water

Article

Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction

Minwoo Cho, Changsu Kim, Kwanyoung Jung and Hoekyung Jung *

Department of Computer Science and Engineering, Paichai University, 155-40 Baejae-ro, Daejeon 35345, Korea; kcjmw1208@kakao.com (M.C.); ddoja@pcu.ac.kr (C.K.); jky5036@korea.kr (K.J.)
* Correspondence: hkjung@pcu.ac.kr; Tel.: +82-425205640

Abstract: The damage caused by floods is increasing worldwide, and if floods can be predicted, the economic and human losses from floods can be reduced. A key parameter of flooding is water level data, and this paper proposes a water level prediction model using long short-term memory (LSTM) and a gated recurrent unit (GRU). As variables used as input data, meteorological data, including upstream and downstream water level, temperature, humidity, and precipitation, were used. The best results were obtained when the LSTM–GRU-based model and the Automated Synoptic Observing System (ASOS) meteorological data were included in the input data when experiments were performed with various model structures and different input data formats. As a result of the experiment, the mean squared error (MSE) value was 3.92, the Nash–Sutcliffe coefficient of efficiency (NSE) value was 0.942, and the mean absolute error (MAE) value was 2.22, the highest result in all cases. In addition, the test data included the historical maximum water level of 3552.38 cm in the study area, and the maximum water level error was also recorded as 55.49, the lowest result. Through this paper, it was possible to confirm the performance difference according to the composition of the input data and the time series prediction model. In a future study, we plan to implement a flood risk management system that can use the predicted water level to determine the risk of flooding, and evacuate in advance.

Keywords: water level prediction; long short-term memory (LSTM); gated recurrent unit (GRU); meteorology data

check for updates

Citation: Cho, M.; Kim, C.; Jung, K.; Jung, H. Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction. *Water* **2022**, *14*, 2221. https://doi.org/10.3390/w14142221

Academic Editor: Chang Huang

Received: 30 June 2022
Accepted: 12 July 2022
Published: 14 July 2022

1. Introduction

Around the world, damage from natural disasters is increasing due to the abnormal climate caused by global warming, and it is expected that the frequency and intensity of natural disasters will continue to increase [1,2]. Compared to the 20 years from 1980 to 1999, both the number of and amount of damage caused by natural disasters during the 20-year period from 2000 to 2019 has increased. In particular, the increase in damage from floods and storms was higher than that of other natural disasters. Figure 1 compares the amount of damage caused by natural disasters during the period from 1980 to 1999 and from 2000 to 2019 [3]. In the case of the United States, the loss due to natural disasters occurred at an all-time high in 2017, with more than $306 billion in losses due to hurricanes and heavy rains [4].

In order to reduce the damage caused by flooding, it is necessary to accurately predict the flood and evacuate people and property located in the flood-damaged area at an appropriate time [5,6]. However, flood prediction has many difficulties in forecasting because there are many variables to consider, and each element has spatial and temporal correlations. Methods for flood prediction are continuously being studied, and there are mainly hydrological models and data-driven intelligent models.

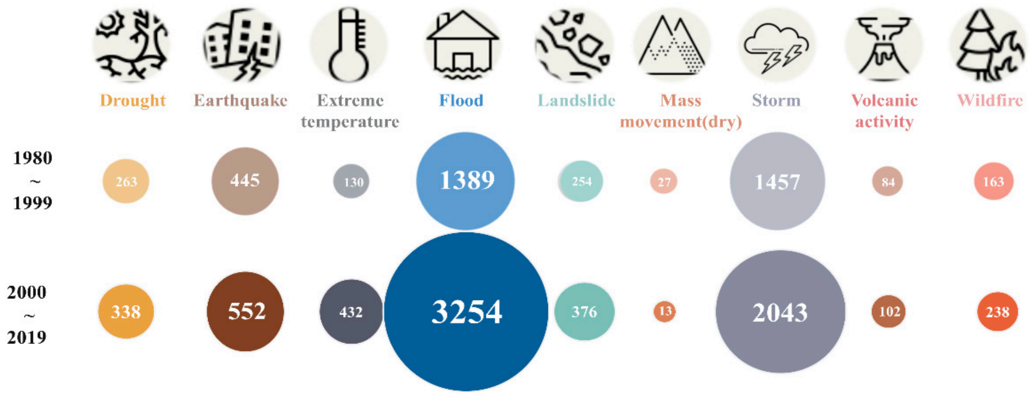


Figure 1. Total disaster events by type: 1980–1999 vs. 2000–2019 from Ref. [3].

In the case of the hydrological model, it is a model that analyzes the characteristics of the hydrological gate and physically describes the discharge confluence point. It is based on the theory of fluid mechanics and is a method for researchers to derive a confluence equation by combining physical laws such as mass, momentum, and conservation of energy. It has been used in the past for flood prediction and management even before studies through machine learning and deep learning were actively conducted, and various studies have been conducted using the model [7–12]. However, the model requires a deep hydrological knowledge of the researcher, and there is variability due to erosion of the topography over time, so it is difficult to use it in the long term. In addition, it is difficult to construct input data with various variables, and low prediction accuracy is obtained due to non-linearity. Furthermore, since each river has different characteristics of the watershed, it is necessary to create an individual model applied to each river, so there is a lack of versatility [13].

In the case of a data-based intelligent model, it is a method of predicting water level and runoff through data analysis based on observed data. Data analysis and data mining technologies are being researched in many fields and are rapidly developing [14]. Regardless of fields such as medicine [15], economy [16,17], environment [18], and tourism [19,20], a great deal of research is being conducted by applying machine learning and deep learning, and it has been confirmed that it shows excellent performance. One of the data-based intelligent models applied in the hydrological field for predicting and managing floods is the artificial neural network (ANN) model. The ANN model developed by [21] uses water level and meteorological data as input data and estimates the water flow by applying harmony search (HS) [22] and differential evolution (DE) [23]. By utilizing HS and DE, overfitting was prevented by updating the parameters of the architecture and selecting important features. Compared with the radial basis function neural network (BRFNN) [24] and multi-layer perceptron (MLP) [25] models, ANN with HS and DE was confirmed to show good performance, and it was proved that the ANN model can be used for water flow prediction. In addition, many studies have been conducted to predict key factors of flooding using models such as ANN-based water level prediction models and runoff prediction models [26–30]. However, most of the related studies utilize hydrological data and meteorological data, which are categories of time series data, as input data [31,32]. In the case of the ANN model, there is a problem of insufficient memory when operating on sequential data and time series data and it is difficult to find an optimal parameter in the learning process [33,34].

Recently, deep neural network (DNN) data-based intelligent models with two or more hidden layers that have improved some problems of the ANN model have been extensively studied. In this case, LSTM and GRU models based on recurrent neural networks (RNN), which show excellent performance in time series data processing, are being widely used. Le et al. [35] proposed an LSTM model for flood forecasting. In this study, daily runoff and rainfall data were used as input datasets, and a performance comparison experiment was conducted according to the input data of the model. An experiment was conducted in the Da River basin of Vietnam, and the performance comparison experiment of the model according to the input data was conducted using the runoff and precipitation data sets. The output of the model was configured to predict the amount of runoff for 1, 2, and 3 days into the future. The performance evaluation was conducted using the Nash–Sutcliffe efficiency (NSE) and root mean squared error (RMSE) indicators [36]. Finally, the NSE value of the first day data was around 99.1%, the NSE value of the second day was 94.9%, and the 3rd day NSE value was 87.2%, so it was possible to obtain more than 87.2% accuracy for all output data. Through this study, it was confirmed that the LSTM model, which is widely used for time series data processing, performed well enough for flood prediction. In addition, RNN-based LSTM and GRU models are continuously being studied for flood prediction and are showing good performance as they are specialized models for time series data prediction [37–42]. In addition, there are many studies using artificial intelligence such as ANN and RNN for flood management which consider the ratio of training data and validation data [43] or which compare the performance according to the size of the input data of the model [44]. All of these studies greatly affect the performance of the model, and the difference in the current paper is that other types of meteorological data are used to conduct a performance comparison experiment according to the meteorological data. As a result of the experiment, the best data set for the study area obtained the best results when water level data and ASOS data were used. There are two or more types of meteorological data measured in most research areas. Therefore, as suggested in this paper, it is considered that there is a possibility that the performance of the model will be improved if appropriate meteorological data are used in the papers related to water level prediction.

In this paper, we propose an LSTM-GRU-based water level prediction model using water level data and meteorological data as input data. In addition, two types of weather data provided by the Korea Meteorological Administration, the Automated Synoptic Observing System (ASOS) data set and the Automatic Weather System (AWS) data set, are used to test which data are more suitable for the study area. In addition, for water level data, data from upstream and downstream of Yeojubo, Icheon-si, Gyeonggi province, Republic of Korea were used, and performance comparison experiments of the multi-LSTM model, multi-GRU model, and LSTM-GRU model were conducted. A total of three evaluation indicators were used: MSE, NSE, and MAE, and the maximum water level prediction error was additionally used. The LSTM-GRU model proposed in this paper showed the best performance among all cases. The mean squared error (MSE) value of the LSTM-GRU model was 3.92, the Nash–Sutcliffe coefficient of efficiency (NSE) value was 0.942, the mean absolute error (MAE) value was 2.22, and the maximum water level error was 55.49, confirming the best performance in all indicators.

Section 2 describes the study area, training data, and deep learning techniques used in model development. Experimental scenarios and model design are described in Section 3, and a model performance comparison is conducted in Section 4. Finally, Section 5 reviews and concludes the results of this study.

2. Methodology

2.1. Study Area

One of the main causes of flooding is localized heavy rain which is accompanied by a large amount of rainfall in a short period of time. To select a test area where extensive damage occurred due to rainfall, the data on the current situation of heavy rain damage by region, published by the Ministry of Public Administration and Security of the Republic of Korea, were used. The Ministry of Public Administration and Security is the central administrative agency of the Republic of Korea in charge of affairs related to disaster prevention, such as establishing, generalizing, and co-ordinating safety and disaster policies. The current status of heavy rain damage by region over the past 30 years was investigated. Figure 2 shows the amount of rain damage and the frequency of occurrence of rain damage for each city in Korea.

