# Sensitivity analysis of network structure in missing streamflow data complementation using Bidirectional Long and Short-Term Memory network

Takeyoshi Nagasato, Kei Ishida, Daiju Sakaguchi,

Masato Kiyama, and Motoki Amagasaki

Kumamoto University, Japan

# 1. Introduction

- ◆ Streamflow data based on the observation may be partially missing due to flood or malfunction of the measuring equipment.
- ◆Linear regression and multiple regression models have been proposed as methods for complementing missing flow rates[1, 2].
- ◆There is a possibility that the missing flow rate can be complemented with high accuracy by using the deep learning method.

⬇

- ◆Therefore, this study implemented deep learning for missing streamflow complementation.
- ◆This study conducted a sensitivity analysis of the network structure.

# 2. Methodology

- Among the deep learning methods, Bidirectional LSTM (Bi-LSTM) was implemented.
- LSTM is a kind of neural network that can learn long-term dependence of time series data.
- Bi-LSTM learns data in both forward and backward directions, compared to Unidirectional LSTM which learns data forward directions (Figure 1, 2) .

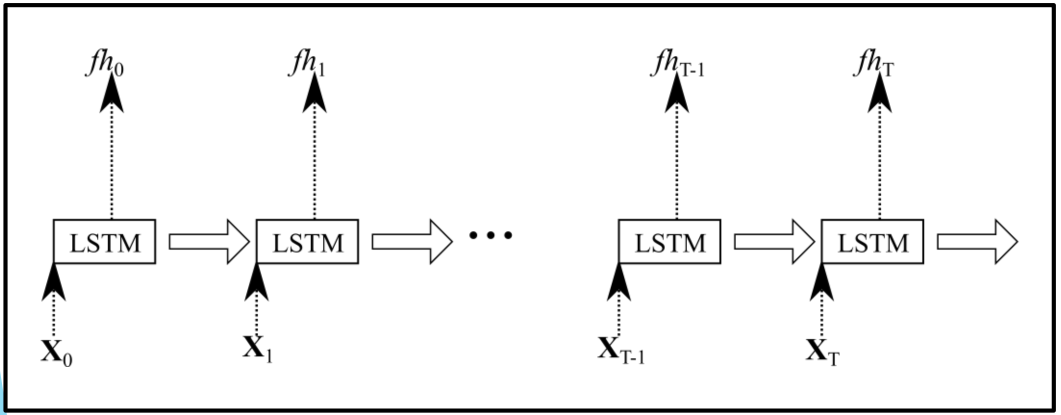


Figure 1 Uni-LSTM

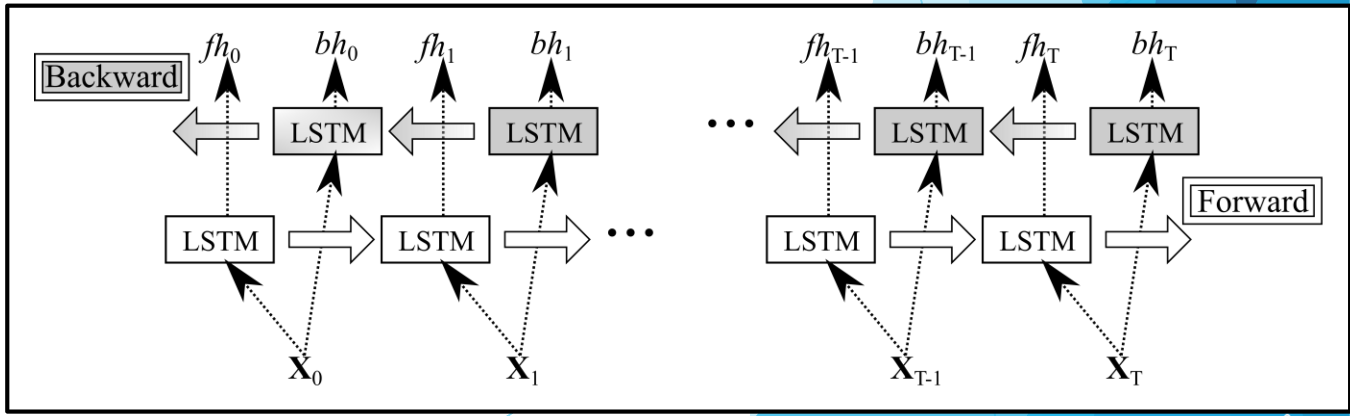


Figure 2 Bi-LSTM

# 3. Study Area and Dataset

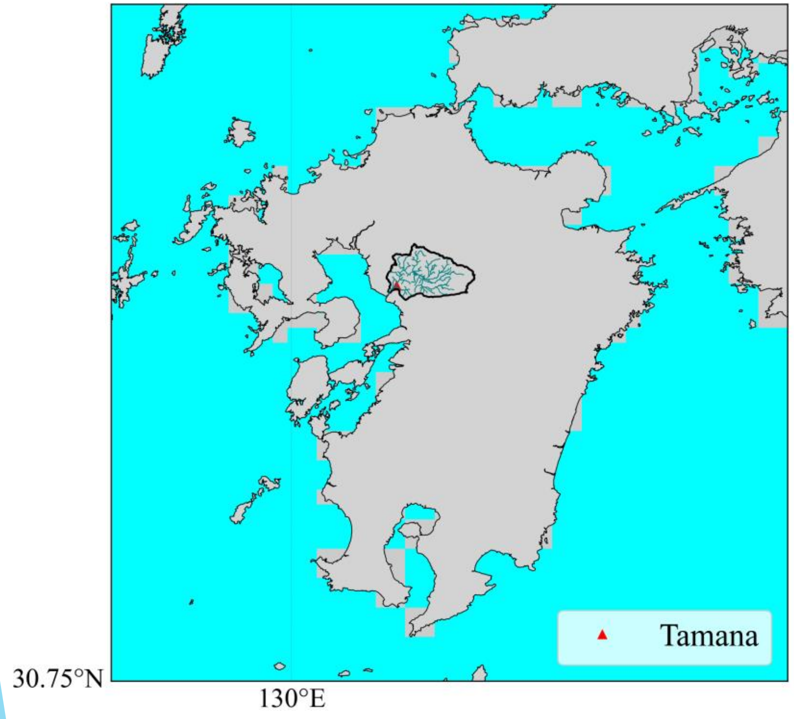


Figure 3 Kikuchi river

- Kikuchi River in Japan was selected as the study area (Figure3).
- Hourly flow data observed at Tamana station were obtained from WIS.
- Hourly precipitation data was obtained from the rainfall analyzed by the Japan Meteorological Agency.

