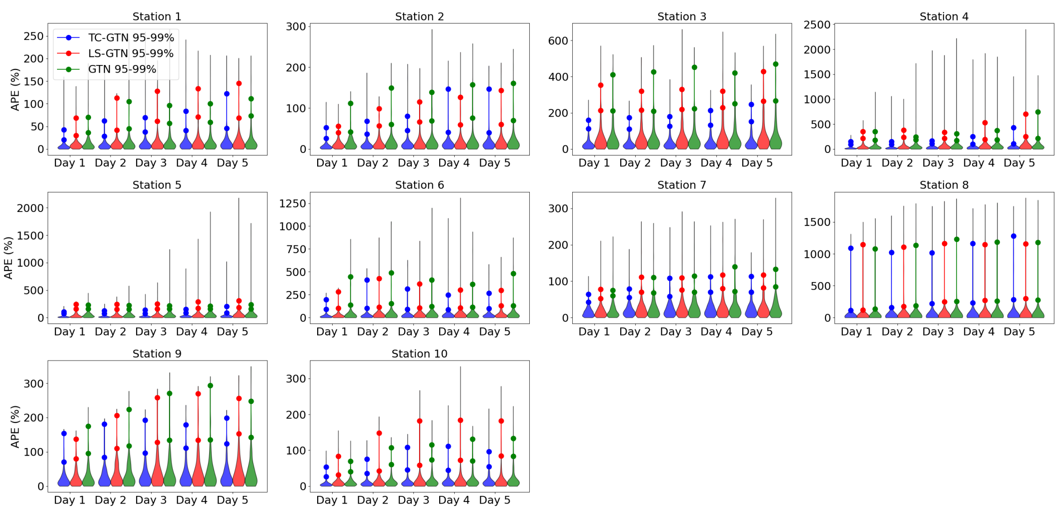


Figure 2: Comparison of APE distributions for different models (TC-GTN, LS-GTN, GTN) over a five-day forecast period across ten hydrological stations. The 95-99 percentile range, indicating high streamflow events, is highlighted.

Table 1: Model comparison (GTN, LS-GTN and TC-GTN) for the streamflow forecasting across multiple metrics.

Station	**Day**	**MAE**			**RMSE**			**R^2**			**MAPE**		
		GTN	LS-GTN	TC-GTN	GTN	LS-GTN	TC-GTN	GTN	LS-GTN	TC-GTN	GTN	LS-GTN	TC-GTN
Station 1	1	7.84	7.37	**5.75**	16.12	16.35	**13.92**	0.949	0.947	**0.962**	0.112	0.109	**0.070**
	2	10.27	9.36	**8.28**	21.82	**19.91**	19.99	0.906	**0.921**	**0.921**	0.137	0.131	**0.102**
	3	11.59	10.43	**9.85**	24.21	21.21	**20.97**	0.884	0.911	**0.913**	0.157	0.148	**0.126**
	4	11.86	11.42	**10.79**	23.38	22.60	**21.78**	0.892	0.899	**0.906**	0.164	0.160	**0.141**
	5	12.40	12.14	**11.32**	23.52	22.83	**22.25**	0.890	0.897	**0.902**	0.176	0.170	**0.150**
Station 2	1	8.02	7.26	**5.46**	15.36	13.83	**11.91**	0.956	0.964	**0.974**	0.129	0.113	**0.067**
	2	10.59	9.94	**8.02**	21.62	20.08	**18.12**	0.913	0.925	**0.939**	0.148	0.133	**0.095**
	3	12.11	11.10	**9.88**	24.30	22.10	**19.88**	0.890	0.909	**0.926**	0.161	0.147	**0.121**
	4	13.02	12.44	**10.76**	24.64	24.27	**21.27**	0.887	0.890	**0.916**	0.174	0.159	**0.135**
	5	14.39	13.89	**11.41**	27.25	26.77	**22.85**	0.861	0.866	**0.903**	0.191	0.172	**0.145**
Station 3	1	3.35	3.16	**1.84**	6.32	4.99	**3.49**	0.599	0.751	**0.878**	0.704	0.697	**0.346**
	2	3.98	3.89	**2.41**	8.55	7.17	**5.12**	0.268	0.485	**0.738**	0.741	0.756	**0.384**
	3	4.52	4.68	**2.94**	9.64	9.63	**6.74**	0.070	0.072	**0.546**	0.794	0.832	**0.438**
	4	4.99	5.22	**3.31**	10.34	10.65	**7.68**	-0.069	-0.135	**0.410**	0.833	0.894	**0.474**
	5	5.73	5.74	**3.85**	12.35	11.54	**8.69**	-0.525	-0.331	**0.245**	0.902	0.952	**0.509**
Station 4	1	3.42	3.17	**1.93**	5.36	4.74	**3.73**	0.813	0.854	**0.909**	0.542	0.526	**0.242**
	2	4.01	3.89	**2.54**	7.09	6.76	**5.46**	0.673	0.703	**0.806**	0.579	0.567	**0.276**
	3	4.36	4.61	**3.04**	7.80	8.92	**6.95**	0.604	0.482	**0.686**	0.616	0.615	**0.321**
	4	4.60	5.09	**3.39**	8.06	9.74	**7.68**	0.578	0.384	**0.617**	0.641	0.659	**0.357**
	5	5.25	5.62	**3.87**	10.10	10.63	**8.18**	0.338	0.266	**0.565**	0.693	0.720	**0.397**
