Multivariate Time Series Classification for Flood Prediction

Tianbo Yu
Department of ECE
Texas A&M University
College Station, TX 77840
tianbo@tamu.edu

Tianyang Ding
Department of ECE
Texas A&M University
College Station, TX 77840
seizethedty@tamu.edu

Junyu Han
Department of CSE
Texas A&M University
College Station, TX 77840
jyhan@tamu.edu

Abstract

Over the past decade, multivariate time series classification has been applied in various fields and received great attention. We propose an innovative way of taking advantage of the Multivariate Long Short Term Memory Fully Convolutional Network (MLSTM-FCN) and Multivariate Attention LSTM-FCN (MALSTM-FCN) to predict flood areas based on time-series data as input. Due to the unavailable of proper open-source datasets, we script the rainfall and water elevation data on Harris County Flood Warning System to form our own dataset and train and validate the model on it. The accuracy above 99% was achieved and high speed for both training and test process was enabled by the model. Furthermore, the model is shown to be able to fit by a small amount of data and accomplish robust real-time flood prediction task.

1 Introduction

1.1 Background

As one of the most catastrophic disasters, floods are causing tremendous impacts on the economic development and social well-being of communities. One of the main contributors to the flood vulnerability of communities, especially in urban areas, is the malfunction of flood control systems in the face of extreme rainfalls. Robust and accurate prediction contributes highly to disaster management strategies, analysis, and rescue [1]. However, the massive difference in time, occurrence location, and dynamic nature of climate conditions make the prediction complex. Today's major flood prediction models are mainly data-specific and involve various simplified assumptions. Mainly, they can be mainly divided into event-driven, empirical black box, lumped and distributed, stochastic, deterministic, and hybrids methods [2].

The most time-honored and frequently used method for flood prediction is the hydrological approach. As one of the primary studies in the field, a conceptual model called the SAC-SMA model was proposed for operational streamflow forecasting [3]. After that, Beven and Kirkby (1979) [4] introduced a physically-based hydrological model called TOPMODEL to predict basin behavior with measured and estimated parameters. Zhao RJ (1980) [5] introduced XINANJIANG (XAJ) model, which is a conceptual hydrological model to be used in humid and semi-humid basin areas contains fifteen parameters optimized by objective functions. Moore (1985) [6] devised a conceptual lumped hydrological model named HYM model, using two series of linear reservoirs to route surface and subsurface flow. Luo et al., 1992 [7] combined the generation mechanisms of infiltration excess (Horton) runoff and saturation excess (Dunne) runoff with building HYB model. To improve the XAJ model for semi-arid and semi-humid regions, Li et al., 2005[8] designed the XAJ-Green-Ampt model based on the Green-Ampt equation. In these models, a combination of data (including rainfall, elevation, moisture, etc.) and expert knowledge (including evaporation rate, water capacity, coefficients of transition, etc.) are used as the input models. The prediction of hydrological information

(the elevation level, inundation probabilities, etc.) in a specific area of a watershed is the output of models.

In addition to physical models, data-driven models also have a long tradition in flood modeling, and the Bayesian model averaging (BMA) method is one of those used technologies for flood prediction [9]. Various research studies have been conducted using Bayesian Modeling for the analysis of stormwater drainage systems and real-time prediction of inundation areas. For inundation forecasting, it is a scheme for the model combination using likelihood measures as model weights [10, 11, 12]. BMA provides a coherent mechanism for accounting for model uncertainty and usually outperforms other multi-model combination methods, which make it gaining popularity [13, 14, 15, 16].

Recently, deep learning methods are applied in the field of flood prediction. Fang et al. [17] created a system that predicts SMAP level-3 moisture product with Long Short-Term Memory (LSTM) deep neural network, and applied atmospheric force, model-simulated moisture, and static physiographic attributes as input data. K. Hiroi and N. Kawaguchi [18] implemented a CWF prediction system that produces accurate and early predictions using a Linear Regression (LR) of deep learning approach with data assimilation. Liu et at. [19] utilized LSTM networks to predict sequential flow rate values based on a set of collected flood factors.

