Realtime Flood Forecasting: River flow analysis using Machine Learning Techniques

Sudasingha H. [1]*, Koswatte S.[2]

[1]Student, Faculty of Geomatics, Sabaragamuwa University of Sri Lanka

[2]Professor in GIS, Faculty of Geomatics, Sabaragamuwa University of Sri Lanka

himal.sudasingha@gmail.com

Abstract *Kalu Ganga river basin in Sri Lanka is highly susceptible during the monsoon seasons, which frequently causes devastating floods, disrupting the lives of local communities. Addressing this critical issue, this research focusses on enhancing the accuracy of water level predictions in the Kalu Ganga river basin. Traditional methods of water level prediction have proven to be inefficient, highlighting the need for more advanced and accurate forecasting techniques. This study developed a rolling forecasting system aimed at predicting future water levels at the Ratnapura station in the Kalu Ganga using several machine learning algorithms. Data collected over a period of 10 months was utilized, with 75% allocated for training and the remainder for testing and validation. We employed four machine learning models, namely Support Vector Regression (SVR), Random Forest (RF), Artificial Neural Network (ANN), and Long Short-Term Memory (LSTM) were used for prediction. All models demonstrated high accuracy in predicting water levels, with the ANN and LSTM models marginally outperforming the SVR and RF in most cases. However, challenges were noted in accurately predicting peak water levels across all models. The limited 10-month data duration potentially constrained the models' predictive capability over extended periods. In conclusion, the rolling forecasting system developed in this study holds promise for integration into the rivernet.lk system, potentially enhancing flood management capabilities. Further research using a larger dataset spanning over multiple years is recommended to improve the accuracy of the models in predicting water levels over longer periods. This study offers insights that could advance water resource management and flood mitigation efforts in Sri Lanka.*

Keywords: Flood forecasting, LSTM, Kalu Ganga, rolling forecasting, Machine learning

Introduction

Accurate flood flow forecasting is crucial for effective water resource management, particularly in populated regions near major rivers, to mitigate risks to society and the economy (Le et al., 2019). Hydrological forecasting employs mathematical and statistical modeling, including stochastic models like Autoregressive Moving Average (ARMA) and the Markov method (Elsafi, 2014). Recently, Machine Learning (ML) techniques such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) have gained traction for their ability to model complex nonlinear relationships without needing detailed physical process knowledge (Dazzi et al., 2021)

Floods, affecting over 2.3 billion people and causing $662 billion in economic losses in the past two decades, are expected to increase in frequency and severity due to climate change and population growth near water bodies (Kundzewicz et al., 2014; Winsemius et al., 2016). Ratnapura district in Sri Lanka, with its high annual rainfall and vulnerable geography, faces significant flood risks, particularly from the Kalu River, which has a low gradient and bottleneck that exacerbate flooding (Department of Census and Statistics, 2012; Edirisooriya et al., 2018; Nandalal, 2009).

This research presents a short-term real-time rolling forecasting system for floods in Ratnapura, analyzing water levels of Kalu Ganga, Way Ganga, and Denawaka Ganga using ML techniques. Moreover, the study evaluated forecasting parameters to improve the accuracy to implement the system on the rivernet.lk web service for public access, focusing on effective forecasting methods and accuracy enhancements.

Literature Review

Floods are among the most devastating natural disasters, exacerbated by climate change and human activities such as urbanization and deforestation. They result in significant loss of life, damage to infrastructure, and displacement of populations, along with contamination of water sources and disruption of transportation networks. Effective flood mitigation requires comprehensive risk assessment and management strategies that consider uncertainties in climate models and social behaviors (Kundzewicz et al., 2014). Floods can be categorized into several types, including flash floods, urban floods, pluvial floods, and river floods, each with distinct causes. Flash floods occur rapidly due to intense rainfall, particularly in hilly areas. Urban floods arise from heavy rain overwhelming drainage systems, while pluvial floods result from excess rainfall in urban areas. River floods are caused by rivers overflowing, typically due to heavy rain or snowmelt (Merz et al., 2010). Other types include coastal, groundwater, and compound floods, which occur simultaneously or successively, complicating management efforts (Kundzewicz et al., 2014).

In Ratnapura, Sri Lanka, significant flooding has been recorded from 1883 to 2017, particularly during the monsoon season, with the Kalu River being a primary source due to its low gradient and bottleneck downstream (Edirisooriya et al., 2018; Nandalal, 2009).

ML has emerged as a transformative technology in flood risk management, identifying patterns in data to address complex problems through various algorithms (M. I. Jordan & T. M. Mitchell, 2015; Mahesh, 2018). The choice of algorithm is crucial, as it depends on the specific problem, highlighting the need for tailored approaches in flood risk assessment (Mahesh, 2018). ML algorithms are categorized into four types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning (Sarker, 2021).