Station 5	1	4.19	2.96	**2.03**	6.05	4.12	**3.40**	0.789	0.902	**0.933**	0.584	0.479	**0.237**
	2	4.12	3.30	**2.51**	6.26	5.12	**4.53**	0.774	0.849	**0.882**	0.587	0.503	**0.264**
	3	4.26	3.79	**2.97**	6.66	6.21	**5.45**	0.744	0.778	**0.829**	0.622	0.558	**0.309**
	4	4.24	4.07	**3.34**	6.72	6.73	**6.11**	0.740	0.739	**0.785**	0.632	0.596	**0.352**
	5	4.55	4.40	**3.63**	7.58	7.21	**6.71**	0.669	0.701	**0.741**	0.658	0.633	**0.385**
Station 6	1	13.41	12.48	**11.15**	28.38	27.08	**25.56**	0.800	0.818	**0.838**	0.373	0.284	**0.235**
	2	15.63	15.01	**13.91**	32.20	30.91	**30.63**	0.742	0.763	**0.767**	0.422	0.360	**0.316**
	3	15.62	15.07	**14.01**	30.00	28.69	**27.92**	0.776	0.795	**0.806**	0.441	0.381	**0.341**
	4	15.91	15.34	**14.14**	28.69	28.14	**26.89**	0.796	0.803	**0.820**	0.458	0.401	**0.356**
	5	17.05	15.66	**14.33**	30.15	28.03	**27.40**	0.774	0.805	**0.813**	0.492	0.414	**0.370**
Station 7	1	38.60	36.08	**33.75**	55.50	52.81	**47.04**	0.934	0.940	**0.953**	0.203	0.183	**0.165**
	2	48.89	48.03	**43.39**	69.67	70.82	**64.59**	0.896	0.892	**0.911**	0.247	0.240	**0.208**
	3	55.01	52.32	**48.69**	76.81	74.48	**69.91**	0.873	0.881	**0.895**	0.275	0.264	**0.237**
	4	56.24	51.83	**49.95**	76.41	72.46	**69.59**	0.875	0.887	**0.896**	0.288	0.271	**0.251**
	5	54.91	53.91	**51.83**	74.64	76.21	**72.95**	0.880	0.875	**0.886**	0.291	0.278	**0.256**
Station 8	1	56.10	54.30	**50.57**	78.56	75.35	**68.06**	0.889	0.898	**0.917**	0.521	0.494	**0.461**
	2	67.95	65.64	**60.77**	95.21	94.87	**83.34**	0.837	0.838	**0.875**	0.606	0.576	**0.511**
	3	71.88	70.46	**64.58**	98.25	100.82	**87.43**	0.826	0.817	**0.863**	0.613	0.592	**0.530**
	4	73.17	72.95	**68.24**	97.74	102.12	**91.54**	0.828	0.812	**0.849**	0.629	0.596	**0.555**
	5	73.48	73.73	**69.42**	97.03	101.46	**93.22**	0.830	0.815	**0.843**	0.632	0.603	**0.559**
Station 9	1	58.93	52.44	**49.00**	89.02	72.36	**69.49**	0.877	0.919	**0.925**	0.257	0.233	**0.208**
	2	79.41	73.90	**68.91**	119.84	101.27	**95.93**	0.777	0.841	**0.857**	0.341	0.324	**0.277**
	3	88.50	85.60	**76.27**	130.50	117.06	**103.81**	0.735	0.787	**0.833**	0.387	0.381	**0.316**
	4	89.52	89.82	**79.75**	123.96	121.48	**107.89**	0.761	0.770	**0.819**	0.397	0.392	**0.329**
	5	88.43	91.94	**81.8**8	117.94	123.62	**110.60**	0.783	0.762	**0.809**	0.397	0.401	**0.336**