Due to the continuous nature of the hydrological information, flood prediction belongs to the category of time-series prediction in the field of deep learning. We believe the flood data has some similarities with clinic data in terms of their data composition and possible patterns. Therefore we carefully analyzed some works in competitions using clinic data and used their ideologies and model structures for flood prediction. The MIMIC-III and PhysioNet are two open-source and commonly used clinical datasets, and they can be helpful for early detection of a patient's risk of sepsis. The data consist of records from ICU stays, and all patients were adults who were admitted for a wide variety of reasons to cardiac, medical, surgical, and trauma ICUs. Up to 42 variables were recorded, including Demographics, Vital Signs, and Laboratory values. A number of works have been made with these datasets and that of Che et al. [20], Rubanova et al. [21] and Edward et al. [22] are considered in this project. Che et al. [20] developed the GRU-D deep learning model based on Gated Recurrent Unit (GRU). It takes two representations of missing patterns, i.e., masking and time interval, and incorporates them into a deep model architecture so that it not only captures the long-term temporal dependencies in time series but also utilizes the missing patterns to achieve better prediction results. Rubanova et al. [21] generalized RNNs to have continuous-time hidden dynamics defined by ordinary differential equations (ODEs), and came up with a model called ODE-RNNs. ODE-RNNs can naturally handle arbitrary time gaps between observations, and explicitly model the probability of observation times using Poisson processes. Edward et al. [22] proposed a model combining two novel techniques, GRU-ODE and GRU-Bayes, which allows feeding sporadic observations into continuous ODE dynamics describing the evolution of the probability distribution of the data. Additionally, they showed that this filtering approach enjoys attractive representation capabilities. Finally, they demonstrated the value of GRU-ODE-Bayes on both synthetic and real-world data. Moreover, while a discretized version of their model performed well on the full MIMIC-III data set, the continuity prior to their ODE formulation proves particularly important in the small-sample regime, which is particularly relevant for real-world clinical data where many data sets remain relatively modest in size. GRU-ODE-Bayes handles the sporadic data more naturally and can more finely model the dynamics and correlations between the observed features, which results in higher performance in terms of NegLL and MSE than other methods for MIMIC-III dataset.

1.2 Motivation

As we mentioned before, due to the unavailability of hydrological information and expert knowledge required by hydrological models, a feasible solution for flood prediction is sometimes expensive or even impossible to be obtained in some areas. Therefore, it is imperative to improve tools and techniques that enable a less demanding flood prediction method with a reliable performance during the flood events. However, with the development and maturity of LSTM models, an accurate and robust solution is applicable with available time-series data. And this solution meets the demands of flood analysis, management, and rescue.

1.3 Our Work

The paper aims at developing a novel LSTM-FCN modeling approach that enables accurate and robust real-time flood prediction of an existing flood control network in an urban area. To achieve this goal, we first analyzed and reproduced MLSTM-FCN and MALSTM-FCN and experimented on existing datasets to show its correctness. Then the flood data, including rainfall amount and elevation height, were gathered and organized into the right format for the model. The model was trained on these data and fine-tuned until an accuracy above 99% was achieved. It performs similar to or even better than state-of-the-art models in the field of flood prediction.

2 Methodology

2.1 Reproduction of MLSTM-FCN and MALSTM-FCN Model

2.1.1 Network Architecture

Multivariate Long Short Term Memory Fully Convolutional Network (MLSTM-FCN) and Multivariate Attention LSTM-FCN (MALSTM-FCN) have achieved success in classifying multivariate time series [23].

