

---

# Short-term Hourly Streamflow Prediction with Graph Convolutional GRU Networks

---

Muhammed Sit<sup>1</sup> Bekir Demiray<sup>1</sup> Ibrahim Demir<sup>1</sup>

## Abstract

The frequency and impact of floods are expected to increase due to climate change. It is crucial to predict streamflow, consequently flooding, in order to prepare and mitigate its consequences in terms of property damage and fatalities. This paper presents a Graph Convolutional GRUs based model to predict the next 36 hours of streamflow for a sensor location using the upstream river network. As shown in experiment results, the model presented in this study provides better performance than the persistence baseline and a Convolutional Bidirectional GRU network for the selected study area in short-term streamflow prediction.

## 1. Introduction

The number and devastating impacts of natural disasters have grown significantly worldwide. The latest report released by Munich Re (2020) states that natural disasters such as hurricanes, floods, and other disasters caused more than \$210 billion estimated damage worldwide, while \$95 billion of the damage occurred in the US. According to the same report, summer floods in China were the costliest natural disaster worldwide in 2020, and the number of fatalities in floods was higher than other natural disasters (Munich Re, 2020). Recent studies show that the frequency and impact of flooding increases in certain regions as a result of climate change (Davenport et al., 2021; The National Centers for Environmental Information, 2021; Tabari, 2020) due to an increase in sea level (Strauss et al., 2016) and frequency of extreme precipitation (Daffenbaugh et al., 2017), or intensifying hurricane rainfall (Trenberth et al., 2018). Thus, it is crucial to predict streamflow and, consequently, flooding to mitigate its devastating effects in terms of damage and fatalities.

<sup>1</sup>IIHR—Hydroscience & Engineering, The University of Iowa, Iowa City, Iowa, USA. Correspondence to: Muhammed Sit <muhammed-sit@uiowa.edu>.

Many physical and data-driven methods have been proposed to achieve accurate streamflow predictions, and recent studies show that deep learning models often provide more accurate results than physical-based models (Gauch et al., 2019; Xiang & Demir, 2020; Xiang et al., 2021). Recurrent neural networks based approaches are mainly used for the task as a result of the success of these architectures on time series problems (Sit & Demir, 2019). (Hu et al., 2018) proposed a Long short-term memory (LSTM) (Hochreiter & Schmidhuber, 1997) model that predicts the hourly streamflow from 1 to 6 hours lead time. (Kratzert et al., 2018) developed another LSTM model that predicts daily runoff for the first time. (Xiang & Demir, 2020) developed a model that uses multiple GRUs (Cho et al., 2014) and Time-distributed layers in order to predict up to 120 hours of streamflow. Many studies aim to predict streamflow based on various data such as evapotranspiration, current streamflow, or weather data. More detailed information about deep learning studies on streamflow prediction can be found in (Sit et al., 2020).

Graph Neural Networks (GNNs) and variants including Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), or Graph Recurrent Networks (GRNs) have gained much attention as a result of their performance on many deep learning tasks comprising sequenced data points which can be expressed as a graph (Wu et al., 2020; Zhou et al., 2020). For instance, (Seo et al., 2018) proposes a convolutional recurrent neural network architecture that captures spatial dependencies with CNNs and identifies dynamic patterns with GRUs in the structured data sequences by showing two use cases; predicting moving MNIST data and modeling natural language with the Penn Treebank dataset. (Bai et al., 2020) uses a graph convolutional recurrent network with two new modules that allow to learn node-specific patterns and discover spatial correlations from data separately for a traffic prediction task. Based on successful implementations of GNNs on structured sequences, in this paper, we present a model based on Graph Convolutional GRUs for streamflow prediction that we will refer to as StreamGConvGRU. Furthermore, we show our preliminary results for StreamGConvGRU using real-world data for a small network of streamflow sensors. To the best of our knowledge, this is the first work that presents at least some preliminary results for streamflow prediction using

Graph Neural Networks.

## 2. Methodology

GNNs expect the input data as a graph and successively outputs a graph. Since most rivers are connected to each other and form a network of rivers, locations along rivers, more specifically stream gauge locations along a river network, could be converted into a graph and fed to a GNN.