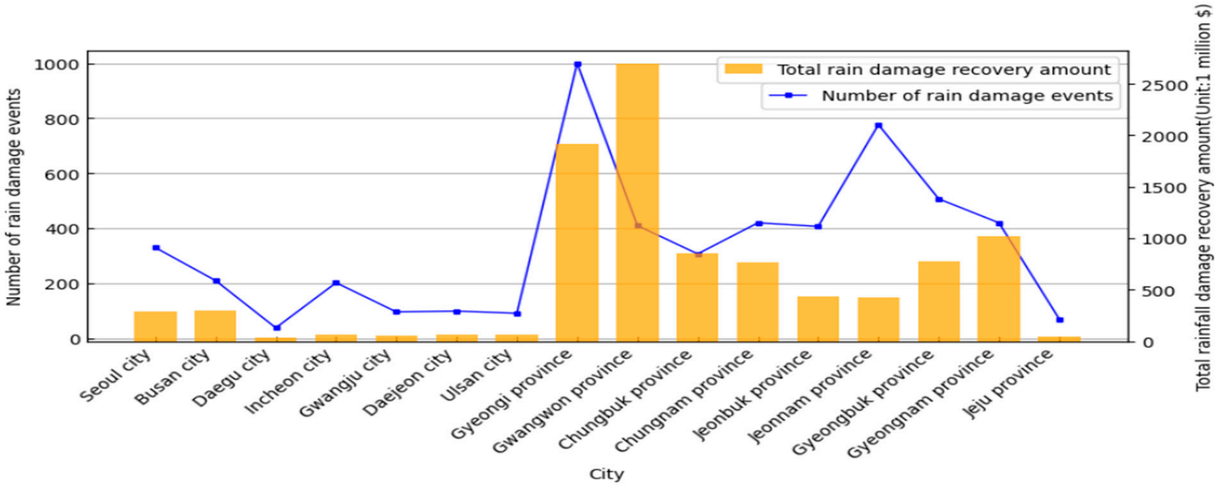


Figure 2. Number of rain damage events and total rain damage recovery amount (1988–2017).

In Gyeonggi province, the amount of rain damage recovery amount was the second highest among all cities, and in the case of occurrence frequency, rainfall occurred with the highest frequency among all cities. In addition, due to the geographical characteristics of the Gyeonggi province region, it surrounds the capital of the Republic of Korea, which has the highest population density, and a large population is also distributed in Gyeonggi-do. Figure 3 shows the population density and total population of each city as of 2020.

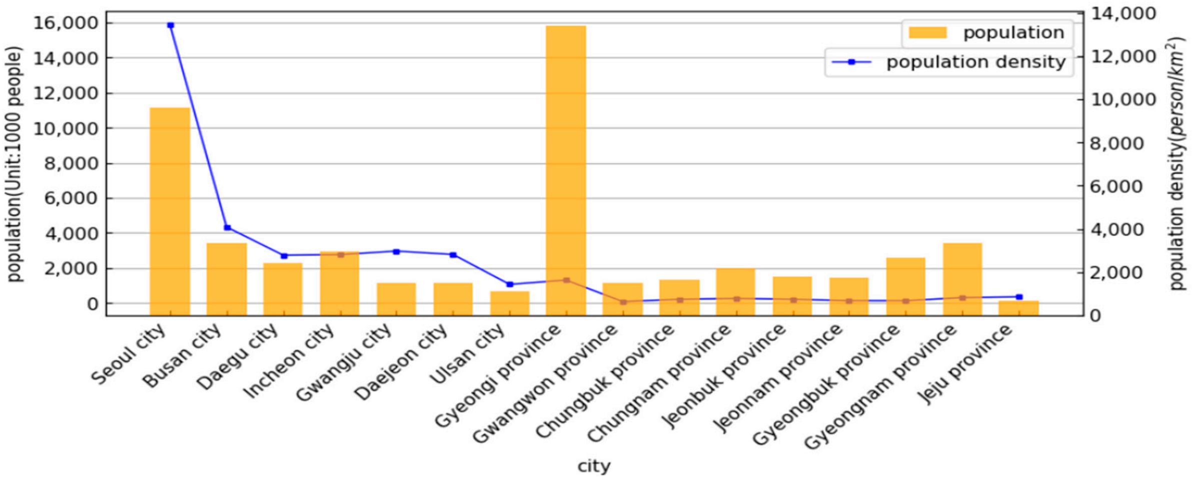
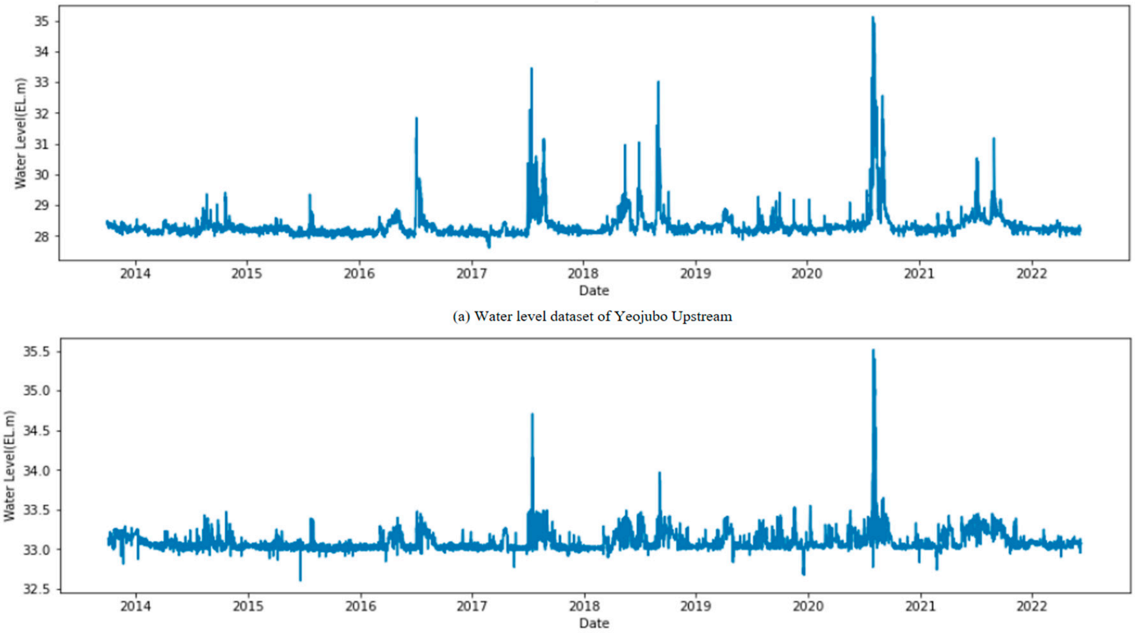


Figure 3. Total Population and Population Density by City/Province of Korea in 2020.

As described above, the Gyeonggi-do region, which is expected to receive extensive damage in case of heavy rain and flood, was selected as the test bed area. Due to the nature of rivers and the water flows from upstream to downstream, excessive runoff due to rainfall occurs, causing flooding in the downstream area. Yeojubo, in Gyeonggi-do, was selected as the final test bed for this study because there are measuring stations upstream and downstream of the city.

2.2. Meteorological Dataset and Water Level Dataset

For the water level dataset, the Korea Water Management Information System (WAMIS) [45] was referred to, and the water level measurement station used data from upstream and downstream water level measurement stations in Yeojubo. The water level dataset and the meteorological dataset are used as input data for the water level prediction model in this paper, and data were measured at 1 h intervals from 2 October 2013 to 9 June 2022. The total dataset for upstream and downstream consists of 71,136 rows each and can be confirmed by visualizing the entire dataset upstream and downstream through Figure 4.



(a) Water level dataset of Yeojubo Upstream

(b) Water level dataset of Yeojubo Downstream

Figure 4. Water level dataset visualization.

The upstream and downstream datasets maintain an appropriate water level except in the summer when there is high rainfall; nevertheless, they have the same characteristic in that the water level increases rapidly in the summer when there is high rainfall. However, in the case of upstream of Yeojubo, it was confirmed that the amount of increase when rainfall occurred compared to the water level in the period when there was little rainfall was significant compared to the downstream recordings.

In this paper, two types of data provided by the Korea Meteorological Administration were used for meteorological training data for the model. The Korea Meteorological Administration uses a variety of meteorological equipment to observe the weather in the field, such as ground, high-rise, marine, and aviation. In this paper, data from the Automatic Weather System (AWS) and the Automated Synoptic Observing System (ASOS) were used to observe the weather on the ground. In the case of the AWS, it is a device designed for automatic observation without human observation, and it automatically handles all processes such as real-time measurement, calculation, storage, and display. The items to be observed include air pressure, temperature, humidity, wind direction, wind speed, precipitation, etc., and are observed in real time. Both ASOS and AWS datasets used temperature, humidity, and precipitation parameters as input data. Furthermore, as with the water level data set, data measurements were carried out at 1-h intervals. Figure 5 shows the quality information of AWS and ASOS data provided by the Korea Meteorological Administration [46]. Korea has a hot and humid climate in summer, which causes high rainfall and a large number of typhoons, and it can be seen that the accuracy of AWS data operated unattended during the corresponding period is lower than that of other periods or ASOS observation equipment. However, the location of the nearest AWS observatory is located about 8 km away from Yeojubo, the actual test area, and the location of the nearest ASOS observatory is about 20 km away from Yeojubo. Figure 6 shows the locations of the water level and meteorological data stations in the test area.

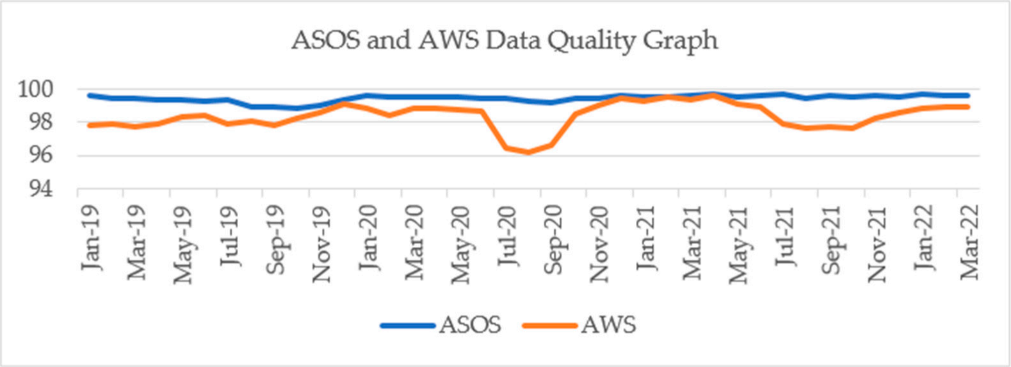


Figure 5. Automated Synoptic Observing System (ASOS) and Automatic Weather System (AWS) Data Quality Graph.

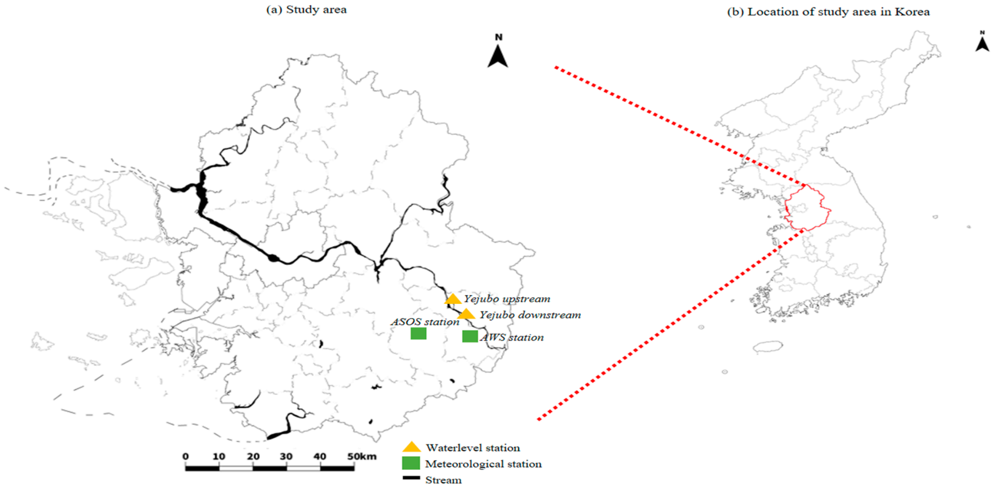


Figure 6. Location of water level and meteorological data stations in the test bed area.