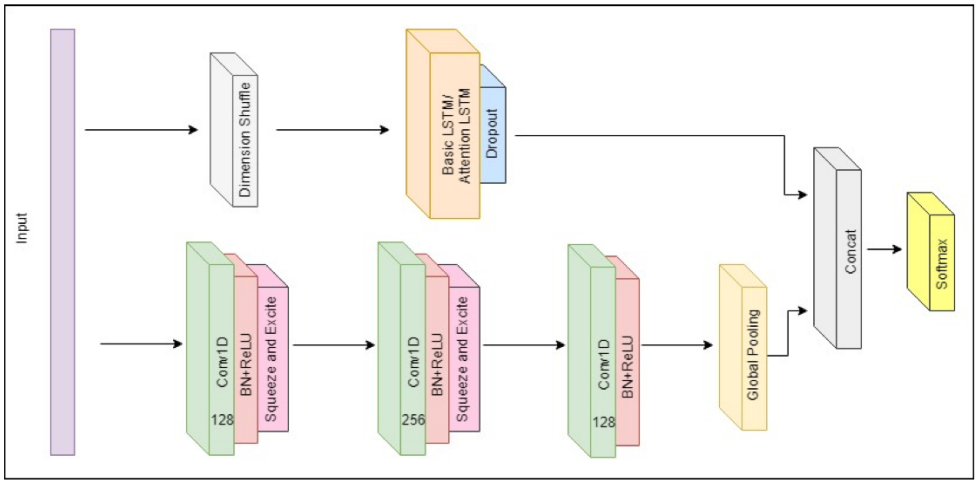

Figure 1: The MLSTM-FCN architecture. LSTM cells can be replaced by Attention LSTM cells to construct the MALSTM-FCN architecture. [23]

As shown in Figure 1, this network is consist of a fully convolutional network (FCN) and a LSTM block. The FCN functioned as a feature extractor is replicated from the original version by Wang et al [24], which contains three temporal convolutional blocks. Each block contains a convolutional layer with 128, 256, and 128 filters respectively and a kernel size of 8, 5, and 3 respectively. Each convolutional layer is succeeded by batch normalization (BN), with a momentum of 0.99 and epsilon of 0.001. The batch normalization layer is followed by the ReLU activation function. In addition, the first two convolutional blocks conclude with a squeeze-and-excite (SE) block replicated from the original version by Hu et al [25].The SE block is used to recalibrates the input feature maps because not all feature maps impact the following layers to the same degree. This adaptive recalibration of the feature maps can be considered as a form of learned self-attention on the output feature maps of prior layers. Figure 2 shows the process of how the SE block is computed. For all SE blocks, the reduction ratio r is set to the default number, 16. The final convolutional block is followed by a global average pooling layer.

In addition, the multivariate time series input is passed through a dimension shuffle layer (explained in Section 2.1.2), succeeded by the LSTM block. The LSTM block is identical to the block by Karim et al [26], comprising either an LSTM layer or an Attention LSTM layer, which is succeeded by a dropout layer.

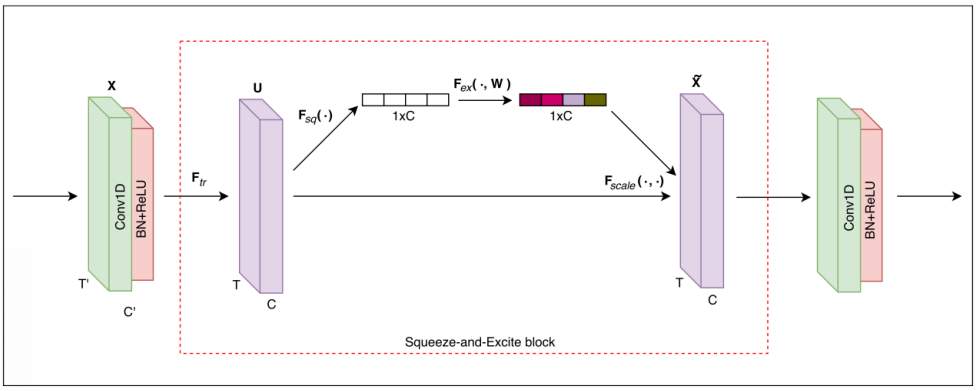

Figure 2: The computation of the temporal squeeze-and-excite block. [23]

2.1.2 Network Input

The time series dataset is a tensor of shape (N, Q, M), where N is the number of samples, Q is the maximum number of time steps amongst all variables and M is the number of variables in each time step. The input to the fully convolutional and LSTM blocks vary according to different datasets of shape (N, Q, M) as stated.

