Machine Learning Models in Water Level Prediction. A Case Study of the Han River

Akerke Galymkyzy*

*Seoul National University, Department of Civil and Environmental Engineering

Abstract

Flooding is a significant natural disaster that can have devastating effects on human lives, infrastructure, and the environment. Accurate and timely flood prediction plays a vital role in mitigating these impacts. This research paper presents a case study focused on flood prediction in the Han River, employing and comparing the mean squared error (MSE) of three different models: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Long Time Series Forecasting Linear (LTSF Linear) models. The study utilizes historical data on water levels from monitoring stations along the Han River. The data is divided into training, validation, and testing sets to develop and evaluate the performance of prediction models. The predictive capabilities are then evaluated using the testing dataset. The MSE is employed as a performance metric to assess the accuracy and reliability of each model's flood predictions, with lower MSE indicating better predictive performance. Overall, this research paper contributes to the advancement of flood prediction techniques by comparing the MSE of LSTM, GRU, and LTSF Linear models in the specific context of the Han River. The insights gained from this case study can aid in developing robust flood prediction models and implementing effective flood mitigation strategies in the future.

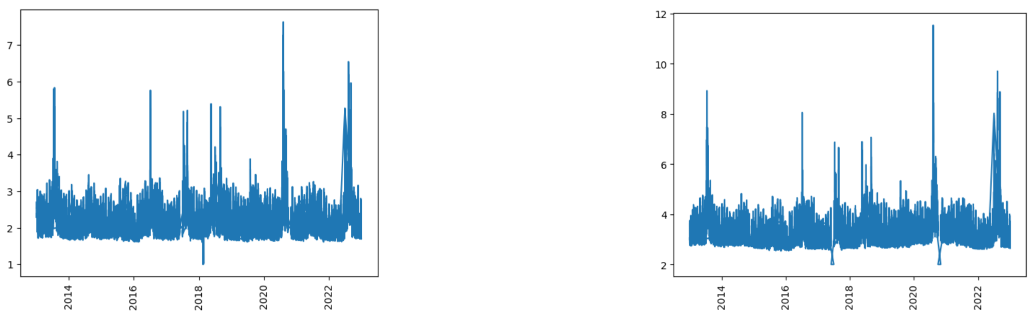
Keywords: flood prediction, machine learning models, LSTM, GRU, LTSF - Linear

1. Introduction

A natural disaster – flood – has tremendous consequences which should be carefully controlled, and immediate actions should be taken when rivers start rising. This causes fatalities and economic damage; it can be prevented by predicting the water level. According to news reports in 2022, the Han River in Seoul, the capital of the Republic of Korea, experienced a flood due to high precipitation. According to the paper review done (Antwi-Agyakwa et. al, 2023) on flood risk prediction tools, the number of research publications in this area annually grew at least by 43 articles between 2017-2021, while in the first quarter of 2022, it was 77. Machine learning (ML) approaches like Artificial Neural Networks (ANN) and Least Squares Support Vector Machine (LSSVM) were widely used in flood prediction for both qualitative and quantitative modeling of hydrological variables in non-linear streamflow. A recent popular method is a recurrent neural network (RNN) which includes Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU). RNNs became the leading tool as it can save the short-past and long-past data for complex and highly non-linear time series forecasting (TSF). Both LSTM and GRU go alongside, however, GRU has demonstrated a simpler structure and a higher operation compared to LSTM (Guo et. al, 2021a). Recent publication about long term series forecasting (LTSF) on TSF has shown simpler, faster, and more precise results. The research done comparing Transformer-based LTSF and LTSF-Linear, the former is a widely used method, but high effectiveness proved for the latter and purposed as a new further research area (Zeng et. al, 2022). This graduation thesis focuses on 3 machine learning tools: LTSF-Linear, GRU, and LSTM which were found to be most prominent according to the literature review.

2. Methods and Methodology

The Google Colab is used as an environment to proceed with the modeling with Python programming language. PyTorch and TensorFlow are deep learning frameworks for ML modeling. The training and test data will be taken from the official website of the Han River Flood Control Office. The modeling tools are described as follows. The data uses the water level from the basins at Haengju Bridge and Jamsu Bridge every 10 minutes from January 2012 to December 2022.