In the case of torrential downpours in summer, there is a characteristic that a large amount of rainfall continues in a narrow range. Due to these characteristics, it was determined that the distance between the actual test bed and the meteorological data observation point would affect the model performance. Therefore, we intend to conduct a performance comparison experiment of the model according to the type of meteorological data used as input data. Table 1 shows the information of the hydrology and meteorology stations used and the datasets used.

Table 1. Information of meteorological station and hydrological station.

Stations	Type	Latitude	Longitude	Measurement Variable	Period (Measured Hourly)
Yeojubo upstream	Hydrology	37°19′56	127°36′34	Water level	2 October 2013~ 9 June 2022
Yeojubo downstream	Hydrology	37°19′12	127°36′43	Water level	
AWS Yeoju	Meteorology	37°16′07	127°38′22	Temperature, humidity, precipitation	
ASOS Icheon	Meteorology	37°15′50	127°29′03	Temperature, humidity, precipitation	

2.3. LSTM (Long Short-Term Memory)

LSTM is one of the recurrent neural network (RNN) architectures used in many fields for time series data processing [47]. RNN is mainly used for temporally correlated data. It considers the correlation between previous data and current data and predicts future data through past data while having a structure in which signals cycle to predict future data. However, there is a problem that the past data cannot be remembered for a long time [48,49]. LSTM is an architecture that has emerged to compensate for these problems. There are a total of six parameters, and through the structure of four gates, not only short-term memory but also long-term memory can be resolved, and the structure of the LSTM can be confirmed in Figure 7.

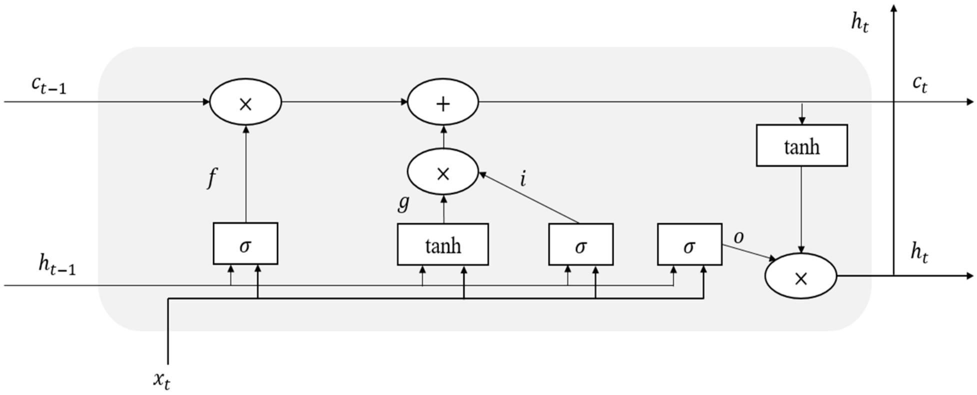


Figure 7. Structure of Long Short-Term Memory (LSTM).

The LSTM network has the same chain structure as the RNN, but the repeating modules of the RNN are structured to exchange information with each other through four layers, not just one tanh layer. The state in the LSTM cell is divided into two vectors, where h_t means a short-term state and C_t means a long-term state. Data can be added to or removed from the cell state through sigmoid gates, each gate having an individual weight, similar to a layer or series of matrix operations. It is designed to solve the problem of long-term dependencies, as gates can also retain information from long-term historical data.

The first step in an LSTM network is to identify and decide which unnecessary information to omit from the cell. The corresponding cell is called a forget gate, and the process determines the output of the last LSTM cell (h_{t-1}) at $t-1$ and the current input (x_t) at the current time t by a sigmoid function. In this instance, the value that comes out through the sigmoid function has a value between (0~1). In this case, the larger the value, the more intact is the information of the previous state that is memorized. The smaller the value, the more information about the previous state is forgotten, and the part to omit from the previous output is decided.

After going through the forget gate, it goes through the process of selecting information to store. Through the forget gate, the memory cell (c_{t-1}) of the previous time is forgotten, new information to be remembered is added, and the value of each element is determined as newly added information. In this case, appropriate choices are made rather than unconditionally accepting new information. A gate that performs a corresponding role is called an input gate. In this instance, the sigmoid function is taken through the last LSTM cell (h_{t-1}) and the current value (x_t), and the *tanh* function, which is an activation function, is added. In this instance, after going through the sigmoid layer, the value is between (0~1) and represents the degree to which new information is updated, and the value after going through the tanh function has a value between (−1~1) and has a weight representative of the importance given to it. Finally, the input gate performs the Hadamard product operation on the two values, and the corresponding new memory is added to the previous cell state (C_{t-1}) to become C_t.

After judging the value of new information through the input gate and selecting the information to remember, the next process is to select the output information. The corresponding gate is called an output gate, and a sigmoid function is obtained through the value of the current time (x_t) and the value of the last LSTM cell (h_{t-1}). The result value through the sigmoid function performs the operation of the current cell state (C_t) and Hadamard product, and the effect of filtering the value occurs and it becomes a hidden state.

$$f_t = \sigma\left(x_t W_f + h_{t-1} W_f + b_f\right), \tag{1}$$

$$g_t = tanh\left(x_t W_g + h_{t-1} W_g + b_g\right), \tag{2}$$

$$i_t = \sigma(x_t W_i + h_{t-1} W_i + b_i), \tag{3}$$

$$c_t = f_t \odot c_{t-1} + g_t \odot i_t, \tag{4}$$

$$o_t = \sigma(x_t W_o + h_{t-1} W_o + b_o), \tag{5}$$

$$h_t = o \odot tanh(c_t), \tag{6}$$

Here, σ denotes a sigmoid function, W denotes a weight matrix, and b denotes a bias. c_t denotes the cell state of the current time, c_{t-1} denotes the cell state of the previous time, and $\odot$ means Hadamard product operation. Equation (1) refers to the process of going through the forgetting gate, and in the process of going through the input data, the cell state is updated through Equations (2)–(4). Next, it goes through the output gate indicated by Equation (5), and the LSTM operates in a structure in which the final hidden layer state is updated through Equation (6).

2.4. GRU (Gated Recurrent Units)

As one of the RNN architectures, GRU is a model inspired by LSTM. It is a model that improves the problem of RNN and reduces the computation of updating the hidden state while maintaining the solution to the long-term dependency problem of LSTM. The performance is similar to that of LSTM [50]. In the case of LSTM, there is a problem that overfitting occurs when there is insufficient data, because more parameters are required compared to the existing RNN to solve the long-term dependency problem. GRU improved this shortcoming through a LSTM structure change. GRU improved this shortcoming through the structural change of LSTM. The structure of GRU is shown in Figure 8.

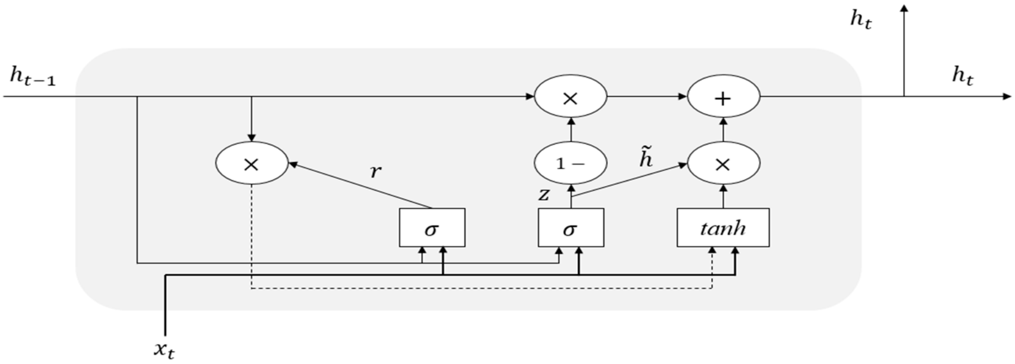


Figure 8. Structure of Gated Recurrent Units (GRU).

In Figure 8 it can see that the structure is more concise than the LSTM structure of Figure 7. The main difference from the LSTM is that GRU integrates the LSTM's forget gate and input gate and replaces it with an update gate. In addition, it has a simpler structure than the LSTM by integrating the cell state and the hidden state, and the calculation cost is low because the number of parameters is smaller than that of the LSTM.

The process corresponding to r in Figure 8 means the reset gate, and the hidden state of the network is calculated through this process. The result of going through the reset gate is calculated with the information of the past hidden layer to calculate the candidate group for the hidden state. Next, the part corresponding to z means the update gate. Through this part, the functions of forget gate and input gate of LSTM are performed, and how much current information is used is decided. The value calculated through the corresponding gate is calculated with the candidate of the previously calculated hidden state to determine the final hidden state. Equation (7) below corresponds to a reset gate, Equation (8) corresponds to an update gate, and Equations (9) and (10) are equations for determining hidden state candidates and the final hidden state.

$$r_t = \sigma(x_t W_r + h_{t-1} W_r + b_r), \tag{7}$$

$$z_t = \sigma(x_t W_z + h_{t-1} W_z + b_z), \tag{8}$$

$$\widetilde{h_t} = tanh(x_t W_x + (r \odot h_{t-1}) W_h + b), \tag{9}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widetilde{h_t}, \tag{10}$$

Since it cannot be concluded that either GRU or LSTM is better in terms of model performance, in this paper, we conduct an experiment based on LSTM and GRU, and compare their performance.

2.5. Performance Indicators

In this paper, three types of performance comparison indicators of the water level prediction model are MSE (mean squared error), NSE (Nash–Sutcliffe coefficient of efficiency), and MAE (mean absolute error). MSE is a widely used method for evaluating the performance of a regression model. It is the square of the difference between the actual observed value and the predicted value. The indicator is sensitive to outliers because the difference between the observed and predicted values is squared. In the case of the hydrologic model, the MSE index was selected because it can cause human casualties if an outlier occurs in the prediction. In the case of NSE, it is a widely used indicator to evaluate the performance of hydrological models. NSE has a value of $(-\infty$~$1)$ as an index frequently used to evaluate the performance of hydrological models. A value closer to 1 means better model performance. In the case of MAE, it means the average of all absolute errors of the observed and predicted values and has the advantage of intuitively checking the performance of the model. The equations for the performance comparison metrics are given as follows:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (OBS_i - SIM_i)^2, \ \ MSE > 0. \tag{11}$$

$$NSE = 1 - \left(\frac{\sum_{i=1}^{N} OBS_i - SIM_i)^2}{\sum_{i=1}^{N} OBS_i - \overline{OBS})^2} \right), -\infty < NSE \leq 1, \tag{12}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |OBS_i - SIM_i|, \ \ MAE > 0, \tag{13}$$

In this paper, performance evaluation is performed according to LSTM, GRU model, and meteorological data through three indicators. Information on each piece of experiment is described in Section 3.