The dimension shuffle which transposes the temporal dimension of the input data can improve the efficiency of the model. The reason is that in most cases M is less than Q, now the LSTM requires M time steps to process Q variables instead of Q time steps to process M variables per time step, without losing information.

2.2 Implement on Flood Network

We implemented MLSTM-FCN and MALSTM-FCN and did several modifications on the original networks:

1. In the original implementation, the training process was monitored based on training loss, which is incorrect, we changed it to validation loss.

2. We introduced early stopping as a callback, which can reduce the probability of overfitting and save lots of computation resource without losing accuracy.

3. We tried out LeakyReLU and PReLU as activation functions, which shows the potential to achieve better performance.

We created our own dataset (Section 3.2) apart from several datasets mentioned in [23] and did experiments on all of them, the results can be seen in Section 4.

3 Case Study: Harris County Flood Control Network

3.1 A brief introduction of Flood Events and Watersheds in Harris County

(a) Study Area
Harris County, home to Houston, the fourth largest city in the United States, is the third-largest county in the United States and has more than 4,023 km of flood control networks. Harris County can be divided into 22 major watersheds, and a watershed is the area of land that catches rain and drains into a marsh, bayou, creek, river, lake, or bay. In Houston, the water collected in each watershed drains into a major river or bayou and eventually drains into Galveston Bay. In terms of the flood control network, only about 6% of 4,023 km of humanmade or improved channels are concrete-lined; most are grass-lined. Figure 3 shows the distribution.

The average annual rainfall within Harris County is approximately 48 inches, which makes this area a flood-prone region. The flood control system in Harris County is designed to perform well in carrying the stormwater runoff resulting from average rainfall; However, under extreme conditions