In traditional ML, selecting the appropriate feature space is critical. For binary classification, instances may be mapped to an intermediary feature space if they cannot be separated directly. Deep learning (DL) architectures, which utilize multiple-layer neural networks, can efficiently model complex problems and represent non-linear functions (Ludovic Arnold et al., 2016). The backpropagation algorithm, developed in the 1970s and widely recognized in the 1980s, facilitated self-directed learning and automated feature extraction (Shrestha & Mahmood, 2019).

DL algorithms include CNNs for image recognition, Recurrent Neural Networks (RNNs) for sequence-based applications, and Deep Belief Networks (DBNs) for unsupervised learning tasks (LeCun et al., 2015). Generative Adversarial Networks (GANs) are employed for generating synthetic data and have advanced fields such as computer vision and natural language processing (Goodfellow et al., n.d.). RNNs and LSTM networks excel in sequential data processing, with LSTMs addressing the vanishing gradient problem (Hochreiter & Schmidhuber, 1997).

Flood risk management is essential for mitigating the impacts of floods, involving both structural and nonstructural measures such as early-warning systems (Jongman et al., 2012; Liu et al., 2018). Accurate river water level prediction is crucial for effective water resource management and flood control (Kim et al., 2022). Forecasting tools include conceptual models, physically based models, and "black box" models, with the latter relying on historical data to capture complex interactions without requiring physical process understanding (Mosavi et al., 2018; Zounemat-Kermani et al., 2020).

Short-term flood prediction is particularly challenging in densely populated areas, necessitating timely warnings to minimize damage (Zhang et al., 2018). Various ML/DL models have been implemented for flood prediction, including ANN, Multilayer Perceptron (MLP), Adaptive Neuro-Fuzzy Inference System (ANFIS), Wavelet Neural Networks (WNN), SVM , Decision Tree (DT), Random Forest (RF), and hybrid models (Dazzi et al., 2021).

The integration of ML into flood risk management offers promising advancements in predicting and mitigating flood events. As climate change continues to influence flood patterns, the need for reliable and accurate forecasting methods becomes increasingly critical. The ongoing development and application of ML techniques will play a pivotal role in enhancing flood risk management strategies globally.

Methodology

a. Data Used

The Irrigation Department of Sri Lanka has implemented water level gauges along the Kalu Ganga river and its seven tributaries to monitor water levels at a one-minute resolution. The gauges transmit data via GSM technology to a public website that displays real-time water levels, improving data collection in the basin. For this research, 10 months of one-minute interval data was collected from December 2020 to October 2021 to analyze water levels and flood risk in the Kalu Ganga basin.

b. Study area

The study area for this research project is the Kalu Ganga river basin (Figure 1), focusing on water level gauges in Ratnapura, Kahawatta, and Palmadulla. Ratnapura is the most flood-affected area in the basin, making it critical for study. The Denawaka Ganga at Palmadulla and the Wey Ganga in Kahawatta are also important tributaries for this research. Data collected from these three gauges will be used to predict the water level of the Kalu Ganga in Ratnapura, providing insights into flood risk in the area. This study is vital for informing water management and flood mitigation decisions in the Kalu Ganga river basin.