(a) Haengju Bridge (b) Jamsu Bridge

Fig 1. Water Level at stations over the period

2.1.1 LSTM

LSTM is one of the RNNs, which was improved to resolve gradient issues in long-term memory and backpropagation. It has three gates, named forget gate, input gate, and output gate to maintain and fix cell and hidden states. The former decides what information should be deleted from the cell state. The middle one determines what information should be used to update the cell state, while the latter one manages information saved in the current cell state and goes into the new hidden state (Guo et. al, 2021b).

2.1.2 GRU

GRU model is a modified version of LSTM with a simpler structure. It has only two control gates: the reset gate and the update gate. The latter manages the data flow of the previous step to the next one. The former gate controls the importance of information from the previous state. In comparison to LSTM, GRU shows better performance in modeling time and parameter updates (Guo et. al, 2021c).

2.1.3 LTSF-Linear

LTSF-Linear is a regression-type ML tool of linear models with one layer used in TSF. It distributes weights for variates and does not build spatial correlation. It has several types, but for this research, NLinear is considered. NLinear enhances the performance of LSTF-Linear, when a dataset has a distribution shift, by removing the last input value of the sequence and passing through a linear layer. This input is returned before making the final prediction. This process is primitive normalization for sequencing inputs. In long term TSF LTSF-Linear showed better results as a DMS forecasting baseline than a transformer-based LTSF according to the studies by Zeng et.al in 2022.

2.2 Forecast Performance measurements

In this research two forecast performance measurements are used: mean square error (MSE) and mean absolute value (MSE). This value is calculated directly during the train and test state.

2.2.1 Mean Square Error

Comparing the MSE values obtained from different regression models helps to analyze performance. A lower MSE indicates better performance, as it indicates that the model's predictions are closer to the actual target values. If one model has a significantly lower MSE compared to others, it suggests that it is better at capturing the underlying patterns in the data and making accurate predictions according to PyTorch documentation.

$$RMSE = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - Y_p\right)^2 \quad (1)$$

Here Y_t indicates real values, while Y_p indicates predicted data. T is a number of time intervals.

3. Case Study

Below is the description of the case used in this study. The study was conducted using the data from the Haengju Bridge.

3.1 Case Description

The Han River or Hangang is a major river in South Korea and the fourth longest river on the Korean peninsula. The river begins as two smaller rivers in the eastern mountains of the Korean peninsula, which then converge near Seoul, the capital of the country. The total length of the Han River is approximately 494 kilometers. Although it is not a long river, the lower Hangang is sufficiently wide for such a relatively short river. Within Seoul City, the river is estimated to be more than 1 kilometer (0.62 mi) wide (Wikipedia, 2023). The data is taken from Haengju (upper river side) and Jamsu (lower river side) Bridges. The distance between them is approximately 19.5 km according to the Naver Map measurements.

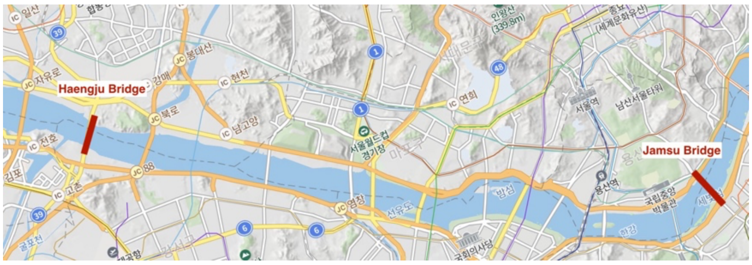


Fig 2. Locations of Haengju and Jamsu Bridges

3.2 Parameters setting

The data used water level as an input parameter as other data was difficult to obtain. The train data was taken from the beginning of the period, January 1st of 2013, till the end of 2020. The test data was from January 1st to December 31st of 2022. The data also has a validation set of the one-year period in 2021. The sequence length was used as 1000 minutes, which is approximately 16,67 hours. Then the data for 2023 until June was predicted and compared with ground data.

4. Results

The following was obtained for each three models. Moreover, hyperparameters of each model, including epoch size, learning rate, and optimizer type, were set differently while running the models.

4.1 Long Short Term Memory

LSTM took about 8 minutes to finish modeling for an epoch size of 5. Adam optimizer and learning rate of 0,01 were found to be optimal for good results. LSTM scored a mean value error of 0,01 and 0,056 for Haengju and Jamsu bridges, respectively. While running it several times, the highest error was around 0,12 but each time the lowest was around 0,01 for Haengju bridge.