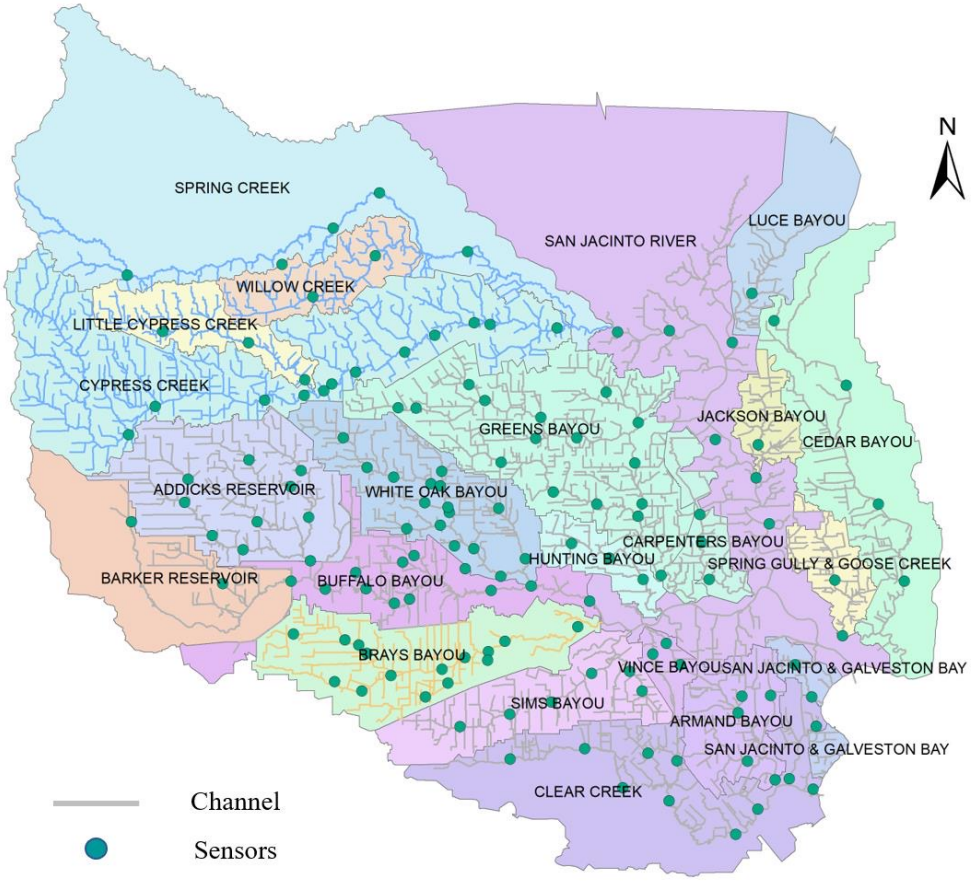

Figure 3: The watershed division of Harris County

(such as Hurricane Harvey) when rainfall exceeds the capacity of the system, many areas tend to flood. There are sensors placed along channels in each watershed, collecting data regarding rainfall amount, inundation status, and water elevation height during extreme events, which can be used as input data for machine learning models. We take three watersheds as an example, and Table 1 shows a characteristic of them.

Table 1: Characteristics of Research Watersheds

Name	Area (Mile^2)	Population	Length (Mile)	Num. of Sensors
Northwest watershed	371	422103	362	20
Brays bayou watershed	127	717198	121	15
White Oak bayou watershed	111	1433250	146	14

(b) Three studied flood events
From 2016 to 2017, Harris Count experienced three relatively severe floods, i.e., Tax Day Flood (2016), Memorial Day Flood (2016), and Hurricane Harvey (2017), as is shown in Table 2. The data required for developing the proposed MALSTM-FCN model was gathered for these flood events.

Tax Day Flood arrived on April 16th, 2016, and continued for two days. It has a 12-hour rainfall amount of 12-16 inches widespread across northwest Harris County. Estimated damages contains 9,800+ homes, 2,700 apartment units, 40,000 vehicles and 9 fatalities.

Six weeks later, Memorial Day Flood attacked again on May 27th, 2016, and sustained for two days. Since the detention basins and water channels had previously been impacted by the Tax Day Flood, the ground was overfilled with stormwater, and some slow-draining parts of the channel network in

Table 2: Influence of Inundation Events in Harris County

Flood Name	Date	Precipitation	Main Influenced Areas
Tax Day Flood	4/16-4/17(2016)	12-16 inches in 12 hours	Whole Harris county
Memorial Day Flood	5/27-5/28(2016)	8-13 inches totally	North Harris county
Hurricane Harvey	8/25-8/31(2017)	28-44 inches in 4 days	Whole Harris county

Harris County were still suffering from higher water elevation levels than usual. The north Harris County received 8-13 inches of rainfall, which was too much for drenched grounds, and some former floodplain was inundated once again.

Hurricane Harvey landed on August 25th, 2017 as a Category 4 hurricane (130 mph). Rainfall occurred over seven days, and the vast scope of inundation was a result of an extraordinary 4-day rainfall event of 28-44 inches widespread in Harris County. Average rainfall during the time over the entire county was 33.7" (69% of the probable maximum precipitation), and the total volume was about one trillion gallons. Hurricane Hervey impacted everyone in Harris County (4.7 million people), and lead to damage of 140,000-160,000 homes or business building, 300,000 vehicles and 36 fatalities.

3.2 Training and test data

The time-series data of rainfall amount and water elevation height are used as the input data. Specifically, the data during Hurricane Harvey and Memorial Day Flood are treated as training data as well as that of Tax Day Flood is considered as test data. The input data for both training and test are in the form of matrix, where each row stands for the corresponding data at a certain time interval and each column represents a sample. In particular, labels are binary values standing for inundation status (water level is higher than the top of spillway) or not and each sample has a label. The rainfall amounts are the sum of rainfall amount in the future 6 hours of the corresponding areas. Due to the absence of rainfall forecast during inundation time, the actual rainfall amounts observed by the Harris County Flood Warning System are considered as the predictive values. Water elevation heights are the distance between water height and top of the spillway of the corresponding channel components.