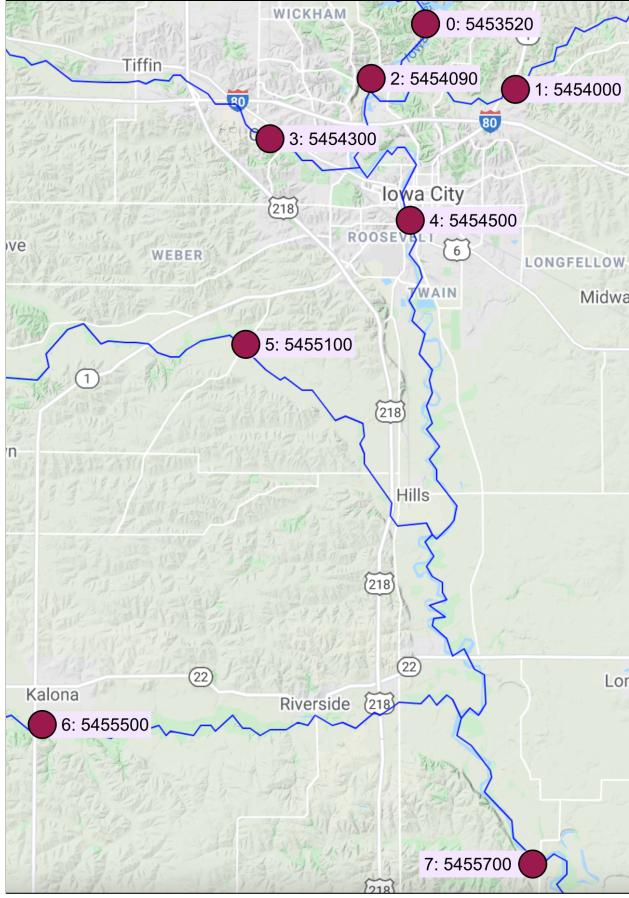


Figure 1. Visualization of study area and USGS sensor locations on Google Maps with their *id\_in\_graph*: *usgs\_id*

### 2.1. Dataset

The United States Geological Survey (USGS) maintains a network of stream gauges while actively deploying new ones all around the United States. The historical measurements for each of these sensors with a temporal resolution of 15 minutes are publicly available. Even though there are hundreds of USGS stream gauges in each watershed, in order to provide a proof of concept model setup, we chose to use eight sensors in Johnson County, Iowa, that form a small subnetwork. Seven of these stream gauges are within

the watershed of the eighth one, and they feed the water to the eighth stream gauge located at the outlet of the watershed. In this study, we aim to predict the next 36 hours of measurements for the eighth stream gauge at the outlet by predominantly using 36 hours of historical streamflow data for all eight stream gauges. USGS sensor ids of these stream gauges are as follows, 05453520, 05454000, 05454090, 05454300, 05455100, 05454500, 05455500, 05455700, the last one being the sensor for that we aim to predict the future streamflow values (Figure 1).

Along with historic stream measurements, the StreamGConvGRU is also fed with precipitation measurements for the past 36 hours and the next 36 hours for all the stream gauge locations. The resource we employed for the rainfall data is the Stage IV hourly multi-sensor precipitation analysis. Since all the precipitation that falls into a river's watershed affects the streamflow, the rainfall data within every sensor's watershed were aggregated by summing up. Also, in order to match stream gauge temporal resolution with Stage IV, the USGS sensor measurements were aggregated to hourly by averaging them.

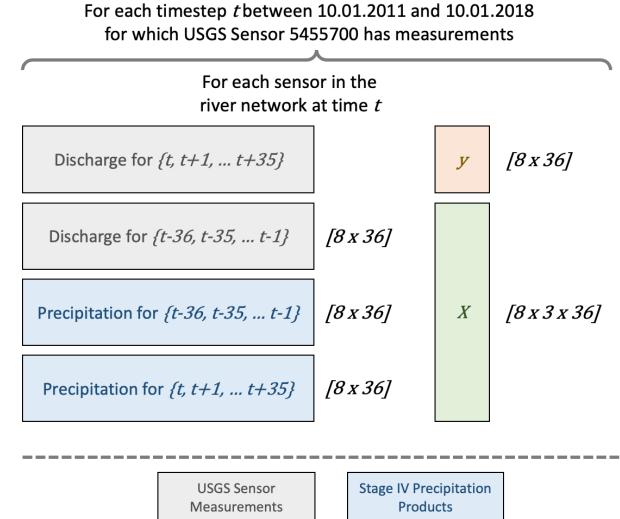


Figure 2. Summary of data sources and how each snapshot was formed.

Table 1. Dataset splits and, start and end dates for each split.