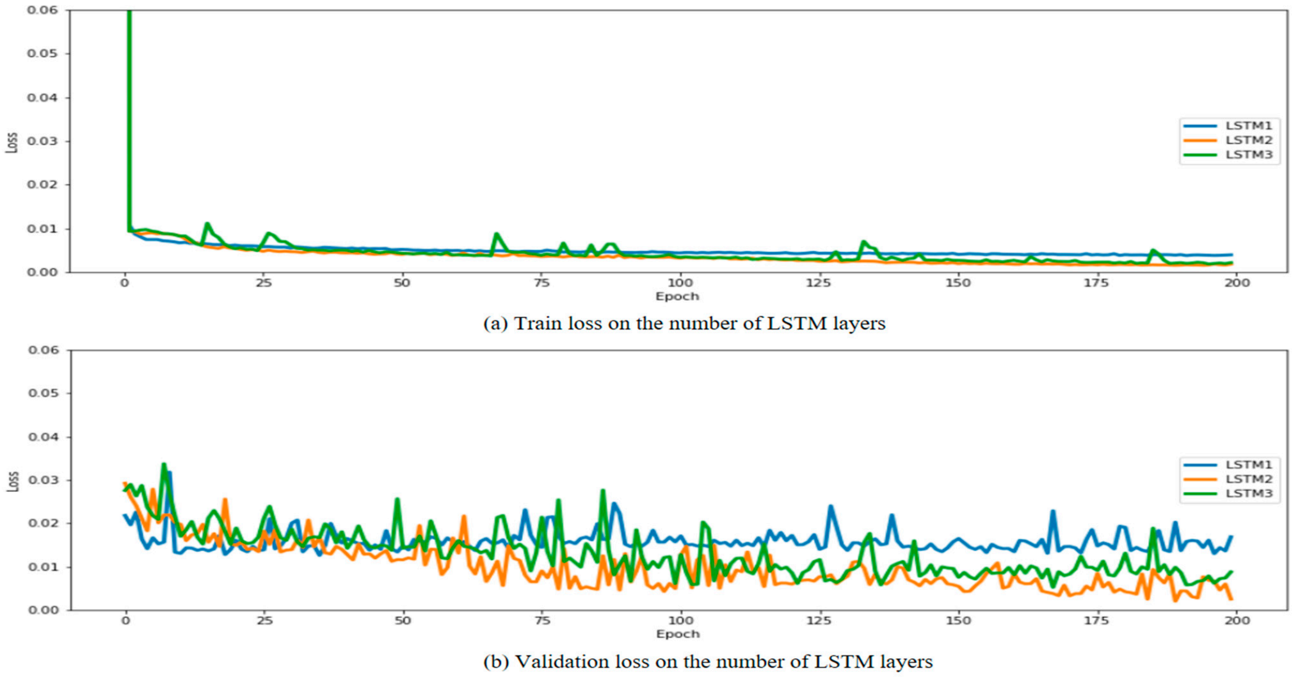
3. Experiment and Model Design

3.1. Preliminary Experiment

A preliminary study was conducted on the LSTM–GRU-based water level prediction model proposed in this paper. A preliminary experiment was conducted in three main approaches. First, for comparison with the LSTM–GRU model, an experiment was conducted to find the optimal configuration of the model using only the LSTM and the model using only the GRU. The second experiment was conducted to find the optimal number of units for each LSTM layer and GRU layer of LSTM–GRU, and finally, the optimal input data size was determined by comparing the performance according to the size of the input data. All factors have a great influence on the performance of the model, and the value that gives the optimal result is used as the model parameter for this experiment through the preliminary experiments. In this case, the water level prediction model is a regression model and performance comparison is performed through the MSE value, which is a loss function.

3.1.1. Model Composition Experiment for Selecting a Performance Comparison Model

A comparison experiment was conducted according to the number of layers to find the most suitable layer configuration of the water level prediction model composed of only LSTM and the water level prediction model composed only of GRUs. This approach was used because the model composed of LSTM and GRU is already showing good results in the field of time series data processing. A large number of hidden layers in the model does not necessarily show good performance, and if it is configured with excessive hidden layers there may be side effects such as overfitting. Therefore, to compare with the LSTM-GRU model proposed in this paper, a performance comparison experiment was performed on a model composed of 1 to 3 hidden layers to determine the appropriate number of layers for the LSTM and GRU models. Figures 9 and 10 show the experimental results according to the number of hidden layers of LSTM and GRU.



(a) Train loss on the number of LSTM layers

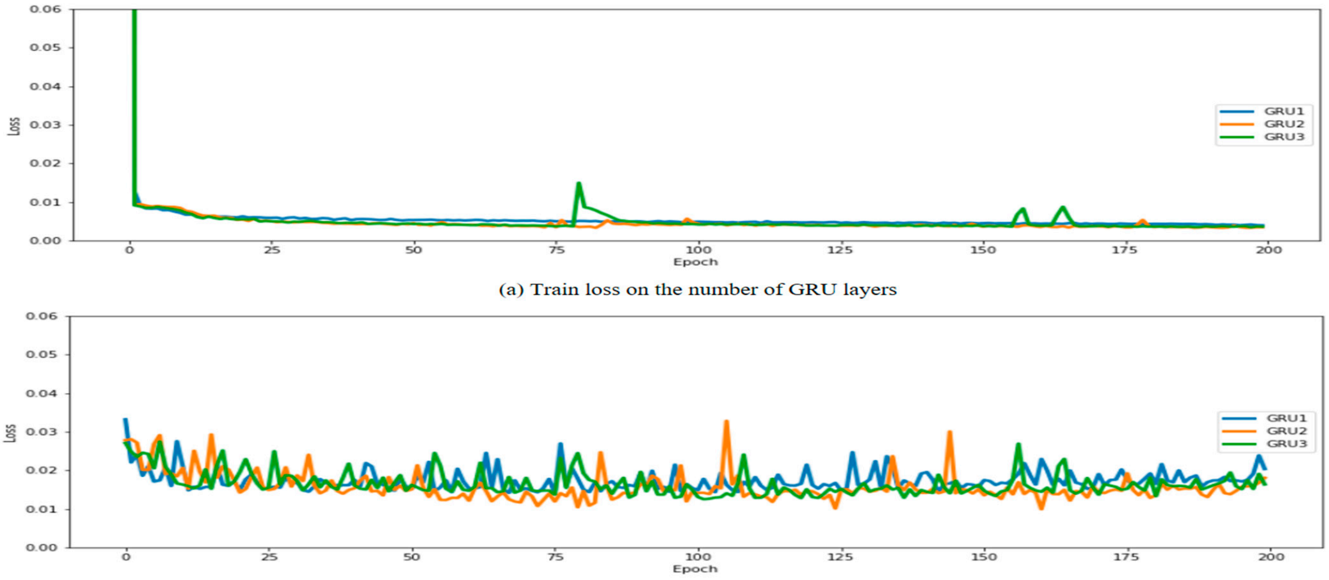
(b) Validation loss on the number of LSTM layers

Figure 9. Train and validation loss values according to the number of LSTM layers.

In the case of the LSTM 1-layer model, the number of hyperparameters was 268,545 and the training time was 1578.38 s. The final training loss value was 0.39 and the validation loss value was 1.68. In the case of the LSTM 2-layer model, the number of hyperparameters was 793,857 and the training time was 2315.98 s. The final training loss value was 0.18 and the validation loss value was 0.59. In the case of the LSTM 3-layer model, the number of hyperparameters was 1,319,169 and the training time was 3401.99 s. The final training loss value was 0.21 and the validation loss value was 0.88. The LSTM 3-layer model had a slower learning rate and more hyperparameters than the LSTM 2-layer model but showed lower performance than the LSTM 2-layer model due to overfitting. Therefore, the model composition that showed the best performance was the LSTM 2-layer model and it is used as a comparative model in the experiments in Section 4.

In the case of a model with GRU 1-layer, the number of hyperparameters was 202,241 and the training time was 1372.81 s. The final learning loss value was 0.39 and the validation loss value was 2.03. In the case of the model with GRU 2-layers, the number of hyperparameters was 596,993 and the training time was 211,319 s. The final learning loss value was 0.34 and the validation loss value was 0.18. In the case of the model consisting of GRU 3-layers, the number of hyperparameters was 991,745 and the training time was 3009.27 s. The final learning loss value was 0.36 and the validation loss value was 1.91. The model composed of GRU 2-layers showed the best performance among the GRU models.

Although it was confirmed that the GRU 1-layer showed the lowest performance, it was confirmed that the difference according to the number of layers was smaller than that of the LSTM. GRU 2-layer is also used as a comparative model in the experiments in Section 4.



(a) Train loss on the number of GRU layers

(b) Validation loss on the number of GRU layers

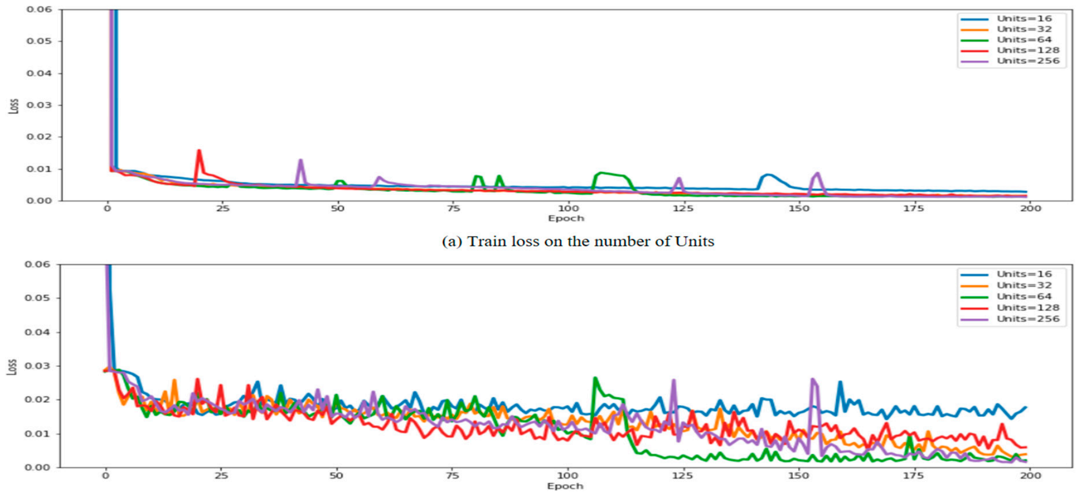
Figure 10. Train and validation loss values according to the number of GRU layers.