The time periods of 7 days and intervals of 30 minutes are set such that different data about rainfall amount and elevation can be gathered according to as time series. Therefore, there are a totally of 336 intervals for each inundation event. For each interval, the data of inundation status, rainfall amount and water elevation height are obtained for both training and test dataset. And the model only considers the rainfall amounts and water elevation heights in the last two days such that each sample has a sequence length of 96 and the number of variates of 2. After doing this to all the sensors in Harris County, a training dataset in the shape of (56334, 2, 96) and a test dataset in the shape of (28167, 2, 96) were obtained.

4 Results Analysis

4.1 Datasets

We selected 6 datasets to validate the reproductive models and then test the model on the flood control network. 4 of the 6 datasets are benchmark datasets that are most recently utilized by Schafer and Leser [27]. The other 2 of the 6 datasets are from the UCI repository. These datasets are summarized in Table 3.

Table 3: Properties of all selected datasets. The yellow cells are datasets used by [27]; the blue cells are UCI repository.

Dataset	Num. of Classes	Num. of Variables	Max Length	Tasks	Train-Test Split
Japanese Vowels	9	12	26	Speech Recognition	42-58 split
LIBRAS	15	2	45	Sign Language Recognition	38-62 split
LP2	5	6	15	Robot Failure Recognition	36-64 split
PenDigits	10	2	8	Digit Recognition	2-98 split
AREM	7	7	480	Activity Recognition	50-50 split
HAR	6	9	128	Activity Recognition	71-29 split

4.2 Reproduction Result

Firstly, we reproduced MALSTM-FCN model (MLSTM-FCN can be implemented by just removing the attention block) to validate its correctness. Then we compared with modified models, which implements early stopping, LeakyReLU, and PReLU separately. Table 4 shows the comparison on test accuracy, green cells show better performance than original network while the red cell shows bad performance. Table 5 shows the number of training epochs on original training setting and with early stopping mechanism ($patient = 25$).

Table 4: Comparasion on testing accuracy

Dataset	MALSM-FCN	MALSM-FCN(EarlyStop)	MALSM-FCN(LeakyReLU)	MALSM-FCN(PReLU)
Japanese Vowels	98.92	98.65	98.38	98.65
LIBRAS	96.92	96.92	96.24	96.41
LP2	73.33	66.67	73.33	76.66
PenDigits	96.18	96.24	96.28	96.10
AREM	84.62	46.15	87.18	84.62
HAR	92.60	92.91	92.70	91.89

Table 5: Comparison on training epochs

Dataset	Original Epochs	Early Stopping
Japanese Vowels	600	145
LIBRAS	1000	149
LP2	1000	163
PenDigits	1000	126
AREM	600	34
HAR	600	29

On most datasets, the early stopping mechanism works well. However, on the AREM dataset, with early stopping we got quite low test accuracy. After visualizing the training process without early stopping shown in Figure 4, we found that there exists a local minimun, with early stopping we were not able to achieve the global optimization. In addition, different activation functions show a potential to further improve the model performance.

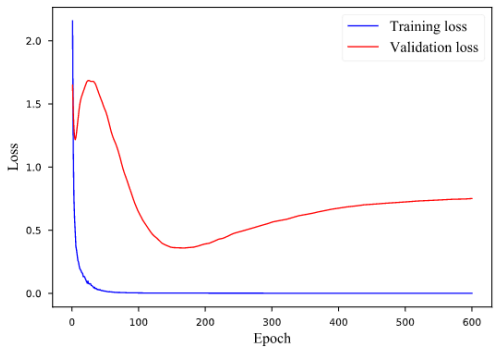

Figure 4: Training and validation loss vs. Epochs on AREM

4.3 Result of Our Study Case

After validating the correctness of the reproducted model, we implemented the model on the flood dataset, which is gathered and organized into the right format for the model. The training process with different techniques is shown in Figure 5. This model manifests its solid ability for flood prediction tasks, with a final test accuracy of above 99%.