SPLIT	START DATE	END DATE	# OF SNAPSHOTS
TRAIN	10.01.2011	10.01.2016	32789
VALIDATION	10.01.2016	10.01.2017	6896
TEST	10.01.2017	10.01.2018	8253

Within the selected date range (10/1/2011 - 10/01/2018),

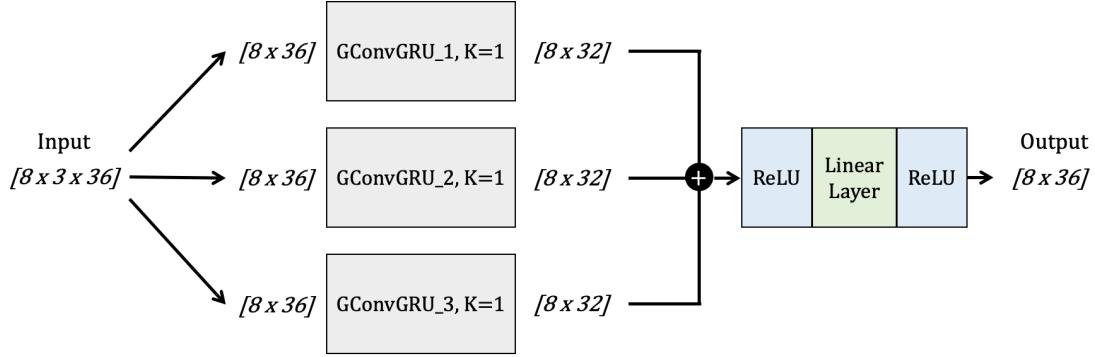


Figure 3. Architecture proposed for StreamGConvGRU.

a snapshot was created for each hour except for the times when the USGS sensor 05455700 is down. In order to ensure a continuous dataset and introduce noise to the training process, we simply used zeros for when any of the seven upstream sensors did not have any measurements. The data acquisition step ends with 47,938 total snapshots. Train, validation, test split (Table 1) then was made by snapshot timestamps, and a normalization to map values to  $[0 - 4]$  was applied to all subsets using minimum and maximum values of streamflow and aggregated precipitation observations for the USGS sensor id 05455700 in the training subset. In the end, each snapshot had an input with the size of  $[8 \times 3 \times 36]$  (number of stream gauges  $\times$  number of time-series sequences used  $\times$  length of the sequence). The output size similarly was  $[8 \times 36]$  (Figure 2). Since a GNN expects graph properties, the graph for sensors was built by considering the hydrological connectivity of the nodes and the distance between USGS sensors as weights.

## 2.2. Network

StreamGConvGRU is a GNN model based on Graph Convolutional Gated Recurrent Unit Networks (GConvGRUs) (Seo et al., 2018). While GConvGRUs have the ability to understand spatial structures with convolutional layers, they also incorporate Gated Recurrent Unit Networks (GRUs) to understand dynamic temporal patterns. For details about GConvGRUs and Graph Convolutional Recurrent Neural Networks, please refer to (Seo et al., 2018).

The StreamGConvGRU uses three separate GConvGRU subnetworks (Figure 3) for each of the three sequences explained in the previous subsection, that is, previous 36 hours of discharge measurements, previous 36 hours of precipitation observations, and next 36 hours of precipitation measurements. The outputs of each of these subnetworks then are summed up and fed to a linear layer that outputs 36 hours of predictions for each of the eight sensors. It should

be noted that since a GNN outputs a graph, the product of the network is also a set of sequences, but since the goal is to predict the next discharge levels for the USGS sensor 05455700, the training is done to optimize one sequence while other seven sequences were simply ignored.

We also built various GRU networks for comparison purposes. The best performing GRU network for our dataset was a Convolutional Bidirectional GRU model, which we will refer to as ConvBiGRU (Figure 4). ConvBiGRU was trained by feeding a matrix with the shape of  $[24 \times 36]$  for each snapshot instead of the graph. The matrix was created for each snapshot by combining all sequences for all USGS sensor locations within the river network. The output of the GRU network is only a sequence of subsequent discharge measurements rather than measurements for each of the sensors in the sensor network.

Networks described here were trained using RMSprop (Hinton et al., 2012) optimizer with L1 loss as the cost function on NVIDIA Titan V GPUs. During the training, the best scoring networks' weights over the validation set were saved, and the results that will be presented in the next section were generated by predicting streamflow for the test dataset. While the ConvBiGRU was implemented using PyTorch numeric computing library (Paszke et al., 2017), the StreamGConvGRU was implemented using pytorch-geometric-temporal library (Rozemberczki et al., 2021) that is built on top of Pytorch, and pytorch-geometric libraries (Fey & Lenssen, 2019).