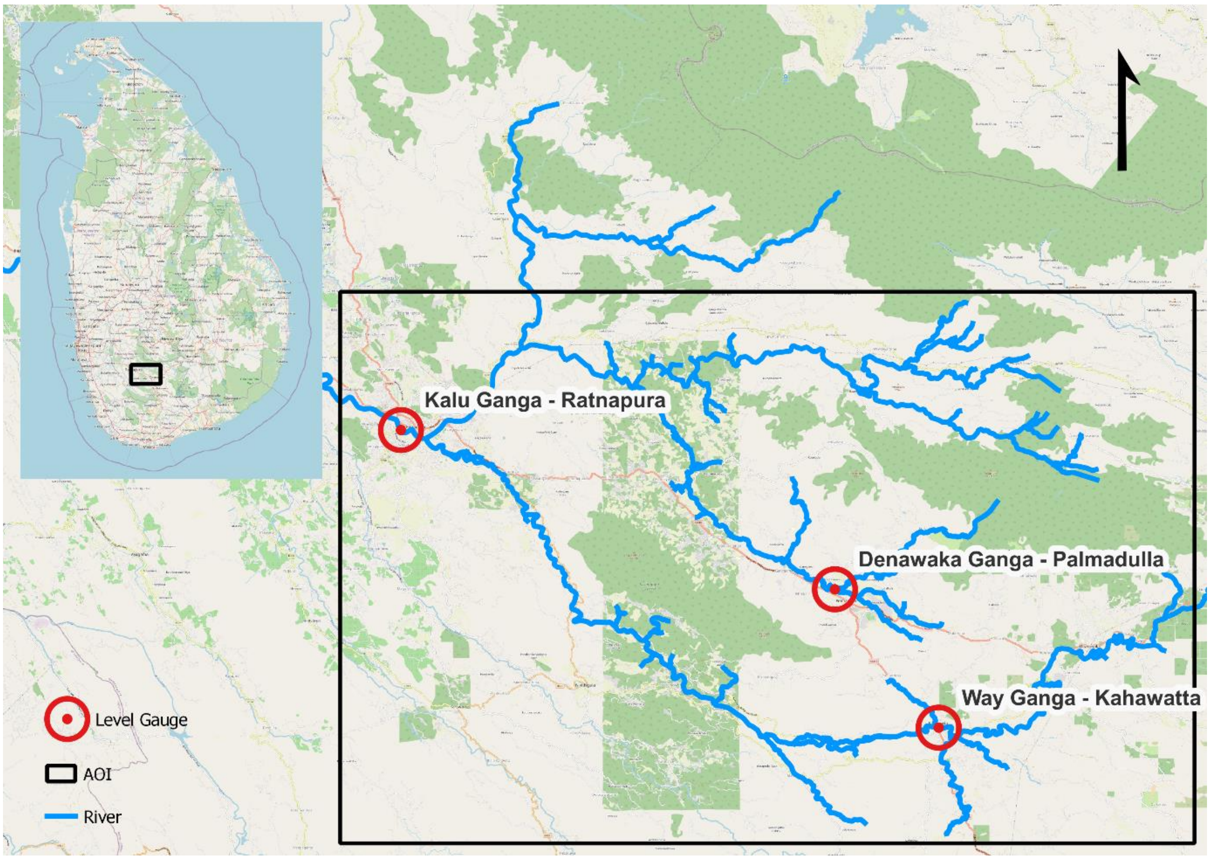

Figure 1 Study Area

c. Software used in processing and analyzing

The initial examination of the dataset using Microsoft Excel provided a first glance at its structure and contents. However, Excel has limitations for complex data manipulation, transformation, and analysis. Python's pandas and NumPy libraries were chosen for preprocessing due to their powerful data manipulation tools, allowing efficient cleaning and transformation of large datasets. The tensorflow and keras libraries were used to develop ML models, including ANN, LSTM, SVR, and RF, based on their suitability for the specific problem. The keras.metrics library calculated metrics such as R-squared (R^2), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) to evaluate model performance. Matplotlib and seaborn libraries were employed to visualize the data, and Jupyter Notebook through Anaconda facilitated the entire process.

d. Data Pre-Processing

The study involved preparing a dataset received as a spreadsheet with columns for various locations, with each row representing a data point at 1-minute intervals. Initial inspection revealed that some values in the Ratnapura column were zero, not missing. Linear interpolation was applied to replace these zeros with approximated values.

Next, a Python script was used to extract 10-minute interval data from the spreadsheet, which was saved as a CSV file for further analysis. Outlier detection using Python's describe() function showed no outliers, likely due to prior data correction during collection. Finally, data irrelevant to the study area was removed, focusing on water level data from specific tributaries, including Ratnapura, Pelmadulla, and Kahawaththa, while excluding others.

e. Model Training and Saving Model Parameters

The study utilized Python 3.11.3 and Jupyter Notebook to train four ML/DL models: Support Vector Regression (SVR), RF, ANN, and LSTM Network. A total of 16 scripts were developed for predicting water levels at various future timestamps, including 30 minutes to 3 hours ahead. The code was organized into sections, each covering data preprocessing, model training, and evaluation. The dataset was prepared as outlined in the Data Preparation section, ensuring optimal accuracy for the models. The ANN model for 30-minute predictions was highlighted, beginning with the importation of essential libraries, followed by defining future and lookback periods for time series analysis. Data was imported using Pandas, and the first 3000 data points were visualized using Matplotlib to identify trends. A custom function, df_to_x_y(), was created to separate dependent and independent variables based on the defined periods. The dataset was divided into training (75%), validation (15%), and testing (10%) sets. The LSTM model was constructed using Keras, featuring an input layer, an LSTM layer, and dense layers for output. The model was compiled with MSE as the loss function and Adam as the optimizer, with a learning rate of 0.0001. The training process involved fitting the model to the training data over 50 epochs, utilizing a callback function to save the best model based on performance.

f. Model Loading and Prediction

This section outlines the Python code used to load a saved ML/DL model and predict future water levels. Developed in Python 3.11.3 within Jupyter Notebook, the code is organized into sectioned snippets for clarity.