3.1.2. Comparative Experiment according to the Size of the Number of Units in the LSTM-GRU Model

The number of units in the LSTM and GRU is also an important factor in learning the model. Experiments were conducted to find the optimal number of units in the LSTM–GRU. In order to find the optimal number of units, a result comparison experiment was conducted when the number of units was set to 16, 32, 64, 128, and 256. Figure 11 shows the experimental results according to the number of units of LSTM and GRU.

When comparing the results according to the number of units of five types, the best results were obtained when the number of units was 256. It was confirmed that the learning time was somewhat proportional to the number of units. As a result, there was no significant difference in training loss values for the other four types except when the number of units was 16. With a slight difference, it was confirmed that the best performance was shown when the number of units was 256 and the second-best performance was confirmed when the number of units was 32. When the number of units were 128 and 64, the training loss value was the same.

The validation loss value also obtained the best performance when the number of units was 256, performance then followed the order of 64, 32, 128, and 16 units. When the number of units was 256, it was confirmed that both learning and verification showed the best performance, and when the number of units was 16, it was confirmed that the worst performance was shown. According to the experimental results, the number of units was finally determined to be 256.



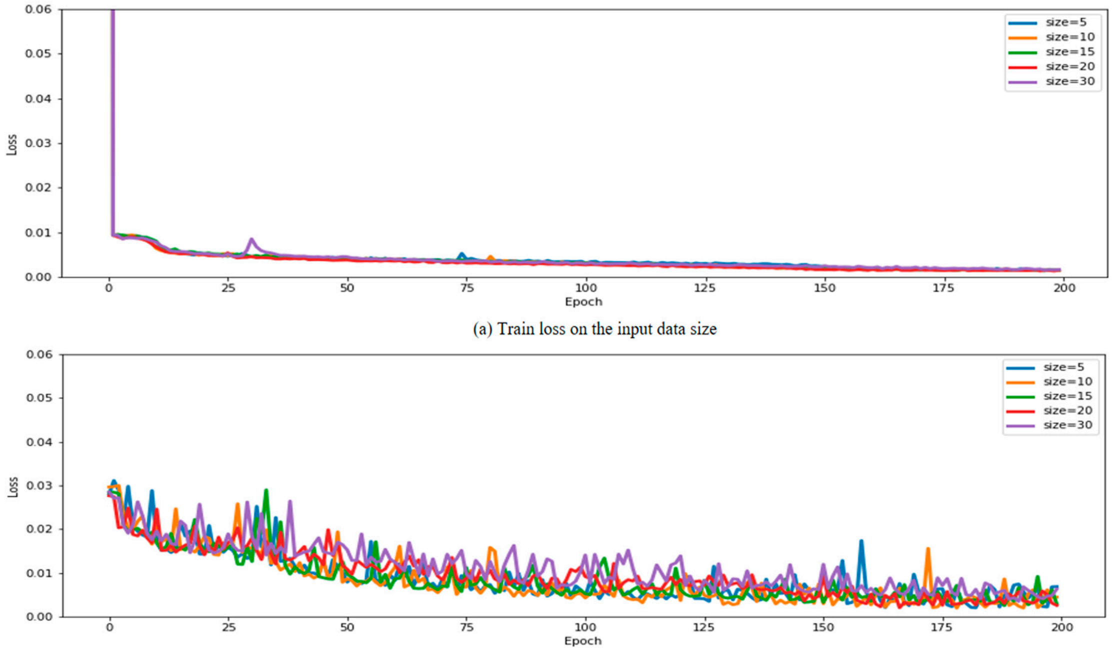
(a) Train loss on the number of Units

(b) Validation loss on the number of Units

Figure 11. Train and validation loss values according to the number of units.

3.1.3. Comparative Experiment according to Input Data Size

The size of the input data used for water level prediction is also one of the factors that greatly affects the performance of the model. In order to find the appropriate size of input data, we conducted a performance comparison experiment according to the size of various input data. The size of the input data was composed of the past 5 h, 10 h, 15 h, 20 h, and 30 h, and the performance was compared according to each input data. The experimental results can be seen in Figure 12.



(a) Train loss on the input data size

(b) Validation loss on the input data size

Figure 12. Train and validation loss values according to input data size.

Finally, the case with the best performance used the past 20 h of data as input data. The learning time increased as the size of the input data increased. There was no significant difference in training loss values in all cases. However, when the size of the input data was 5 and 10, the loss value increased due to overfitting in the model. The learning loss value had a slight difference, and the best performance was obtained when using the past 20 h of data.

Furthermore, in the case of the verification loss value, it was possible to check the performance difference according to the input data. Similarly, when the input data size was 5 and 10, a phenomenon in which the loss value bounced due to overfitting was found. Even when the size of the input data was 30, the final convergence loss value was high. Among the five types, it was confirmed that excellent performance was shown when the input data size was 20 and 15, with the input data size of 20 showing slightly better performance. Therefore, the size of the final input data was selected as 20.

3.2. Experimental Design

In this paper, an experiment was conducted according to the input data using the LSTM model and the GRU model. There were three model configurations to be used in the experiment as follows: 1. the multi-LSTM model with 2 LSTM layers; 2. the multi-GRU model with 2 GRU layers; 3. the LSTM–GRU model with LSTM and GRU layers. As for the dataset to be used for model training, a total of three datasets were used: a dataset consisting of water level data only, a dataset consisting of water level data and AWS meteorological data, and a dataset consisting of water level data and ASOS meteorological data. We compared the performance of a total of nine models for three models and three datasets. Table 2 shows the details of the experiment.

Table 2. Experiment Details.

Model	Case	Details of Input Variable	Meteorology Data Type
Multi LSTM	S1_LSTM	Upstream water level, downstream water level	—
	S2_LSTM	Upstream water level, downstream water level, temperature, humidity, rainfall	AWS
	S3_LSTM	Upstream water level, downstream water level, temperature, humidity, rainfall	ASOS
Multi GRU	S1_GRU	Upstream water level, downstream water level	—
	S2_GRU	Upstream water level, downstream water level, temperature, humidity, rainfall	AWS
	S3_GRU	Upstream water level, downstream water level, temperature, humidity, rainfall	ASOS
LSTM-GRU	S1_LSTM_GRU	Upstream water level, downstream water level	—
	S2_LSTM_GRU	Upstream water level, downstream water level, temperature, humidity, rainfall	AWS
	S3_LSTM_GRU	Upstream water level, downstream water level, temperature, humidity, rainfall	ASOS

S1 means a scenario using the water level dataset as training data, S2 means a scenario using the water level dataset and AWS weather dataset as training data, and S3 uses the water level dataset and the ASOS dataset as training data. Therefore, we conducted a performance comparison experiment according to nine scenarios.

3.3. Model Structures

There are three structures of the water level prediction model used in this paper, and the overall model configuration can be confirmed through Figure 13.

All three models commonly use data from the past 20 h as input data. In this instance, when the learning data corresponding to scenario S1 is used, the form of the input data consists of [None,2], and in the case of S2 and S3, it consists of [None, 5]. Finally, it goes through the dense layer and eventually predicts the data of the 21st hour through the data of the past 20 h. In the case of the multi-LSTM, all hidden layers are composed of LSTM layers, and in the multi-GRU, both layers are composed of GRUs. In the LSTM–GRU model, the first hidden layer is composed of an LSTM layer and the second hidden layer is composed of a GRU layer. The performance comparison results according to the input data through the three models are described in Section 4.

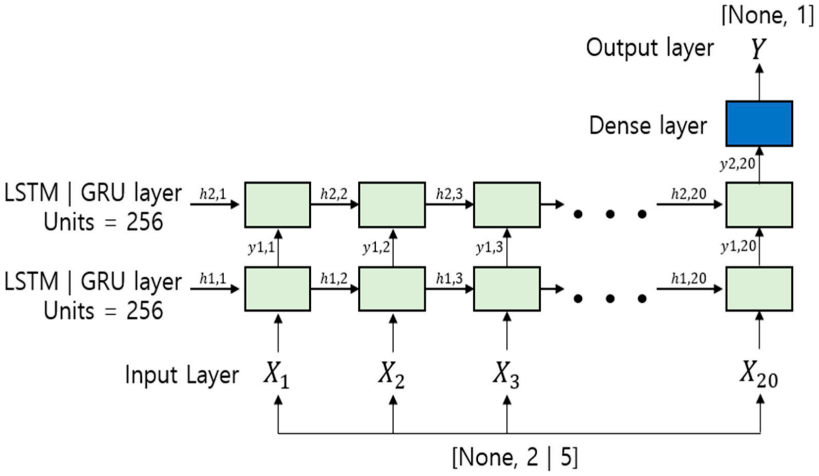


Figure 13. Construction of the proposed model.

4. Results

The water level and meteorological data sets described in Section 2 were used for the data for the experiment, and the experiment was conducted using the input data and model consisting of nine types of input data using the three models described in Section 3.

The total collection period of all training and test data was measured at 1-h intervals from 2 October 2013 to 9 June 2022. Of the data consisting of a total of 76,152 rows, 57,114 rows (75%) were used as training data and validation data, and the remaining 19,018 rows (25%) were used as test data. In the case of a flood prediction model, it is important to predict a high-water level due to a rapid increase. In this instance, the maximum value among all the measured water levels exists in the test data, so that the validity of the model can be verified.