Station 10	1	8.33	7.61	**6.11**	19.50	**17.35**	19.19	0.872	**0.899**	0.876	0.133	0.132	**0.084**
	2	10.71	10.09	**7.90**	24.58	23.07	**21.26**	0.797	0.821	**0.848**	0.171	0.164	**0.118**
	3	12.12	11.56	**9.41**	27.39	26.42	**23.69**	0.748	0.765	**0.811**	0.198	0.190	**0.150**
	4	12.99	12.14	**10.16**	27.28	26.22	**24.02**	0.750	0.769	**0.806**	0.218	0.206	**0.170**
	5	15.30	13.27	**11.19**	32.37	29.07	**24.83**	0.647	0.715	**0.792**	0.252	0.223	**0.191**

Table 2: Model comparison for streamflow forecasting for filtered metrics (true values above 75th percentile).

Station	**Day**	**MAE 75%**			**RMSE 75%**			**R^2 75%**			**MAPE 75%**		
		GTN	LS-GTN	TC-GTN	GTN	LS-GTN	TC-GTN	GTN	LS-GTN	TC-GTN	GTN	LS-GTN	TC-GTN
	1	18.22	17.63	15.15	27.99	30.34	25.22	0.878	0.856	0.901	0.113	0.096	0.084
	2	25.12	23.05	21.69	37.43	35.56	36.46	0.781	0.803	0.793	0.154	0.130	0.124
Station 1	3	27.96	24.72	25.03	40.69	35.39	36.47	0.742	0.805	0.793	0.178	0.150	0.153
	4	27.97	27.26	26.74	38.52	38.14	37.55	0.769	0.773	0.780	0.182	0.168	0.168
	5	28.32	28.96	28.18	37.85	38.51	38.95	0.777	0.769	0.763	0.187	0.184	0.180
	1	17.56	14.65	13.87	27.12	25.00	21.85	0.890	0.907	0.929	0.101	0.080	0.078
	2	25.49	23.57	21.17	37.37	36.32	32.90	0.792	0.803	0.838	0.153	0.132	0.121
Station 2	3	30.38	26.53	25.55	41.91	38.52	35.20	0.738	0.779	0.815	0.187	0.158	0.154
	4	31.67	30.80	27.09	42.29	43.18	37.38	0.733	0.722	0.792	0.196	0.188	0.167
	5	34.29	34.73	28.22	46.74	47.66	40.13	0.674	0.661	0.760	0.206	0.209	0.175
	1	5.68	5.47	3.60	10.05	8.43	5.81	0.109	0.374	0.702	0.236	0.236	0.145
	2	7.71	7.69	5.33	14.74	12.57	9.00	-0.914	-0.393	0.286	0.300	0.305	0.208
Station 3	3	9.23	10.05	6.99	17.09	17.33	12.08	-1.575	-1.648	-0.287	0.355	0.383	0.268
	4	10.52	11.53	8.23	18.41	19.30	14.31	-1.987	-2.284	-0.805	0.404	0.441	0.318
	5	12.40	12.77	9.87	21.88	20.92	16.30	-3.221	-2.859	-1.343	0.473	0.491	0.377
	1	5.50	5.31	4.00	8.12	7.69	6.77	0.591	0.633	0.716	0.180	0.179	0.121
	2	7.32	7.60	5.94	11.73	11.58	10.08	0.147	0.170	0.370	0.225	0.236	0.176
Station 4	3	8.21	9.88	7.33	13.06	15.84	12.88	-0.057	-0.555	-0.028	0.245	0.295	0.215
	4	8.98	11.19	8.30	13.86	17.48	14.38	-0.191	-0.894	-0.281	0.272	0.335	0.246
	5	10.50	12.24	9.61	17.49	18.83	15.09	-0.897	-1.198	-0.412	0.311	0.370	0.286