# 4. Model Implementation

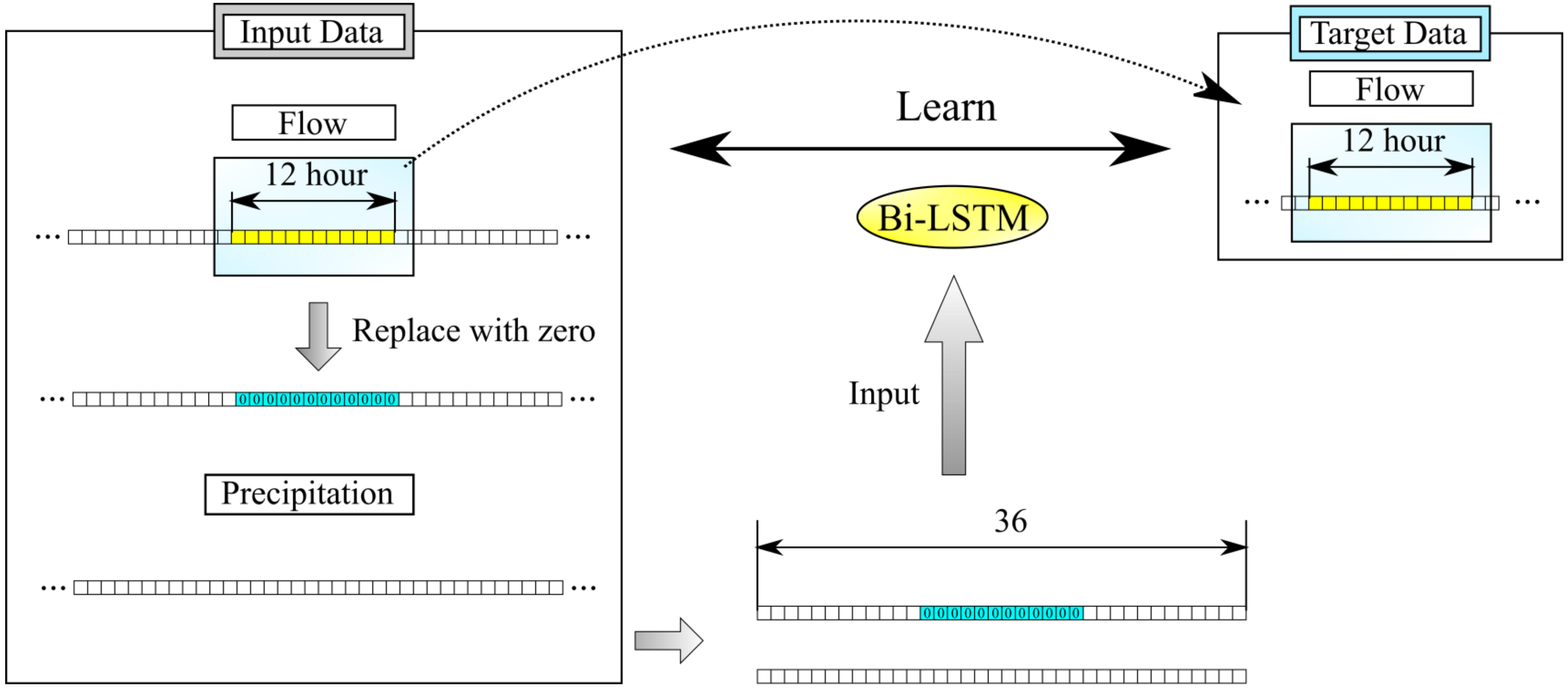


Figure 4 training dataset

- Missing flow data were artificially generated by replacing the non-missing part of the flow data with zero (Figure 4).
- The length of the flow data used for input before and after the missing was 12 hours.

# 4. Model Implementation

- The dataset is divided into three datasets: the training dataset (2006-2008), the validation dataset (2009-2010), and the test dataset (2011-2012).
- For the evaluation of model accuracy this study selected the Nash-Sutcliffe efficiency (NSE), RMSE, RMSE99.
- In addition, Time/Epoch was used to compare the computational costs.
- The training processes with the training and validation datasets is repeatedly conducted 50 times.
- This study compared the estimation accuracy of 50 learning results using a box plot.

# 4. Model Implementation

- This study conducted a sensitivity analysis on the effect of the network on the estimation accuracy (Table 1).

Table 1. Setting conditions

| Case | Number of LSTM Layer | Input | Length of missing time |
|---|---|---|---|
| 1 | 1 | Precipitation, Flow | 12h |
| 2 | 2 | Precipitation, Flow | 12h |
| 3 | 3 | Precipitation, Flow | 12h |