A custom function, plot_predictions, is defined to visualize predicted and actual water levels. It takes parameters for the model, independent variables (x), actual dependent variable values (y), optional start and end indices, a title for the plot, and a folder path for saving results. The function predicts values using the trained model, creates a Pandas DataFrame to compare predicted and actual values, and calculates performance metrics such as R^2, MSE and RMSE. It generates a scatter plot and a line plot to illustrate the predictions and actual trends, saving the results as a CSV file if a folder path is provided.

The function is subsequently called to predict water levels using the chosen model on the training, validation, and test datasets, each time plotting the actual versus predicted values and saving the results as specified. This structured approach ensures effective visualization and evaluation of the model's performance in predicting water levels.

Results and Discussion

The results obtained in this study can be divided into two main parts: (a) the intermediate results and (b) the final results that were derived from intermediate results. The intermediate results include the performance metrics and validation results obtained during the training and evaluation of various ML and deep learning models. The final result of this study denotes the best performing model that can accurately predict the future water levels based on the historical data. This was achieved by analyzing the intermediate results and selecting the model with the best performance metrics, such as the lowest MSE and highest correlation coefficient (R^2 score), among others.

a. Intermediate Results

The research used four major approaches for predicting future water levels: SVR, RF, ANN, and LSTM. Each approach has its own intermediate results that can be analyzed to evaluate the performance of the models. This section discuss the results obtained from each approach separately.

Support Vector Regression (SVR) Approach

Figure 2 below displays the actual versus predicted results of SVR models for forecasting water levels 30min, 1h, 2h, and 3h into the future.

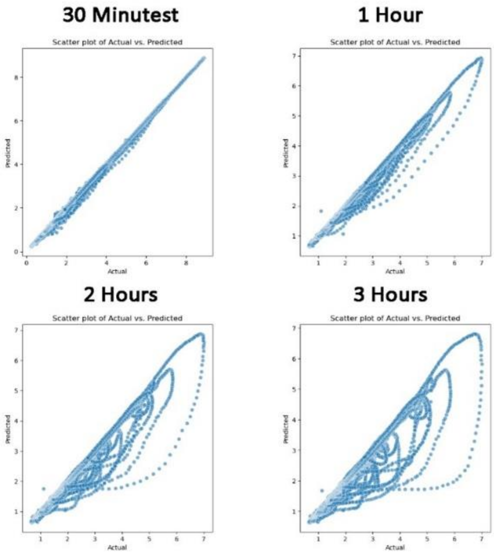

Figure 2 SVR actual and predicted scatter plots

Table 1 presents the evaluation metrics for the SVR models for each prediction time frame. The metrics used for evaluation include MSE, RMSE, and coefficient of determination (R^2 score).

Table 1 SVR evaluation criteria

Evaluation Criteria	Duration			
	30m	1h	2h	3h
R^2	1.00	0.99	0.96	0.91
MSE	0.00	0.01	0.05	0.11
RMSE	0.06	0.11	0.22	0.33

The obtained intermediate results of the SVR approach reveal that it shows promising results for 30-minute and 1-hour into the future predictions, as depicted in Figure 2. Additionally, it provides better insights into peak values. However, for 2-hour and 3-hour into the future predictions, the SVR model was not performing well and could not accurately predict peak water levels above 5 meters.

The evaluation criteria for the SVR model are presented in Table 1. It can be observed that the model performed exceptionally well for the 30-minute and 1-hour into the future predictions with R^2 values of 1.00 and 0.99, respectively. This implies that the SVR model can explain 100% and 99% of the variability in the water level data for the 30-minute and 1-hour forecasts, respectively. The MSE and RMSE values for these timeframes are also quite low, which indicates the model's high level of accuracy. However, for the 2-hour and 3-hour into the future predictions, the R^2 values are relatively lower at 0.96 and 0.91, respectively, indicating a decrease in model performance. Additionally, the MSE and RMSE values are higher for these timeframes, suggesting lower accuracy.

Random Forest Regression Approach

Error! Reference source not found. displays the actual versus predicted results of RF models for forecasting water levels 30min, 1h, 2h, and 3h into the future.