4.2 Gated Recurrent Unit

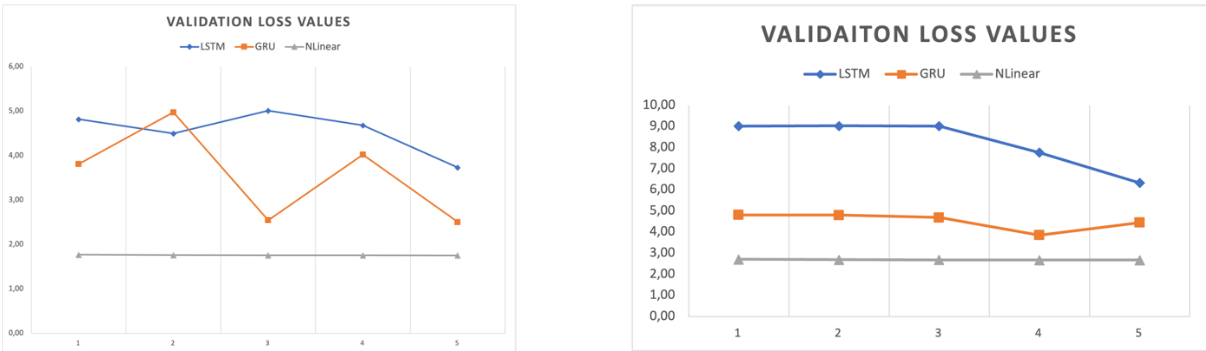
GRU performed 4 minutes with 5 epochs for each bridge. Adamax optimizer with a learning rate of 0,01 was found to be effective. The MSE score was 0,034 and 0,0065 for Haengju and Jamsu bridges, respectively.

4.3 Linear Time Series Forecast – Linear

LTSF – Linear scored the lowest mean square value of 0,0073 and 0,0039 for Haengju and Jamsu bridges, respectively, just in 1,5 minutes for each with 5 epochs. Stochastic gradient descent (SGD) was used to model with a learning rate of 0,01. It is important to note that Adam optimizer was not effective. Also, since the initial MSE of SGD optimizer was small, the following epochs were decreasing in MSE value by a fraction of 1000, therefore, 5 epochs were enough.

	LSTM %	GRU %	LTSF – Linear %
Haengju Bridge	1.02	0.65	0.39
Jamsu Bridge	5.66	3.14	0.73

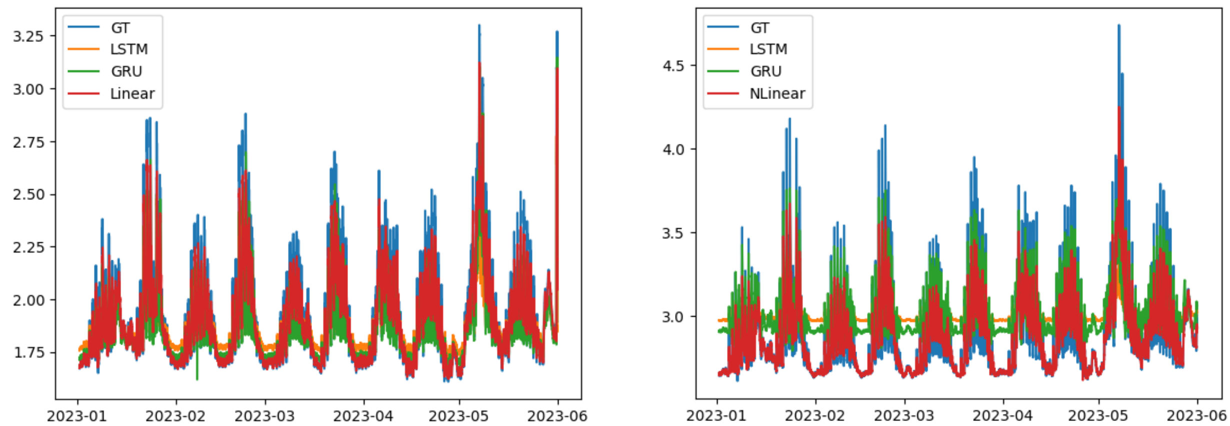
Table 1. MSE values of validation loss of 3 models in % for two bridges