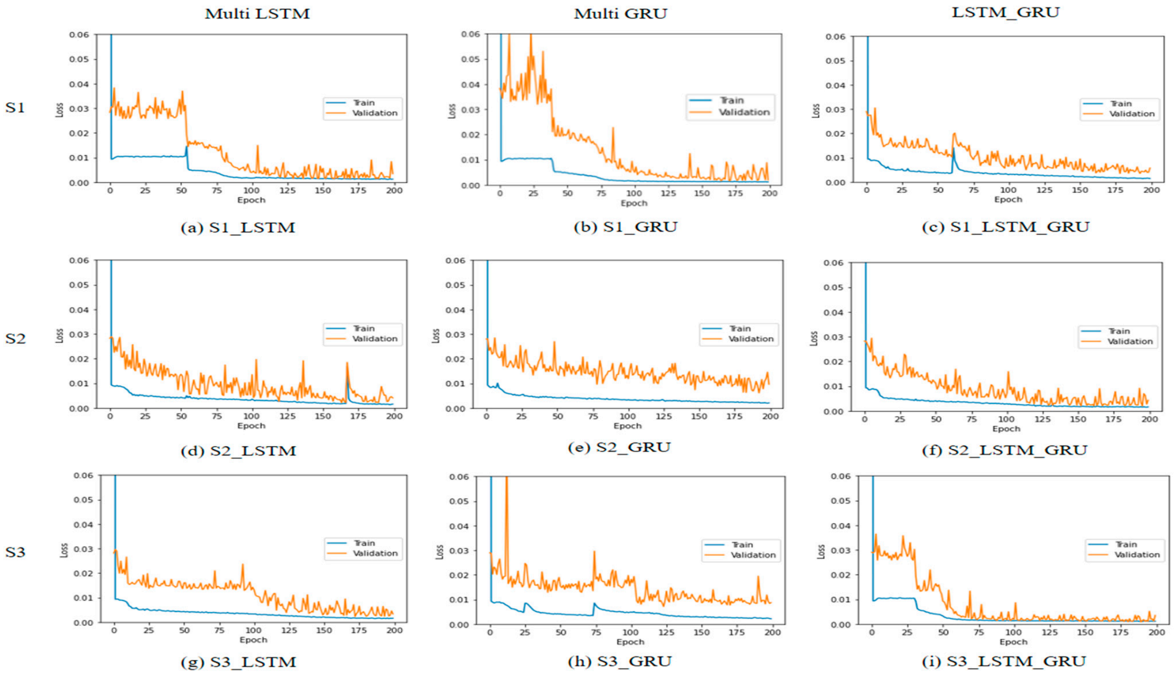
4.1. Training and Validation Result

MSE was used as the basic loss function for model training, and NSE and MAE indicators were used as auxiliary indicators for actual test data to compare observed values and predicted values. In addition, Adam was used as the optimization function, the number of units of the LSTM and GRU models was 256, and the number of training iterations (epoch) was 200, which was performed under the same conditions. The reason for selecting the parameters of the model was to derive the optimal values through preliminary experiments. In the case of the learning model, a total of nine cases were carried out as in the case summarized in Table 2, information on each model training and validation can be checked in Table 3, and the change in loss value according to epoch during model training is shown in Figure 10. In the case of the learning model, a total of nine cases were performed as in the case summarized in Tables 2 and 3 and shows information about each model learning and validation. Figure 10 shows the change in loss value according to epoch during model training.

According to the characteristics of the LSTM model, the multi-LSTM model had the largest number of hyperparameters, and the multi-GRU model had the smallest number of hyperparameters. In the case of training time, the GRU model took the least amount of time, but it was confirmed that the number of hyperparameters and the training time were not proportional. In addition, looking at Figure 14, the loss value converges to close to 0 according to the epoch in all cases, and it was determined that the learning was well accomplished. It was considered that overfitting occurred, but for equality with other models, the epochs were maintained at 200 and the experiment was conducted. In the case of the MSE value used as the loss function, differences were found according to the input data characteristics for each model. In the case of the multi-LSTM model, there was no significant difference depending on the input data; however, in the case of multi-GRU and LSTM–GRU, it was confirmed that the performance difference occurred depending on the input data. In the case of the multi-GRU model, it was confirmed that the performance of the S2 and S3 cases, including the weather data in the training data, was lower than that of the S1 case, and it was confirmed that the performance was poor among the nine cases. When checking the training and validation results, it was confirmed that GRU showed good performance for S1 data with a small number of dimensions, but poor performance for S2 and S3 with a large number of dimensions. However, in the case of the LSTM–GRU model, it was confirmed that the verification MSE of the S1 model with a small dimension of training data showed the lowest performance. Finally, the worst performing case was S3_GRU which showed the worst results with 0.39 for training and 1.80 for validation, and the case with the best performance showed the best results with S3_LSTM_GRU, 0.15 for training and 0.19 for validation.

Table 3. Training and validation results.

Case	Hyperparameter	Training Time	Train MSE (cm)	Validation MSE (cm)
S1_LSTM	790,785	2510.91 s	0.22	0.31
S2_LSTM	793,857	2492.76 s	0.19	0.38
S3_LSTM	793,857	2606.77 s	0.18	0.34
S1_GRU	594,689	2072.76 s	0.22	0.49
S2_GRU	596,993	2141.85 s	0.48	1.75
S3_GRU	596,993	2305.55 s	0.39	1.80
S1_LSTM_GRU	660,225	2331.33 s	0.60	1.74
S2_LSTM_GRU	663,297	2343.19 s	0.17	0.28
S3_LSTM_GRU	663,297	2493.50 s	0.15	0.19



(a) S1_LSTM (b) S1_GRU (c) S1_LSTM_GRU (d) S2_LSTM (e) S2_GRU (f) S2_LSTM_GRU (g) S3_LSTM (h) S3_GRU (i) S3_LSTM_GRU

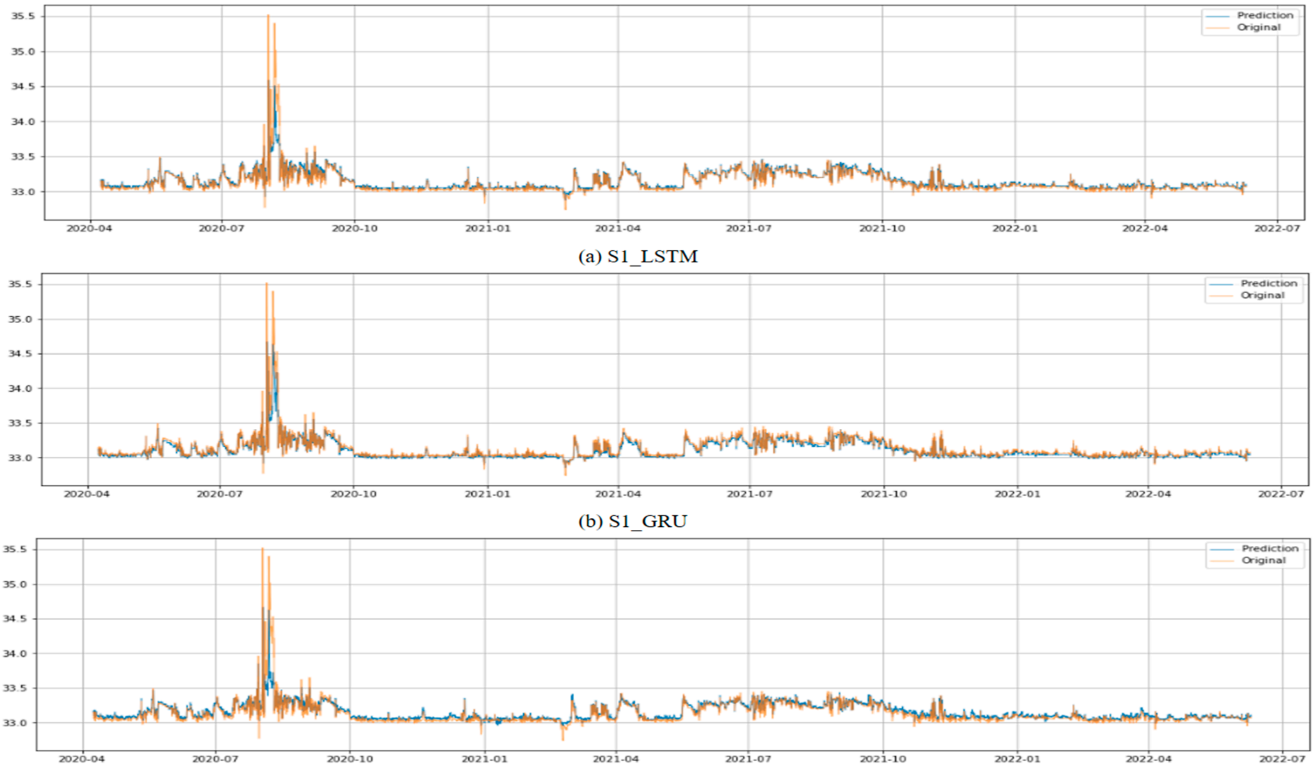
Figure 14. Loss value graph according to epoch during model training by scenario.

4.2. Test Result

The data set for the test has the highest water level data during the entire data collection period, and the performance of the model is judged by comparing the MSE, NSE, and MAE indicators with the highest water level prediction error. In this instance, the highest water level in the observed data was 3552.38 cm. The test results are shown in Table 4, and the difference between the observed value and the predicted value according to the training data for each model are shown in Figures 15–17.

Table 4. Test results.

Case	Test MSE (cm)	Test NSE	Test MAE (cm)	Maximum Prediction(cm)	Highest Water Level Prediction Error (cm)
S1_LSTM	5.02	0.917	2.74	3461.11	91.27
S2_LSTM	4.73	0.834	3.32	3434.51	117.87
S3_LSTM	4.88	0.905	2.35	3458.91	93.47
S1_GRU	7.73	0.835	3.93	3467.36	85.02
S2_GRU	5.65	0.430	3.36	3457.38	95
S3_GRU	5.89	0.315	6.22	3340.61	211.77
S1_LSTM_GRU	3.95	0.864	2.84	3466.67	85.71
S2_LSTM_GRU	5.02	0.920	2.88	3466.74	79.64
S3_LSTM_GRU	3.92	0.942	2.22	3490.36	55.49



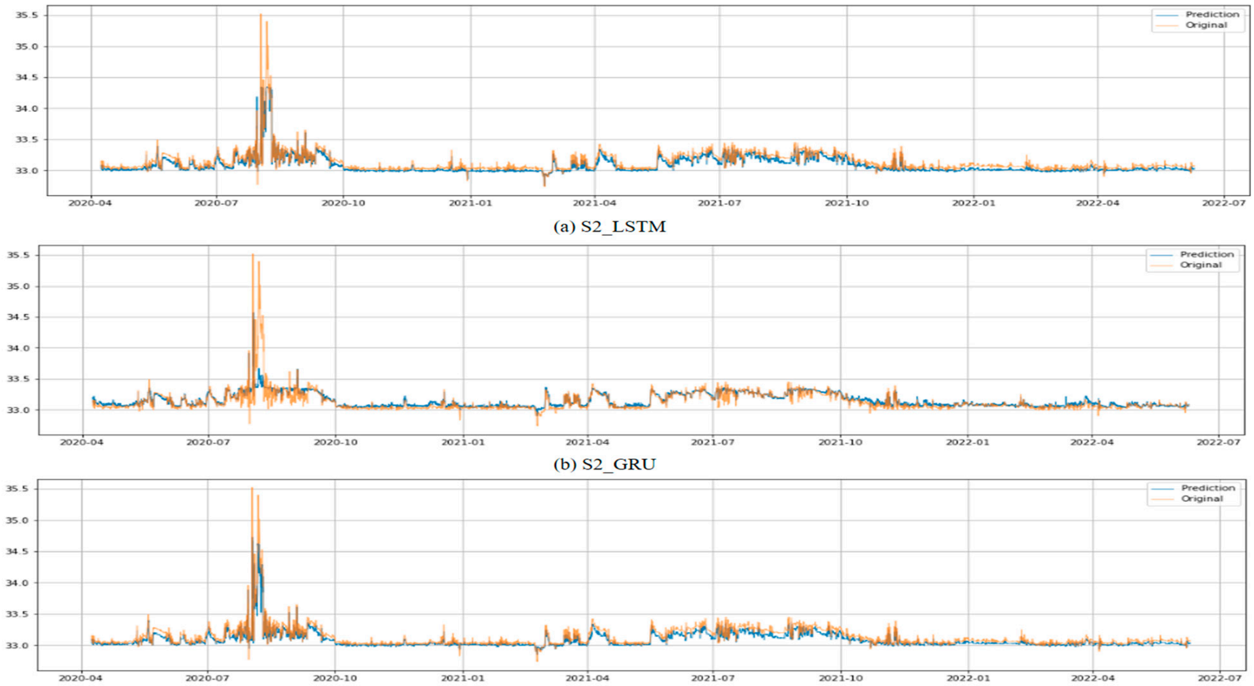
(a) S1_LSTM

(b) S1_GRU

(c) S1_LSTM_GRU

Figure 15. Comparison graph of observed and predicted values for each model in S1 case.