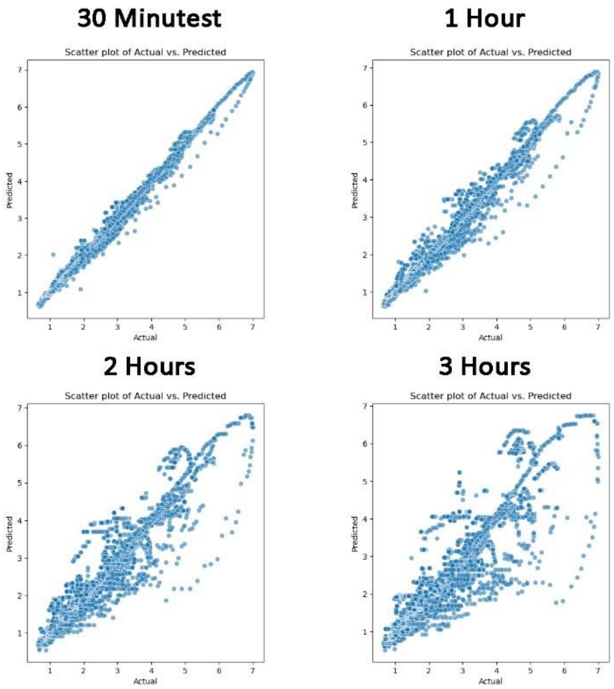

Figure 3 RF actual and predicted scatter plots

Figure 3 presents the evaluation metrics for the RF models for each prediction time frame. The metrics used for evaluation include MSE, RMSE and coefficient of determination (R^2 score).

Table 2 RF evaluation criteria

Evaluation Criteria	Duration			
	30m	**1h**	**2h**	**3h**
R^2	1.00	0.98	0.92	0.84
MSE	0.01	0.03	0.1	0.2
RMSE	0.08	0.16	0.31	0.44

The results of RF regression are presented in Figure 3, which shows the scatter plot of actual vs predicted values. The RF model did not perform well in predicting the water levels for any of the forecast horizons. It failed to accurately predict even the lower water level values such as 2m and 3m, which are not the peak values leading to flooding in 2-

hour and 3-hour timeframes. Therefore, the RF approach is not a suitable technique for forecasting future water levels of Kaluganga at Ratnapura based on the results obtained.

As seen in the evaluation criteria of the RF regressor in Table 2, the model has shown a good performance for the 30-minute and 1-hour into the future predictions with R^2 values of 1.00 and 0.98, respectively. However, the performance of the model significantly decreases for the 2-hour and 3-hour into the future predictions, with R^2 values of 0.92 and 0.84, respectively. The MSE and RMSE also follow a similar pattern, indicating the model's limitations in predicting the peak water levels accurately for longer timeframes.

Artificial Neural Network Approach

Figure 4 displays the actual versus predicted results of ANN models for forecasting water levels 30min, 1h, 2h, and 3h into the future.

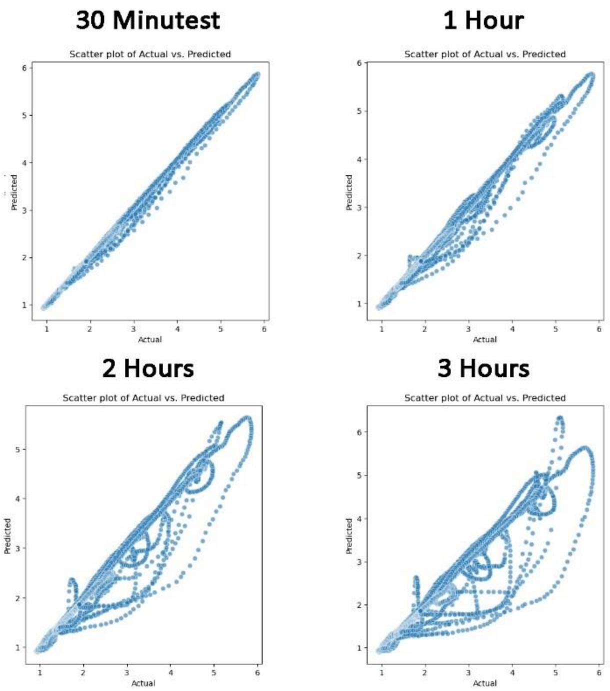

Figure 4 ANN actual and predicted scatter plots

Table 3 presents the evaluation metrics for the ANN models for each prediction time frame. The metrics used for evaluation include MSE, RMSE, and coefficient of determination (R^2 score).

Table 3 ANN evaluation criteria

Evaluation Criteria	Duration			
	30m	1h	2h	3h
R^2	1.00	0.99	0.96	0.90
MSE	0.00	0.01	0.06	0.13
RMSE	0.06	0.12	0.24	0.37

The ANN approach, as depicted in Figure 4, shows promising results for 30-minute, 1 hour, and 2-hour predictions. However, it appears to struggle with accurately predicting peak water levels for the 3-hour timeframe. This may be due to the complexity of the network and the amount of data used, as well as the potential for overfitting to the training data. Further optimization and tweaking of the model may improve its accuracy for longer-term predictions.