(a) Haengju Bridge

(b) Jamsu Bridge

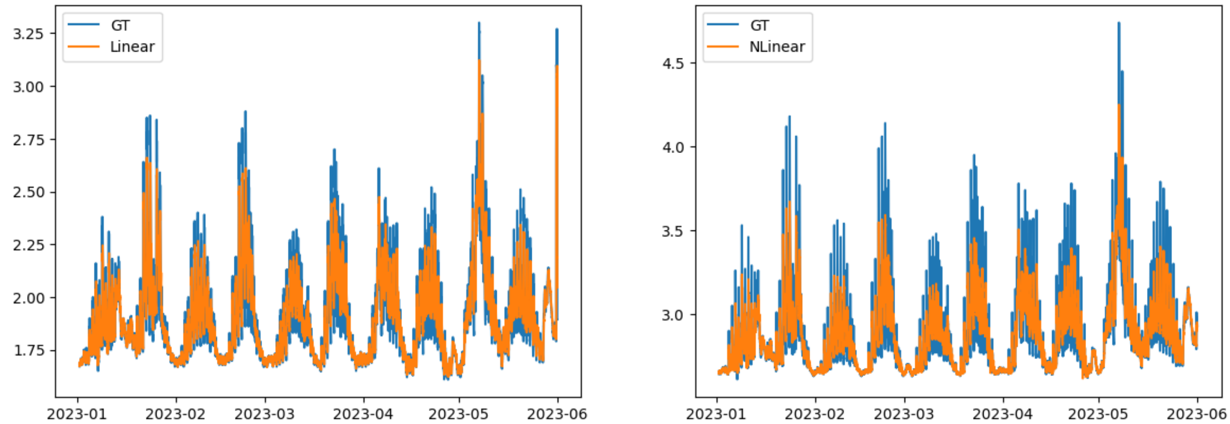
Fig 3. Validation Loss of three models



(a) Haengju Bridge

(b) Jamsu Bridge

Fig 4. Water level prediction for 2023



(a) Haengju Bridge (b) Jamsu Bridge

Fig 5. Water level prediction for 2023 of LTSF-Linear with ground truth value

5. Discussion

As expected, the LSTF – Linear was found to be the fastest and most accurate model to predict the water level through the time series forecast. However, surprisingly GRU scored just above the LSTF – Linear after several epochs, which means that more iterations with more epochs are suggested. However, it takes more time, as we should also find the optimal epoch size with best-suiting parameters, as the model can overfit. There were some missing data on water levels, especially, for those periods when the flood occurred.

5.1 Limitations

There are several limitations in this work. First of all, the data was accessed through the request from the center through the API key. However, the models lack parameters, only the water level variable was used, while other research papers suggest correlations between water temperature and precipitation (Supatmi et. al, 2019). The data has null values which were detected during the modeling which can affect the prediction. Another point to mention is the sequence time which is 16,67 hours. This is quite a short time to get prepared for a flood if there is a high probability of it. Therefore, a longer period is suggested. Finally, the data is not enough as it covers only the past 10 years.

6. Conclusion

Among machine learning tools to predict time series LSTF – Linear showed high performance. As LSTF – Linear is a suggestion for new research from previous papers, it can be further implemented on a bigger scale.

Reference

Antwi-Agyakwa, K. T., Afenyo, M., & Angnuureng, D. B. (2023). Know to Predict, Forecast to Warn: A Review of Flood Risk Prediction Tools. *Water*, *15*(3), 427. https://doi.org/10.3390/w15030427

Guo, Y., Yu, X., Xu, Y., Chen, H., Gu, H., & Xie, J. (2021). AI-based techniques for multi-step streamflow forecasts: application for multi-objective reservoir operation optimization and performance assessment. *Hydrology and Earth System Sciences*, *25*(11), 5951–5979. https://doi.org/10.5194/hess-25-5951-2021

MSELoss — PyTorch 2.0 documentation. (n.d.). https://pytorch.org/docs/stable/generated/torch.nn.MSELoss.html

South Korea's capital Seoul hit with serious floods. (2022, August 9). BBC Newsround. https://www.bbc.co.uk/newsround/62479716

Supatmi, S., Huo, R., & Sumitra, I. D. (2019). Implementation of Multiplicative Seasonal ARIMA Modeling and Flood Prediction Based on Long-Term Time Series Data in Indonesia. In *Lecture Notes in Computer Science*. Springer Science+Business Media. https://doi.org/10.1007/978-3-030-24265-7_4

Zeng, A., Chen, M., Zhang, L., & Xu, Q. (2022). Are Transformers Effective for Time Series Forecasting? *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2205.13504