# 5. Results and Discussion

## Effects of the network structure on the accuracy (NSE, RMSE)

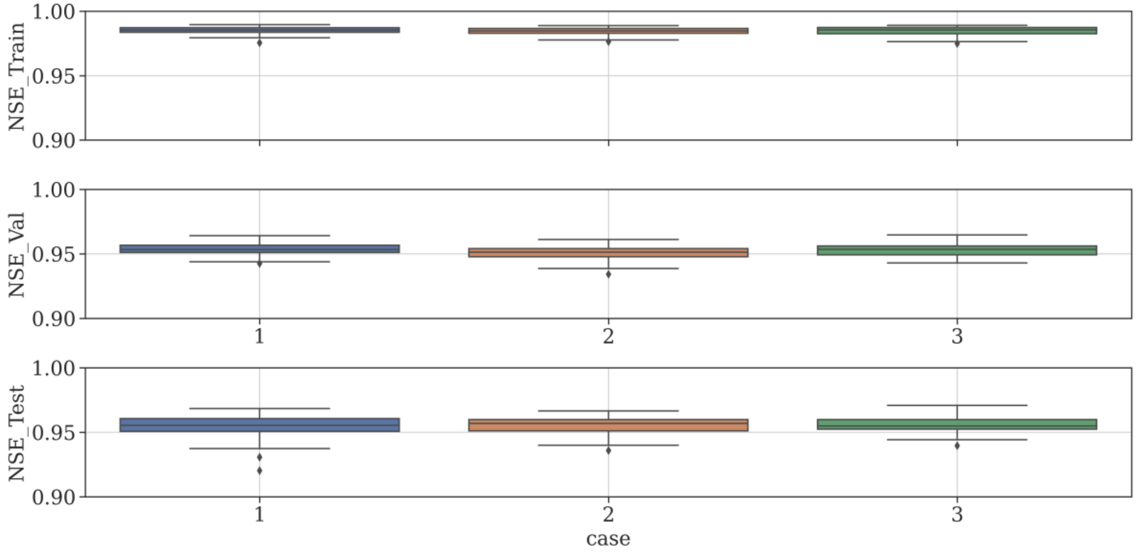


Figure 5 Result (NSE)

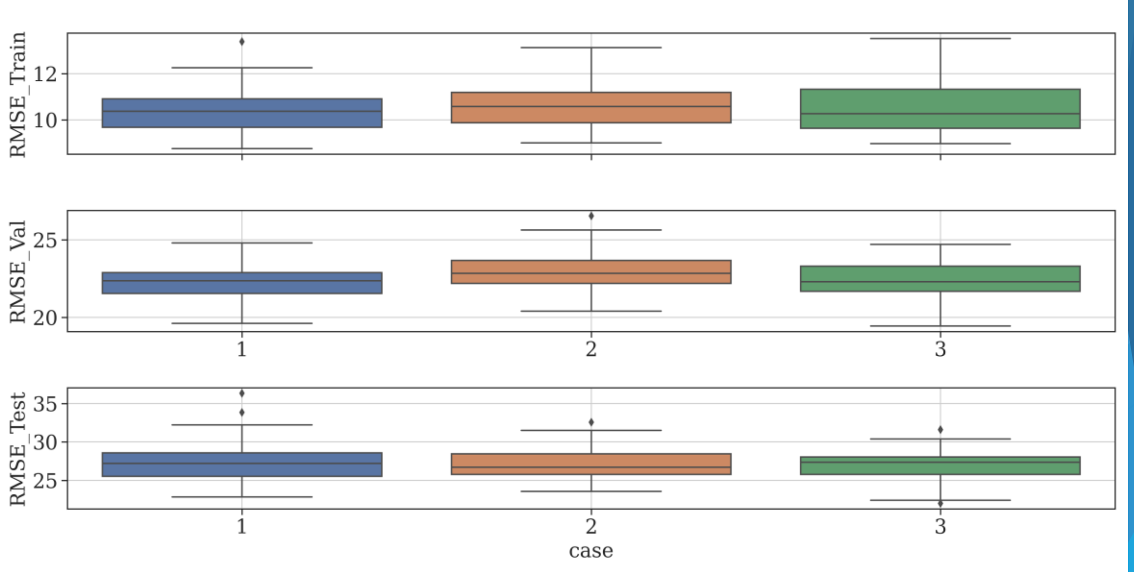


Figure 6 Result (RMSE)

- ◆Bi-LSTM has potential to estimate the missing flow data with high accuracy(NSE = 0.9 or more).
- ◆Even if the structure of LSTM deeper, the estimation accuracy may not be improved.