Based on the evaluation criteria of R^2, MSE and RMSE for the ANN model shown in Table 3, the results indicate that ANN has performed well in predicting water levels for the 30-minute, 1-hour, and 2-hour timeframes. However, for the 3-hour timeframe, the ANN model showed a decrease in accuracy as it could not accurately predict peak water levels. The values for all timeframes are above 0.9, indicating a strong correlation between the actual and predicted values. The MSE and RMSE values are also relatively low for the 30-minute and 1-hour predictions but increase for the 2-hour and 3-hour predictions, indicating a decrease in model performance for longer timeframes. Overall, the ANN model has shown potential for accurately predicting water levels, but further improvements are needed for longer-term forecasting.

Long-Short Term Memory Network Approach

Figure 5 displays the actual versus predicted results of LSTM network models for forecasting water levels 30min, 1h, 2h, and 3h into the future.

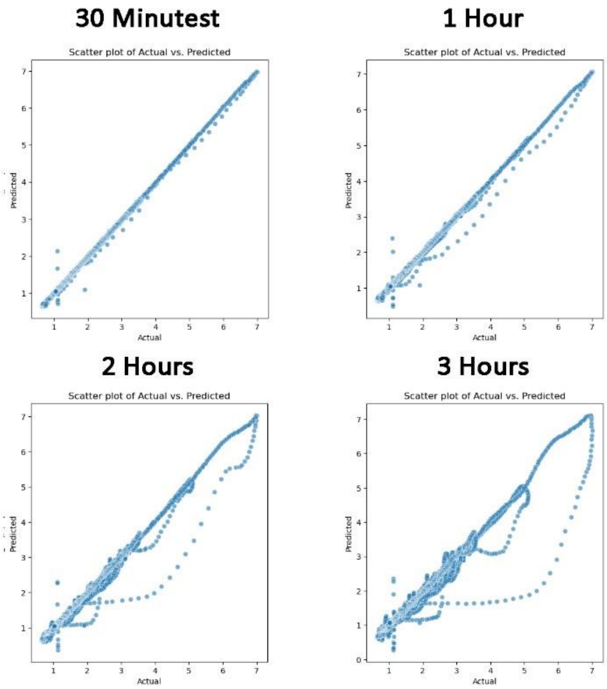

Figure 5 LSTM actual and predicted scatter plots

Table 4 presents the evaluation metrics for the LSTM network models for each prediction time frame. The metrics used for evaluation include MSE, RMSE, and coefficient of determination (R^2 score).

Table 4 LSTM evaluation criteria

Evaluation Criteria	**Duration**			
	30m	**1h**	**2h**	**3h**
R^2	1.00	1.00	0.99	0.96
MSE	0.00	0.00	0.01	0.04
RMSE	0.03	0.05	0.11	0.20

shown in Figure 5, the LSTM model performed very well for all timeframes. The actual and predicted scatter plots show that the LSTM model was able to predict water levels accurately for 30 minutes, 1-hour, and 2-hour into the future. While there were some difficulties in predicting peak water levels accurately for the 3-hour timeframe, overall, the LSTM model outperformed the other ML models in terms of accuracy and reliability for this specific water level forecasting task.

The LSTM network approach evaluation criteria, as shown in Table 4, exhibited highly accurate results for all timeframes. The 30-minute, 1-hour, and 2-hour predictions showed exceptional accuracy, while the 3-hour predictions had some complications in accurately predicting peak water levels. However, the overall performance of the LSTM model was highly promising, as evidenced by its R^2 values of 1.00, 1.00, 0.99, and 0.96 for the 30-minute, 1-hour, 2-hour, and 3-hour predictions, respectively. The MSE values for the 30-minute and 1-hour predictions were 0.00, indicating perfect accuracy, and the MSE values for the 2-hour and 3-hour predictions were 0.01 and 0.04, respectively, indicating highly accurate predictions. The RMSE values for the LSTM model were 0.03, 0.05, 0.11, and 0.20 for the 30-minute, 1-hour, 2-hour, and 3-hour predictions, respectively, further supporting the high accuracy of the model.

b. Final Results

This section presents the final results that were used to determine the best-performing model for the task based on the evaluation criteria of all the ML models trained and validated. The following figures (Figure 4.9 and Figure 4.10) illustrate a comparison of the evaluation criteria of all the trained models and timeframes, which led to selecting the best performing model for the work.