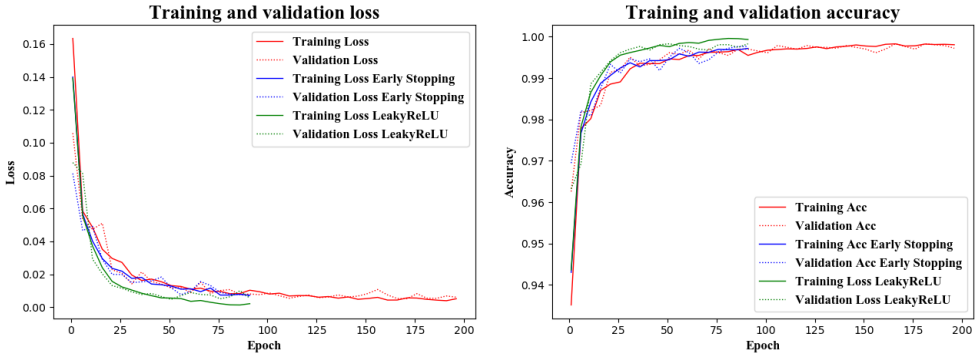

Figure 5: Training process with different techniques

5 Conclusion

This paper presents a modified MLSTM-FCN and MALSTM-FCN model to simulate the failure cascade process and contribute to flood control network management. The proposed model enables minimal preprocessing and feature extraction with state-of-the-art accuracy. Also, it generates robust results on our complex multivariate time series classification task, even in the case that the training samples are relatively limited.

The proposed model was applied to the Harris County, Texas, flood control network. Hurricane Harvey, Tax Day Flood, and Memorial Day Flood data were combined and split into training and test set to form our unique dataset, which trained and validated the model. The simulation of failure cascades achieves above 99% accuracy.

The limitation of our model is that due to the absence of prediction rainfall data, we take the observed rainfall amounts in the future 6 hours as input, which may not be available in reality. But with the development of model meteorology, the rainfall forecast is already very near to the real rainfall, such that our simplification would not make a substantial deviation.

In the future, one of the possible improvements to our work is the involvement of the topological structure of the channel network. It enables the consideration of the mutual influence between channel components for our model and might achieve higher accuracy.

Author Contributions

All authors contributed equally to the reproduction of the original MLSTM-FCN and MALSTM-FCN model. In addition, Tianbo Yu prepared the dataset, Tianyang Ding and Junyu Han did several modifications on the original models including implementation of early stopping and different activation functions.

All authors contributed equally to the sections of the report. In addition, Tianbo Yu did the literature review, Junyu Han performed the experiments and collected the results, Tianyang Ding wrote the manuscript. All authors discussed the results and commented on the manuscript.

Acknowledgments

This project is a course project instructed by Dr. Shuiwang Ji from Department of Computer Science & Engineering, Texas A&M University. We thank Dr. Ji for his insight and expertise that greatly assisted the research.

We thank Dr. Ali Mostafavi from Zachry Department of Civil Engineering, Texas A&M University for providing raw data we used in Harris County Flood Case Study.

We would also like to show our gratitude to Dr. Anxiao Jiang from Department of Computer Science & Engineering, Texas A&M University for the GPU resourse.

References

[1] Kun Xie, Kaan Ozbay, Yuan Zhu, and Hong Yang. Evacuation zone modeling under climate change: A data-driven method. *Journal of Infrastructure Systems*, 23(4):04017013, 2017.

[2] Amir Mosavi, Yatish Bathla, and Annamária Varkonyi-Koczy. Predicting the future using web knowledge: State of the art survey. In *International conference on global research and education*, pages 341–349. Springer, 2017.

[3] Robert JC Burnash, R Larry Ferral, and Robert A McGuire. *A generalized streamflow simulation system: Conceptual modeling for digital computers*. US Department of Commerce, National Weather Service, and State of California . . . , 1973.

[4] Keith J BEVEN and Michael J Kirkby. A physically based, variable contributing area model of basin hydrology/un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, 24(1):43–69, 1979.

[5] R-J Zhao. The xinanjiang model. In *Proceedings of the Oxford Symposium*, 1980.

[6] RJ Moore. The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, 30(2):273–297, 1985.

[7] WS Luo, CQ Hu, and JT Han. Research on a runoff yield reflecting excess infiltration and excess storage simultaneously. *Journal of Soil and Water Conservation*, 6:6–13, 1992.