The test results are similar to the model training and validation results. In the multi-LSTM model, no significant difference was found depending on the input data, but the NSE value obtained the best result when the data corresponding to the S1 case were used as the input data. In addition, it was confirmed that the difference in the predicted maximum water level included in the test data was 91.27 cm in the S1 case, which obtained the best result. Among the LSTM models, when the S2 case data were used, the NSE was 0.834, which was slightly lower than those of the S1 and S3 cases, and the maximum water level prediction error was also the lowest at 117.87 cm.

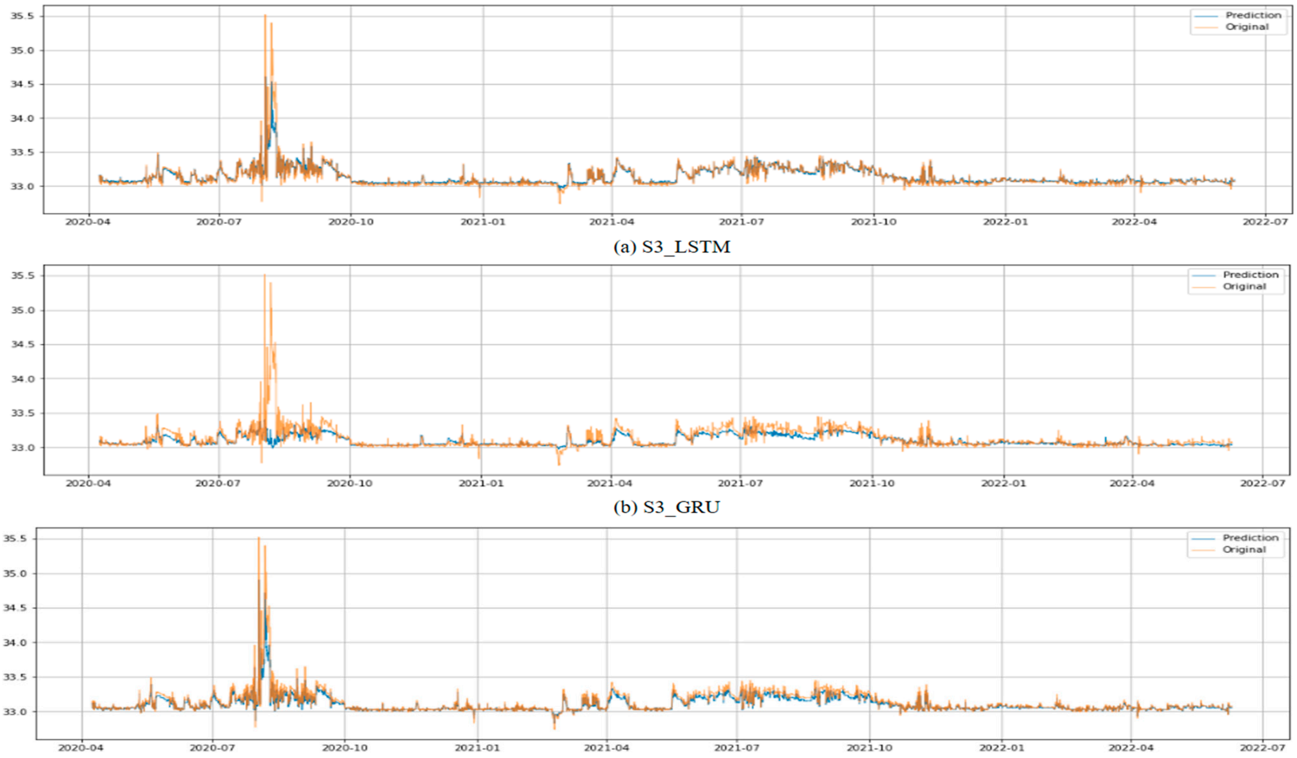


(a) S2_LSTM

(b) S2_GRU

(c) S2_LSTM_GRU

Figure 16. Comparison graph of observed and predicted values for each model in S2 case.



(a) S3_LSTM

(b) S3_GRU

(c) S3_LSTM_GRU

Figure 17. Comparison graph of observed and predicted values for each model in S3 case.

In the case of the multi-GRU model, the lowest performance was confirmed in all cases compared to other models. Similarly, in the results of verification it was confirmed that the GRU-based water level prediction model effectively learned S1 with a small number of dimensions from the experimental results but showed poor performance when the input data had a large number of dimensions. When the data of S1 case were used, the best results were obtained in the GRU model, but it was confirmed that the results were inferior compared to other models. For S2 and S3, the NSE values were 0.2 and 0.3, respectively, which were the lowest among the nine cases. Furthermore, in the case of S3, it was confirmed that the maximum water level prediction error was 211.77, being the lowest result among all scenarios.

The LSTM–GRU model showed the best performance on average among the three models, and in the case of S1, it was confirmed that the NSE was slightly lower than that of the LSTM model. However, in the case of S2 and S3, the NSE values were 0.920 and 0.942, respectively, and the maximum water level error was 79.64 and 55.49, which showed much better performance than other models.

Of these, when the S3 case was used, the best results were obtained among the nine models in all evaluation indicators, including MSE, NSE, MAE, and maximum water level prediction error. Therefore, it is determined that the most suitable weather data for the test bed, Yeojubo, is the ASOS data corresponding to the S3 case in view of the model's performance.

5. Discussion

Due to global warming and abnormal climate changes, the frequency of and damage from floods are steadily increasing. Therefore, many studies are being conducted for flood management to minimize flood damage. Hydrological models and data-based intelligent models are mainly used to predict the water level, the most important parameter of flooding. In this paper, we propose the LSTM–GRU water level prediction model included in the data-based intelligent model. In this study, Yeojubo in Gyeonggi-do, Korea was selected as the test area, and the water level datasets upstream and downstream of Yeojubo were used. In addition, using meteorological data, comparison experiments between ASOS data and AWS data were conducted.

First, we conducted two experiments to find the most suitable parameters for the LSTM–GRU model. The two experiments were as follows: 1. experiment according to the number of units of the LSTM and GRU layers; and 2. performance comparison experiment according to input data size. As a result of the experiment, the best results were obtained when the number of units was 256, and the best performance was obtained when the size of the input data was used for the past 20 h. Next, to select a comparative model, a model composed of LSTM and GRU, which shows good performance in time-series data processing, was used. In this case, a performance comparison experiment was conducted according to the number of hidden layers to find the most suitable model configuration. When both LSTM and GRU were composed of two hidden layers, the best performance was shown, and multi-LSTM and multi-GRU models were selected as comparative models.

Next, a comparison experiment was conducted according to the three input data sets, and the three data sets consist of the following: 1. water level dataset; 2. water level dataset + AWS dataset; and 3. water level dataset + ASOS dataset. Experiments were conducted on nine scenarios using three datasets for a total of three models, and the model performance was evaluated using MSE, NSE, MAE, and maximum water level error. In the case of MSE, as an evaluation index sensitive to outliers, it was selected because a fatal error can occur if an outlier occurs in response to a rapidly rising water level. In the case of NSE, it is a model widely used for performance evaluation of hydrological models, and MAE has the strength of intuitively checking errors. Lastly, the reason for using the maximum water level error is that the historical highest water level of Yeojubo in the test data set occurred, so it was used as an evaluation index to check to what extent the water level can be predicted. In this instance, the LSTM–GRU model proposed in this study obtained the best results among nine scenarios when the water level dataset and the ASOS dataset were combined as input data. If the LSTM–GRU model proposed in this paper is used, it is considered that it can be used as a preliminary study of a flood management system that can predict the water level in advance, and evacuate through risk assessment.

6. Conclusions

In this paper, we propose an approach for data-based water level prediction using actual observation data. In order to predict the data of the 21st hour using the data of the past 20 h through the data of the upstream and downstream of Yeojubo in Gyeonggi-do, Republic of Korea, a model was constructed using the LSTM and GRU models, and a performance comparison experiment was conducted according to the input data. For comparative experiments, a model composed of LSTM and GRU was used. In this instance, the optimal model configuration was selected through experiments, and LSTM composed of two layers and GRU composed of two layers were selected as final comparative models. The existing hydrological method used for flood prediction has a problem in that it is difficult to build a model because various factors, including soil and watershed characteristics, must be considered, and there is a disadvantage that requires high hydrological competence of the researcher. The water level prediction model proposed in this paper has the advantage of being able to predict the water level more easily by using only the meteorological data set and the water level data set. Multi-LSTM, multi-GRU, and LSTM–GRU models were used to conduct a performance comparison experiment according to a total of nine cases, and it was confirmed that there was a significant performance difference when the input data was configured differently according to each model. The three types of input data are: S1 case using water level data only, S2 case using water level data and AWS data provided by the Korea Meteorological Administration, and ASOS data provided by the Korea Meteorological Administration. When looking at the data quality information provided by the Korea Meteorological Administration, the quality of the ASOS data was higher, but due to the concentrated heavy rain that fell in a narrow range in a short period of time in summer due to the seasonal nature of Korea, the AWS data were used in the comparative experiment. The LSTM and LSTM–GRU-based models confirmed good results for all three data types, but in the case of the GRU model, it was confirmed that the results were slightly inferior when the data of S2 and S3 cases with large input data dimensions were used. It was determined that the GRU, which reduced the number of parameters by simplifying the LSTM, did not effectively learn the characteristics of the training data used in this paper. In this instance, the case that obtained the best results used ASOS data and water level data, which are S3 cases, as training data, and the water level prediction model learned through the LSTM–GRU model obtained the best results. The test MSE of this model was 3.92, NSE was 0.942, and MAE was 2.22, showing the best performance in all evaluation indicators, and the maximum water level prediction error was 55.49 cm, which was the best result. In order to predict floods, the ability to predict a sudden rise in water level due to heavy rainfall, etc. is a very important factor, so the model is judged to be the most suitable among the nine cases. The NSE of the LSTM–GRU model proposed in this paper confirmed a high score of 0.942, and it is considered that it can be used for flood management.

Author Contributions: Conceptualization, M.C.; methodology, M.C. and C.K.; formal analysis, M.C. and K.J.; writing—original draft, M.C.; supervision, H.J. All authors have read and agreed to the published version of the manuscript.