Table 4 Evaluation Criteria comparison for all models

Model	Evaluation Criteria	Duration			
		30m	1h	2h	3h
SVR	**R^2**	1.00	0.99	0.96	0.91
	MSE	0.00	0.01	0.05	0.11
	RMSE	0.06	0.11	0.22	0.33
RF	**R^2**	1.00	0.98	0.92	0.84
	MSE	0.01	0.03	0.10	0.20
	RMSE	0.08	0.16	0.31	0.44
ANN	**R^2**	1.00	0.99	0.96	0.90
	MSE	0.00	0.01	0.06	0.13
	RMSE	0.06	0.12	0.24	0.37
LSTM	**R^2**	1.00	1.00	0.99	0.96
	MSE	0.00	0.00	0.01	0.04
	RMSE	0.03	0.05	0.11	0.20

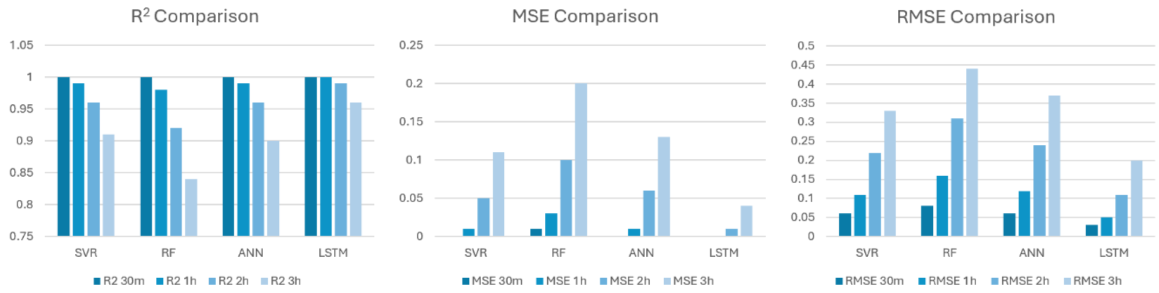

Figure 6 Evaluation Criteria comparison chart

The evaluation criteria in Table 5 and Evaluation criteria comparison in Figure 6 shows the performance of four different models used for predicting the water levels at different timeframes. SVR shows very high accuracy in predicting the water levels for 30 minutes into the future, with an R^2 value of 1.00 and an MSE value of 0.00, indicating perfect predictions. However, its performance decreases as the timeframe increases, with an R^2 value of 0.91 and an RMSE value of 0.33 for the 3-hour timeframe.

RF model shows good results for 30 minutes and 1-hour predictions with an R^2 value of 1.00 and 0.98, respectively. However, its performance also decreases significantly as the timeframe increases, with an R^2 value of 0.84 and an RMSE value of 0.44 for the 3-hour timeframe.

The ANN model also shows promising results for 30 minutes, 1 hour, and 2 hours predictions, with an R^2 value above 0.96. However, similar to SVR and RF models, its performance decreases for the 3-hour timeframe with an R^2 value of 0.90 and an RMSE value of 0.37.

On the other hand, the LSTM model shows the best performance across all timeframes, with perfect predictions for the 30-minute timeframe. Moreover, it continues to perform very well for the longer timeframes, with an R^2 value of 0.99 and an RMSE value of 0.11 for the 2-hour timeframe and an R^2 value of 0.96 and an RMSE value of 0.20 for the 3-hour timeframe. Therefore, the LSTM model is the most suitable model for predicting water levels accurately across different timeframes.

Conclusion and Recommendation

Water resources management is a critical issue in many regions of the world, including Sri Lanka. Effective management of water resources requires accurate and timely information about the water levels in rivers and other bodies of water. In recent years, advances in data collection and ML techniques have made it possible to build predictive models that can forecast water levels with high accuracy. This research focused on the use of ML techniques to predict water levels in the Kalu Ganga river at Ratnapura station.

The dataset used for training, testing and validating the models only covers a period of 10 months, which may not be sufficient to capture the full range of variability in water level changes in the Kaluganga basin. The basin is subject to the influence of two distinct monsoon seasons that occur at different times of the year, resulting in complex and dynamic hydrological conditions. By training the models on a limited dataset, there is a risk of overfitting to the available data and failing to generalize to other time periods or environmental conditions. This could lead to inaccurate predictions when the models are used for real-time flood forecasting or other applications. To address this limitation, it would be beneficial to collect and incorporate data from multiple years into the model training process. By doing so, the models can learn to capture the longer-term trends and patterns in water level changes and better account for the effects of interannual variability and changing environmental conditions.

The developed rolling forecasting system can be implemented into the rivernet.lk system, which is currently managed by a private firm, as an embedded system to access real-time data from the Kalu Ganga river and its tributaries and provide future predicted water levels at the Ratnapura station. The purpose of this integration is to provide access to real-time water level data and display future predicted water levels of Kalu Ganga at Ratnapura station. The trained models can be implemented on rivernet.lk servers as an embedded system, enabling users to access accurate and reliable information on future water levels. By incorporating this system, the management of water resources in the region can be significantly improved, leading to better decision-making and more effective planning for future events. This system can be a valuable tool for decision-makers, water resource managers, and the public to make informed decisions related to flood preparedness, and other related activities.