[8] Lianfa Li, Jinfeng Wang, Hareton Leung, and Chengsheng Jiang. Assessment of catastrophic risk using bayesian network constructed from domain knowledge and spatial data. *Risk Analysis: An International Journal*, 30(7):1157–1175, 2010.

[9] Shasha Han and Paulin Coulibaly. Bayesian flood forecasting methods: A review. *Journal of Hydrology*, 551:340–351, 2017.

[10] Hongxiang Yan and Hamid Moradkhani. Toward more robust extreme flood prediction by bayesian hierarchical and multimodeling. *Natural Hazards*, 81(1):203–225, 2016.

[11] Shanhu Jiang, Liliang Ren, Chong-Yu Xu, Shuya Liu, Fei Yuan, and Xiaoli Yang. Quantifying multi-source uncertainties in multi-model predictions using the bayesian model averaging scheme. *Hydrology Research*, 49(3):954–970, 2017.

[12] Wenbo Huo, Zhijia Li, Jingfeng Wang, Cheng Yao, Ke Zhang, and Yingchun Huang. Multiple hydrological models comparison and an improved bayesian model averaging approach for ensemble prediction over semi-humid regions. *Stochastic environmental research and risk assessment*, 33(1):217–238, 2019.

[13] Adrian E Raftery and Yingye Zheng. Discussion: Performance of bayesian model averaging. *Journal of the American Statistical Association*, 98(464):931–938, 2003.

[14] Newsha K Ajami and Chuanhui Gu. Complexity in microbial metabolic processes in soil nitrogen modeling: a case for model averaging. *Stochastic Environmental Research and Risk Assessment*, 24(6):831–844, 2010.

[15] KM Mok, KV Yuen, KI Hoi, KM Chao, and D Lopes. Predicting ground-level ozone concentrations by adaptive bayesian model averaging of statistical seasonal models. *Stochastic environmental research and risk assessment*, 32(5):1283–1297, 2018.

[16] Saeideh Samani, Asghar Asghari Moghaddam, and Ming Ye. Investigating the effect of complexity on groundwater flow modeling uncertainty. *Stochastic environmental research and risk assessment*, 32(3):643–659, 2018.

[17] Kuai Fang, Chaopeng Shen, Daniel Kifer, and Xiao Yang. Prolongation of smap to spatiotemporally seamless coverage of continental us using a deep learning neural network. *Geophysical Research Letters*, 44(21):11–030, 2017.

[18] Kei Hiroi and Nobuo Kawaguchi. Floodeye: Real-time flash flood prediction system for urban complex water flow. In *2016 IEEE SENSORS*, pages 1–3. IEEE, 2016.

[19] Zhaoyang Liu, Weigang Xu, Jun Feng, Shivakumara Palaiahnakote, Tong Lu, et al. Context-aware attention lstm network for flood prediction. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 1301–1306. IEEE, 2018.

[20] Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. Recurrent neural networks for multivariate time series with missing values. *Scientific reports*, 8(1):6085, 2018.

[21] Yulia Rubanova, Ricky TQ Chen, and David Duvenaud. Latent odes for irregularly-sampled time series. *arXiv preprint arXiv:1907.03907*, 2019.

[22] Edward De Brouwer, Jaak Simm, Adam Arany, and Yves Moreau. Gru-ode-bayes: Continuous modeling of sporadically-observed time series. *arXiv preprint arXiv:1905.12374*, 2019.

[23] Fazle Karim, Somshubra Majumdar, Houshang Darabi, and Samuel Harford. Multivariate lstm-fcns for time series classification. *Neural Networks*, 116:237–245, 2019.

[24] Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 international joint conference on neural networks (IJCNN)*, pages 1578–1585. IEEE, 2017.

[25] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[26] Fazle Karim, Somshubra Majumdar, Houshang Darabi, and Shun Chen. Lstm fully convolutional networks for time series classification. *IEEE Access*, 6:1662–1669, 2017.

[27] Patrick Schäfer and Ulf Leser. Multivariate time series classification with weasel+ muse. *arXiv preprint arXiv:1711.11343*, 2017.