	1	6.52	4.49	4.02	8.68	6.43	5.96	0.577	0.768	0.801	0.212	0.138	0.122
	2	6.45	5.65	5.48	9.03	8.35	8.12	0.542	0.609	0.630	0.200	0.167	0.161
Station 5	3	6.66	6.84	6.59	9.43	10.14	9.61	0.502	0.423	0.482	0.202	0.194	0.194
	4	6.77	7.49	7.39	9.98	10.98	10.81	0.442	0.324	0.344	0.210	0.217	0.218
	5	7.40	8.31	7.98	11.39	11.70	11.71	0.273	0.233	0.231	0.222	0.243	0.236
	1	26.00	26.95	24.94	47.20	47.83	45.53	0.542	0.530	0.574	0.160	0.161	0.145
	2	31.49	32.26	31.10	52.53	51.78	53.76	0.433	0.449	0.407	0.204	0.204	0.188
Station 6	3	31.55	32.24	30.80	48.68	48.29	48.65	0.514	0.521	0.514	0.212	0.215	0.195
	4	31.21	32.46	30.50	45.20	46.06	44.95	0.580	0.564	0.585	0.212	0.218	0.195
	5	33.44	32.89	30.62	47.96	46.55	46.51	0.528	0.555	0.556	0.235	0.225	0.202
	1	59.55	62.51	58.45	83.88	86.85	75.15	0.850	0.839	0.880	0.101	0.107	0.104
	2	79.68	82.63	75.52	106.41	115.88	105.46	0.759	0.714	0.763	0.141	0.143	0.134
Station 7	3	90.02	87.00	82.14	117.06	117.31	110.65	0.708	0.707	0.739	0.164	0.155	0.147
	4	88.04	81.73	82.05	111.76	108.94	106.85	0.734	0.747	0.757	0.163	0.153	0.149
	5	79.85	86.80	87.21	102.75	114.99	112.86	0.775	0.718	0.729	0.153	0.158	0.157
	1	70.63	69.46	63.46	99.70	97.82	84.07	0.797	0.804	0.856	0.114	0.115	0.106
	2	87.78	87.06	83.97	118.93	123.59	111.64	0.711	0.688	0.745	0.145	0.146	0.141
Station 8	3	94.18	98.22	92.27	122.15	136.76	120.31	0.695	0.618	0.704	0.159	0.167	0.158
	4	: 93.23	103.46	98.01	116.62	139.65	122.74	0.722	0.601	0.692	0.157	0.172	0.168
	5	91.57	102.55	96.88	113.38	133.63	124.46	0.737	0.635	0.683	0.157	0.173	0.164
	1	85.29	70.79	70.10	136.94	97.41	100.81	0.584	0.790	0.775	0.120	0.104	0.106
	2	114.99	98.50	105.53	183.00	135.12	141.19	0.257	0.595	0.558	0.166	0.145	0.158
Station 9	3	120.24	106.08	106.97	188.80	148.39	142.88	0.209	0.512	0.547	0.173	0.159	0.162
	4	116.05	108.77	110.88	164.26	147.96	143.01	0.402	0.514	0.546	0.167	0.162	0.167
	5	112.46	111.95	114.90	151.85	149.36	146.53	0.489	0.505	0.524	0.167	0.167	0.172
	1	23.01	18.95	18.66	37.84	33.67	38.68	0.774	0.821	0.764	0.197	0.151	0.138
	2	30.20	27.10	23.23	47.06	43.25	41.02	0.651	0.705	0.735	0.276	0.233	0.188
Station 10	3	32.95	30.50	25.49	50.06	46.84	42.63	0.605	0.654	0.714	0.306	0.275	0.215
	4	33.97	31.24	26.10	49.02	45.45	42.29	0.621	0.674	0.718	0.322	0.289	0.228
	5	40.09	34.22	28.40	59.14	52.23	44.14	0.449	0.570	0.693	0.363	0.311	0.256