References

Dazzi, S., Vacondio, R., & Mignosa, P. (2021). Flood stage forecasting using machine-learning methods: A case study on the parma river (italy). *Water (Switzerland)*, *13*(12). https://doi.org/10.3390/w13121612

Department of Census and Statistics. (2012). *Census of Population and Housing*.

Edirisooriya, K. V.D., Vitanage, N. S., Uluwaduge, P., & Senevirathna, E. M. T. K. (2018). Understanding Disaster Risk and Mitigation Special Reference to Ratnapura District. *Procedia Engineering*, *212*, 1034–1039. https://doi.org/10.1016/j.proeng.2018.01.133

Elsafi, S. H. (2014). Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. In *Alexandria Engineering Journal* (Vol. 53, Issue 3, pp. 655–662). Elsevier B.V. https://doi.org/10.1016/j.aej.2014.06.010

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (n.d.). *Generative Adversarial Nets*. http://www.github.com/goodfeli/adversarial

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Jongman, B., Ward, P. J., & Aerts, J. C. J. H. (2012). Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change*, *22*(4), 823–835. https://doi.org/10.1016/j.gloenvcha.2012.07.004

Kim, D., Han, H., Wang, W., & Kim, H. S. (2022). Improvement of Deep Learning Models for River Water Level Prediction Using Complex Network Method. *Water (Switzerland)*, *14*(3). https://doi.org/10.3390/w14030466

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G. R., Kron, W., Benito, G., Honda, Y., Takahashi, K., & Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal*, *59*(1), 1–28. https://doi.org/10.1080/02626667.2013.857411

Le, X. H., Ho, H. V., Lee, G., & Jung, S. (2019). Application of Long Short-Term Memory (LSTM) neural network for flood forecasting. *Water (Switzerland)*, *11*(7). https://doi.org/10.3390/w11071387

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Liu, C., Guo, L., Ye, L., Zhang, S., Zhao, Y., & Song, T. (2018). A review of advances in China’s flash flood early-warning system. *Natural Hazards*, *92*(2), 619–634. https://doi.org/10.1007/s11069-018-3173-7

Ludovic Arnold, Sébastien Rebecchi, Sylvain Chevallier, & Hélène Paugam-Moisy. (2016). *An Introduction to Deep Learning*. Ciaco.

M. I. Jordan, & T. M. Mitchell. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 253–255. https://doi.org/10.1126/science.aac4520

Mahesh, B. (2018). Machine Learning Algorithms-A Review Machine Learning Algorithms-A Review View project Self Flowing Generator View project Batta Mahesh Independent Researcher Machine Learning Algorithms-A Review. *International Journal of Science and Research*. https://doi.org/10.21275/ART20203995

Merz, B., Kreibich, H., Schwarze, R., & Thieken, A. (2010). Review article "Assessment of economic flood damage." *Natural Hazards and Earth System Sciences*, *10*(8), 1697–1724. https://doi.org/10.5194/nhess-10-1697-2010

Mosavi, A., Ozturk, P., & Chau, K. (2018). Flood Prediction Using Machine Learning Models: Literature Review. *Water*, *10*(11), 1536. https://doi.org/10.3390/w10111536

Nandalal, K. D. W. (2009). Use of a hydrodynamic model to forecast floods of Kalu River in Sri Lanka. *Journal of Flood Risk Management*, *2*(3), 151–158. https://doi.org/10.1111/j.1753-318X.2009.01032.x

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. In *SN Computer Science* (Vol. 2, Issue 3). Springer. https://doi.org/10.1007/s42979-021-00592-x

Shrestha, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. In *IEEE Access* (Vol. 7, pp. 53040–53065). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2019.2912200

Winsemius, H. C., Aerts, J. C. J. H., van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., Kwadijk, J. C. J., Ligtvoet, W., Lucas, P. L., van Vuuren, D. P., & Ward, P. J. (2016). Global drivers of future river flood risk. *Nature Climate Change*, *6*(4), 381–385. https://doi.org/10.1038/nclimate2893

Zhang, J., Hou, G., Ma, B., & Hua, W. (2018). Operating characteristic information extraction of flood discharge structure based on complete ensemble empirical mode decomposition with adaptive noise and permutation entropy. *Journal of Vibration and Control*, *24*(22), 5291–5301. https://doi.org/10.1177/1077546317750979

Zounemat-Kermani, M., Matta, E., Cominola, A., Xia, X., Zhang, Q., Liang, Q., & Hinkelmann, R. (2020). Neurocomputing in surface water hydrology and hydraulics: A review of two decades retrospective, current status and future prospects. *Journal of Hydrology*, *588*, 125085. https://doi.org/10.1016/j.jhydrol.2020.125085