## 3. Results and Discussions

Besides the neural network architecture described in the previous section, we define a naive baseline model here. Namely, persistence, which is inherently data-driven, is proposed to be used as a baseline model in streamflow prediction (Ghimire & Krajewski, 2020). Persistence is built

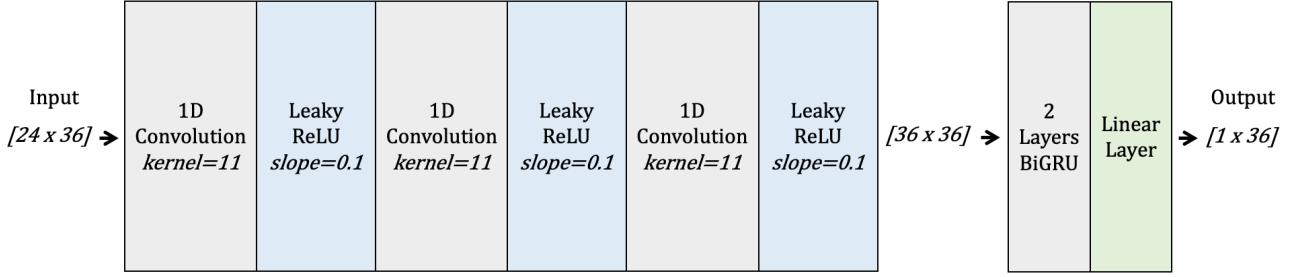


Figure 4. Convolutional Bidirectional GRU network's architecture.

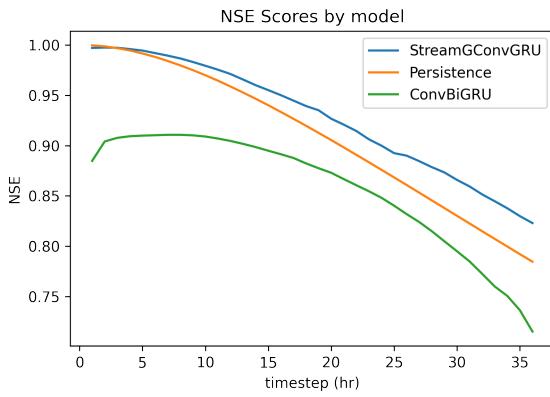


Figure 5. Hourly NSE scores for StreamGConvGRU, persistence and ConvBiGRU.

relying on the idea that “tomorrow will be like today.” It has a straightforward implementation, when prediction is being done at timestep  $t$ , every future measurement of sensor  $s$  for timestep  $t'$  is predicted as  $s[t - t']$ . In other words, persistence assumes all predictions will be the same as the latest measurement at time  $t$ .

To compare persistence, ConvBiGRU, and StreamGConvGRU, we used Nash-Sutcliffe Efficiency (NSE) score (Nash & Sutcliffe, 1970). NSE score is a widely used efficiency metric for streamflow prediction models, both physically based and data-driven applications, and defined as

$$NSE = 1 - \frac{\sum_{t=1}^T (Q_m^t - Q_o^t)^2}{\sum_{t=1}^T (Q_o^t - \bar{Q}_o)^2}$$

where  $Q_m^t$ ,  $Q_o^t$  and  $\bar{Q}_o$  mean modeled discharge at time  $t$ , observed discharge at time  $t$  and mean of observed discharges, respectively. Please note that the best NSE score a streamflow prediction can get is 1.0. Thus, out of two

streamflow prediction models, the model with the NSE score closer to 1.0 is the better one.

The preliminary results for hourly NSE scores are presented in Figure 5. As it can be seen in Figure 5, even though it was the best GRU-based model we built, ConvBiGRU does not get closer to persistence. The StreamGConvGRU model we present in this paper, on the other hand, outperforms both persistence and the ConvBiGRU model by a significant margin after the fifth hour. The performance of the persistence baseline for the first few hours could be explained with the approach it uses. Streamflow rates typically do not change drastically in the first few hours. However, when it starts to change, the StreamGConvGRU model takes the lead as it takes advantage of upstream nodes’ data and connectivity. A similar relationship between the proposed model and persistence can be seen in (Xiang & Demir, 2020). Also, we want to stress that various others RNNs we built to compare with StreamGConvGRU did not produce comparable NSE scores; consequently, we decided not to include them here for the sake of simplicity.