Funding: This study was carried out with the support of R&D Program for Forest Science Technology (Project No. 2021340A00-2123-CD01) provided by Korea Forest Service (Korea Forestry Promotion Institute).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors also greatly appreciate the anonymous reviewers and academic editor for their careful comments and valuable suggestions to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Intergovernmental Panel on Climate Change (IPCC). *Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*; IPCC: Geneva, Switzerland, 2012.
2. IPCC. *Special Report on the Ocean and Cryosphere in a Changing Climate*; IPCC: Geneva, Switzerland, 2019.
3. Centre for Research on the Epidemiology of Disasters(CRED); United Nations Office for Disaster Risk Reduction (UNDRR). *The Human Cost of Disasters: An Overview of the Last 20 Years (2000–2019)*; CRED: Brussels, Belgium; UNDRR: Geneva, Switzerland, 2020.
4. Van Oldenborgh, G.J.; Van Der Wiel, K.; Sebastian, A.; Singh, R.; Arrighi, J.; Otto, F.; Haustein, K.; Li, S.; Vecchi, G.; Cullen, H. Attribution of extreme rainfall from Hurricane Harvey, August 2017. *Environ. Res. Lett.* **2017**, *12*, 124009. [CrossRef]
5. Lee, J.; Eo, G.; Choi, C.; Jung, J.; Kim, H.S. Development of Rainfall-Flood Damage Estimation Function using Nonlinear Regression Equation. *J. Soc. Disaster Inf.* **2016**, *12*, 74–88.
6. Flack, D.L.; Skinner, C.J.; Hawkness-Smith, L.; O'Donnell, G.; Thompson, R.J.; Waller, J.A.; Chen, A.S.; Moloney, J.; Largeron, C.; Xia, X.; et al. Recommendations for improving integration in national end-to-end flood forecasting systems: An overview of the FFIR (Flooding From Intense Rainfall) programme. *Water* **2019**, *11*, 725. [CrossRef]
7. Nam, D.H.; Mai, D.T.; Udo, K.; Mano, A. Short-term flood inundation prediction using hydrologic-hydraulic models forced with downscaled rainfall from global NWP. *Hydrol. Processes* **2014**, *28*, 5844–5859. [CrossRef]
8. Zhao, R.-J. The Xinanjiang model applied in China. *J. Hydrol.* **1992**, *135*, 371–381. [CrossRef]
9. Pierini, N.A.; Vivoni, E.R.; Robles-Morua, A.; Scott, R.L.; Nearing, M.A. Using observations and a distributed hydrologic model to explore runoff thresholds linked with mesquite encroachment in the Sonoran Desert. *Water Resour. Res.* **2014**, *50*, 8191–8215. [CrossRef]
10. Kabir, M.A.; Dutta, D.; Hironaka, S. Estimating sediment budget at a river basin scale using a process-based distributed modelling approach. *Water Resour. Manag.* **2014**, *28*, 4143–4160. [CrossRef]
11. Rogger, M.; Viglione, A.; Derx, J.; Blöschl, G. Quantifying effects of catchments storage thresholds on step changes in the flood frequency curve. *Water Resour. Res.* **2013**, *49*, 6946–6958. [CrossRef]
12. Kaya, C.M.; Tayfur, G.; Gungor, O. Predicting flood plain inundation for natural channels having no upstream gauged stations. *J. Water Clim. Chang.* **2019**, *10*, 360–372. [CrossRef]
13. Ding, Y.; Zhu, Y.; Feng, J.; Zhang, P.; Cheng, Z. Interpretable spatio-temporal attention LSTM model for flood forecasting. *Neurocomputing* **2020**, *403*, 348–359. [CrossRef]
14. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
15. Shen, D.; Wu, G.; Suk, H.I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248. [CrossRef] [PubMed]
16. Zhao, Y.; Li, J.; Yu, L. A deep learning ensemble approach for crude oil price forecasting. *Energy Econ.* **2017**, *66*, 9–16. [CrossRef]
17. Nelson, D.M.Q.; Pereira, A.C.M.; de Oliveira, R.A. Stock market's price movement prediction with LSTM neural networks. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 1419–1426.
18. Qiao, W.; Tian, W.; Tian, Y.; Yang, Q.; Wang, Y.; Zhang, J. The forecasting of PM2.5 using a hybrid model based on wavelet transform and an improved deep learning algorithm. *IEEE Access* **2019**, *7*, 142814–142825. [CrossRef]
19. Li, Y.; Cao, H. Prediction for tourism flow based on LSTM neural network. *Procedia Comput. Sci.* **2018**, *129*, 277–283. [CrossRef]
20. Yanjie, D.; Yisheng, L.; Fei-Yue, W. Travel time prediction with LSTM neural network. In Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, 1–4 November 2016; pp. 1053–1058.
21. Panigrahi, B.K.; Das, S.; Nath, T.K.; Senapati, M.R. An application of data mining techniques for flood forecasting: Application in rivers Daya and Bhargavi, India. *J. Inst. Eng. (India) Ser. B* **2018**, *99*, 331–342. [CrossRef]
22. Geem, Z.W.; Kim, J.H.; Loganathan, G.V. A new heuristic optimization algorithm: Harmony search. *Simulation* **2001**, *76*, 60–68. [CrossRef]
23. Das, S.; Suganthan, P.N. Differential evolution: A survey of the state-of-the-art. *IEEE Trans. Evol. Comput.* **2010**, *15*, 4–31. [CrossRef]
24. Park, J.; Sandberg, I.W. Universal approximation using radial-basis-function networks. *Neural Comput.* **1991**, *3*, 246–257. [CrossRef]
25. Gardner, M.W.; Dorling, S.R. Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ.* **1998**, *32*, 2627–2636. [CrossRef]
26. Bande, S.; Shete, V.V. Smart flood disaster prediction system using IoT & neural networks. In Proceedings of the International Conference on Smart Technologies for Smart Nation (SmartTechCon), Bengaluru, India, 17–19 August 2017; pp. 189–194.
27. Sahagun, M.A.M.; Cruz, J.C.D.; Garcia, R.G. Wireless sensor nodes for flood forecasting using artificial neural network. In Proceedings of the International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, Philippines, 1–3 December 2017; pp. 1–6.
28. Khatibi, R.; Ghorbani, M.A.; Pourhosseini, F.A. Stream flow predictions using nature-inspired Firefly Algorithm Algorithms and Algorithm a Multiple Model strategy–Directions of innovation towards next generation practices. *Adv. Eng. Inform.* **2017**, *34*, 80–89. [CrossRef]
29. Keong, K.C.; Mustafa, M.; Mohammad, A.J.; Sulaiman, M.H.; Abdullah, N.R.H. Artificial neural network flood prediction for sungai isap residence. In Proceedings of the International Conference on Automatic Control and Intelligent Systems (I2CACIS), Shah Alam, Malaysia, 22 October 2016; pp. 236–241.

30. Abu Bakar, M.A.; Abdul Aziz, F.A.; Mohd Hussein, S.F.; Abdullah, S.S.; Ahmad, F. Flood water level modeling and prediction using radial basis function neural network: Case study Kedah. In Proceedings of the Asian Simulation Conference, Singapore, 26 August 2017; pp. 225–234.
31. Maspo, N.A.; Harun, A.N.B.; Goto, M.; Cheros, F.; Haron, N.A.; Nawi, M.N.M. Evaluation of Machine Learning approach in flood prediction scenarios and its input parameters: A systematic review. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *479*, 012038. [CrossRef]
32. Liu, C.; Guo, L.; Ye, L.; Zhang, S.; Zhao, Y.; Song, T. A review of advances in China's flash flood early-warning system. *Nat. Hazards* **2018**, *92*, 619–634. [CrossRef]
33. Govindaraju, R.S. Artificial neural networks in hydrology. I: Preliminary concepts. *J. Hydrol. Eng.* **2000**, *5*, 115–123.
34. Mijwel, M.M. Artificial Neural Networks Advantages and Disadvantages. Available online: https://www.linkedin.com/pulse/artificial-neural-networks-advantages-disadvantages-maad-m-mijwel (accessed on 27 January 2018).
35. Le, X.H.; Ho, H.V.; Lee, G.; Jung, S. Application of long short-term memory (LSTM) neural network for flood forecasting. *Water* **2019**, *11*, 1387. [CrossRef]
36. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models Part I—A discussion of principles. *J. Hydrol. Eng.* **1970**, *10*, 282–290. [CrossRef]
37. Liu, M.; Huang, Y.; Li, Z.; Tong, B.; Liu, Z.; Sun, M.; Jiang, F.; Zhang, H. The applicability of LSTM-KNN model for real-time flood forecasting in different climate zones in China. *Water* **2020**, *12*, 440. [CrossRef]
38. Ding, Y.; Zhu, Y.; Wu, Y.; Jun, F.; Cheng, Z. Spatio-temporal attention LSTM model for flood forecasting. In Proceedings of the International Conference on Internet of Things (IThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Atlanta, GA, USA, 14–17 July 2019; pp. 458–465.
39. Song, T.; Ding, W.; Wu, J.; Liu, H.; Zhou, H.; Chu, J. Flash flood forecasting based on long short-term memory networks. *Water* **2019**, *12*, 109. [CrossRef]
40. Gao, S.; Huang, Y.; Zhang, S.; Han, J.; Wang, G.; Zhang, M.; Lin, Q. Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. *J. Hydrol.* **2020**, *589*, 125188. [CrossRef]
41. Muhammad, A.U.; Li, X.; Feng, J. Using LSTM GRU and hybrid models for streamflow forecasting. In *Machine Learning and Intelligent Communications. MLICOM 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*; Zhai, X., Chen, B., Zhu, K., Eds.; Springer: Cham, Switzerland, 2019; Volume 294, pp. 510–524.
42. Miau, S.; Hung, W.H. River flooding forecasting and anomaly detection based on deep learning. *IEEE Access* **2020**, *8*, 198384–198402. [CrossRef]
43. Ebtehaj, I.; Sammen, S.S.; Sidek, L.M.; Sihag, P.; Al-Janabi, A.M.S.; Chau, K.W.; Bonakdari, H. Prediction of Daily Water Level Using New Hybridized GS-GMDH and ANFIS-FCM Models. *Eng. Appl. Comput. Fluid Mech.* **2021**, *15*, 1343–1361. [CrossRef]
44. Fu, M.; Fan, T.; Ding, Z.; Salih, S.Q.; Al-Ansari, N.; Yaseen, Z.M. Deep Learning Data-Intelligence Model Based on Adjusted Forecasting Window Scale: Application in Daily Streamflow Simulation. *IEEE Access* **2020**, *8*, 32632–32651. [CrossRef]
45. Water Resources Management Information System (WAMIS). Available online: http://www.wamis.go.kr/ (accessed on 10 June 2022).
46. Korea Meteorological Administration Weather Data Service. Available online: https://data.kma.go.kr/ (accessed on 10 June 2022).
47. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
48. Pascanu, R.; Mikolov, T.; Bengio, Y. On the difficulty of training recurrent neural networks. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1310–1318.
49. Hochreiter, S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int. J. Uncertain. Fuzziness Knowl. -Based Syst.* **1998**, *6*, 107–116. [CrossRef]
50. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.