# 5. Results and Discussion

## Effects of the network structure on the accuracy (RMSE99)

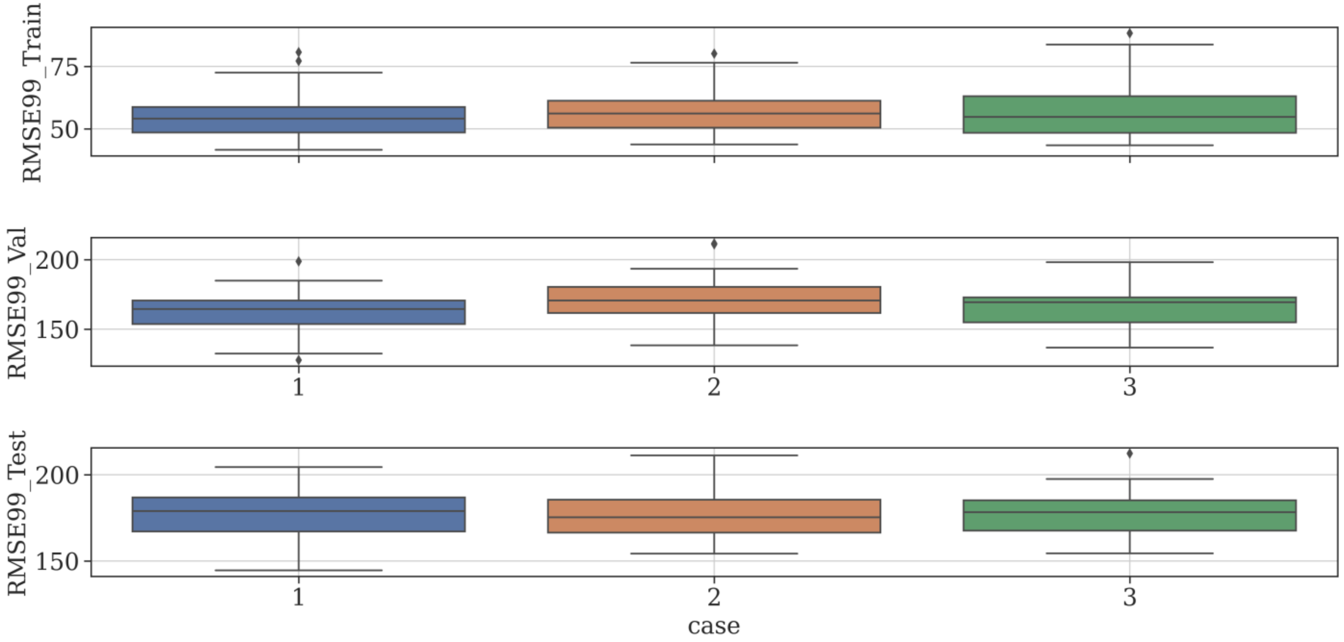


Figure 7 Result (RMSE99)