#### 4. Conclusions

The importance of streamflow and, consequently, flood prediction increases as a result of the devastating effects of climate change. This paper presented an approach where we employed a Graph Convolutional GRU Networks based model, StreamGConvGRU, to predict 36 hours of streamflow for a sensor location using the upstream river network. As shown in the preliminary results, the StreamGConvGRU provides better performance than the persistence baseline and a Convolutional Bidirectional GRU model in our study region for short-term hourly streamflow prediction. For the future work, we aim to focus on the following three points: (1) exploring GConvGRUs’ abilities with a more extensive river network; (2) prediction of greater lead time on streamflow; (3) prediction of streamflow on each node in the river network instead of focusing only on one node at the outlet.

## References

Bai, L., Yao, L., Li, C., Wang, X., and Wang, C. Adaptive graph convolutional recurrent network for traffic forecasting. *arXiv preprint arXiv:2007.02842*, 2020.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

Davenport, F. V., Burke, M., and Diffenbaugh, N. S. Contribution of historical precipitation change to us flood damages. *Proceedings of the National Academy of Sciences*, 118(4), 2021.

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., Charland, A., Liu, Y., Haugen, M., Tsiang, M., et al. Quantifying the influence of global warming on unprecedented extreme climate events. *Proceedings of the National Academy of Sciences*, 114(19): 4881–4886, 2017.

Fey, M. and Lenssen, J. E. Fast graph representation learning with pytorch geometric. *arXiv preprint arXiv:1903.02428*, 2019.

Gauch, M., Mai, J., Gharari, S., and Lin, J. Data-driven vs. physically-based streamflow prediction models. In *Proceedings of 9th International Workshop on Climate Informatics*, 2019.

Ghimire, G. R. and Krajewski, W. F. Exploring persistence in streamflow forecasting. *JAWRA Journal of the American Water Resources Association*, 56(3):542–550, 2020.

Hinton, G., Srivastava, N., and Swersky, K. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited on*, 14(8), 2012.

Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., and Lou, Z. Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. *Water*, 10(11):1543, 2018.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M. Rainfall–runoff modelling using long short-term memory (lstm) networks. *Hydrology and Earth System Sciences*, 22(11):6005–6022, 2018.

Munich Re. Record hurricane season and major wildfires – the natural disaster figures for 2020. <https://www.munichre.com/en/company/media-relations/>

media-information-and-corporate-news/media-information/2021/2020-natural-disasters-balance.html, 2020. [Online; accessed 30-May-2021].

Nash, J. E. and Sutcliffe, J. V. River flow forecasting through conceptual models part i—a discussion of principles. *Journal of hydrology*, 10(3):282–290, 1970.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A. Automatic differentiation in pytorch. 2017.

Rozemberczki, B., Scherer, P., He, Y., Panagopoulos, G., Astefanoaei, M., Kiss, O., Beres, F., Collignon, N., and Sarkar, R. Pytorch geometric temporal: Spatiotemporal signal processing with neural machine learning models. *arXiv preprint arXiv:2104.07788*, 2021.

Seo, Y., Defferrard, M., Vandergheynst, P., and Bresson, X. Structured sequence modeling with graph convolutional recurrent networks. In *International Conference on Neural Information Processing*, pp. 362–373. Springer, 2018.

Sit, M. and Demir, I. Decentralized flood forecasting using deep neural networks. *arXiv preprint arXiv:1902.02308*, 2019.

Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., and Demir, I. A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*, 82(12):2635–2670, 2020.

Strauss, B. H., Kopp, R. E., Sweet, W. V., and Bittermann, K. *Unnatural coastal floods: Sea level rise and the human fingerprint on US floods since 1950*. Climate Central, 2016.

Tabari, H. Climate change impact on flood and extreme precipitation increases with water availability. *Scientific reports*, 10(1):1–10, 2020.

The National Centers for Environmental Information. Us billion-dollar weather and climate disasters. <https://www.ncdc.noaa.gov/billions/>, 2021. [Online; accessed 30-May-2021].

Trenberth, K. E., Cheng, L., Jacobs, P., Zhang, Y., and Fasullo, J. Hurricane harvey links to ocean heat content and climate change adaptation. *Earth's Future*, 6(5): 730–744, 2018.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 2020.

Xiang, Z. and Demir, I. Distributed long-term hourly streamflow predictions using deep learning—a case study for state of iowa. *Environmental Modelling & Software*, 131: 104761, 2020.

Xiang, Z., Demir, I., Mantilla, R., and Krajewski, W. F. A regional semi-distributed streamflow model using deep learning. *EarthArXiv preprint*, 2021.

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020.