# 5. Results and Discussion

## Effects of the network structure on the computational costs

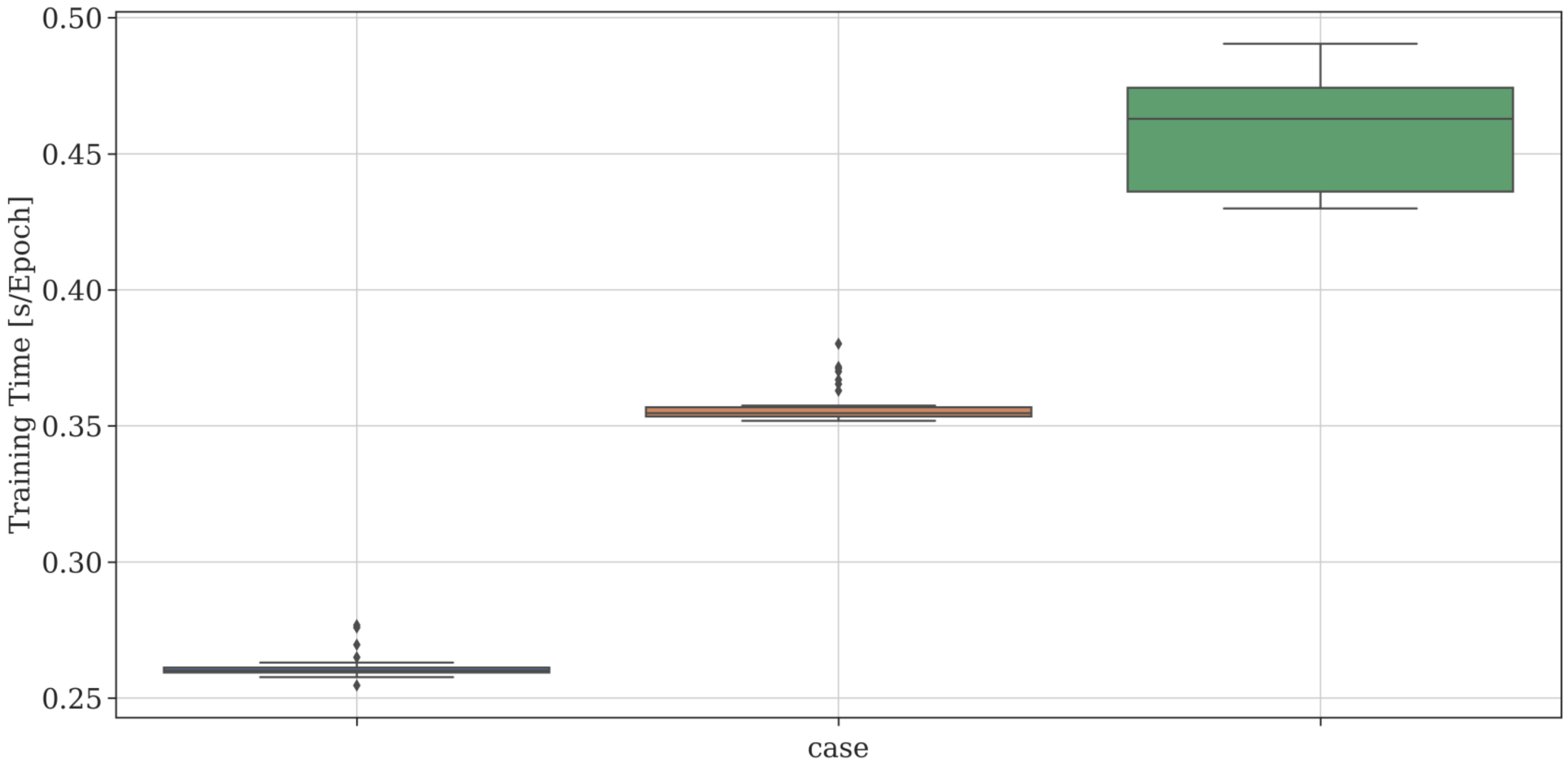


Figure 8 Result

# 6. Conclusion

- This study implemented Bidirectional LSTM (Bi-LSTM) to complement missing flow data and conducted a sensitivity analysis of the network structure.
- To complement the hourly missing flow data, the hourly flow data and basin average precipitation data are used as the input.
- Bi-LSTM has potential to estimate the missing flow rate with high accuracy(NSE = 0.9 or more).
- Even if the structure of LSTM deeper, the estimation accuracy may not be improved.
- Considered the computational cost, optimizing the network structure is important.

# References

1. KOJIRI, T., PANU, U. S. & TOMOSUGI, K. Complement Method of Observation Lack of Discharge with Pattern Classification and Fuzzy Inference. J. Japan Soc. Hydrol. Water Resour. 7, 536–543 (1994).
2. Mfwango, L. H., Salim, C. J. & Kazumba, S. Estimation of Missing River Flow Data for Hydrologic Analysis: The Case of Great Ruaha River Catchment. (2018) doi:10.4172/2157-7587.1000299.