Machine Learning Models for Water Level Prediction in Rapid Urban Streams: Case of Mburicaó, Asunción Paraguay

Mathias Aguilar
Facultad de Ingeniería - Universidad Nacional de Asunción
San Lorenzo, Paraguay
maguilar@fiuna.edu.py

Héctor Velázquez
Facultad de Ingeniería - Universidad Nacional de Asunción
San Lorenzo, Paraguay
hvelazquez@fiuna.edu.py

Diego H. Stalder
Facultad de Ingeniería - Universidad Nacional de Asunción
San Lorenzo, Paraguay
zz

Andres Wehrle
Facultad de Ingeniería - Universidad Nacional de Asunción
San Lorenzo, Paraguay
awehrle@ing.una.py

***Abstract*—Urban streams in rapidly developing cities often experience sudden increases in water levels due to intense precipitation, posing risks to surrounding communities. This study explores a data-driven approach for predicting water level peaks and detecting critical flooding events in the Mburicao stream, located in Asunción, Paraguay. Using ten-minute interval data from three meteorological stations, we develop two predictive models: a linear regression model for estimating peak water levels, and a support vector machine classifier for identifying threshold-exceeding events. To inform model input selection, we apply both mutual information analysis and exhaustive search over sliding temporal windows of accumulated precipitation data. The models are evaluated using standard performance metrics, and the results highlight the relationship between lead time, feature selection, and predictive accuracy. The study offers insights into the integration of machine learning techniques in early warning systems for urban flood risk management.**

***Index Terms*—River Level Forecasting**

I. Introduction

Urban flooding remains a concern for cities that are growing quickly, especially in regions where rainfall patterns are becoming more unpredictable and the existing infrastructure struggles to handle brief but heavy rain events. The Mburicao stream in Asunción, Paraguay, flows through various densely populated neighborhoods and shows quick hydrological reactions to localized rainfall. These dynamics can lead to flash floods that affect daily life, harm infrastructure, and pose safety risks.

Efforts to address flood risks in Asunción have involved urban planning strategies like rainwater harvesting to decrease runoff volume and peak flow [1]. For operational flood response, the ability to forecast river level changes in real time is important. Short-term prediction tools can assist early warning systems and enable local authorities and communities to prepare or respond more effectively.

Hydrological models that rely on physical principles are commonly employed for forecasting. However, the need for calibration and the computational demands can restrict their usability in urban settings, especially when real-time execution is necessary [2], [3]. Additionally, these models might find it challenging to represent the rapidly changing, nonlinear behaviors often seen in urban hydrology.

Machine learning (ML) methods provide a flexible option. Several techniques, such as support vector machines, artificial neural networks (ANN), adaptive neuro-fuzzy inference systems (ANFIS), and ensemble decision trees have been used in river and lake level forecasting [4], [5], [6], [7], [8]. Some studies have incorporated these approaches into early warning frameworks [9], [10], and comparative analyses indicate that data-driven models can perform on par with traditional statistical approaches in certain contexts [11], [12].

In this context, our study explores a practical, data-informed framework for real-time flood prediction in the Mburicao stream. We utilize ten-minute interval precipitation and stream level data from three meteorological stations to create two predictive models: (i) a linear regression model aimed at estimating peak stream levels, and (ii) a support vector machine (SVM) classifier designed to determine if an event surpasses a specified critical threshold.

Feature selection is used to enhance accuracy and ensure interpretability by applying mutual information and exhaustive subset search on rainfall accumulation values within sliding windows. These strategies aim to lower dimensionality and pinpoint relevant temporal lags linked to rainfall patterns that can lead to flooding. The aim is to evaluate if straightforward, affordable models can adequately assist in real-time monitoring and alert systems in urban streams, where prompt and dependable predictions are important.

979-8-3503-4807-1/23/$31.00 ©2023 IEEE

II. Dataset

This study is based on environmental data collected from three meteorological stations located in the metropolitan area of Asunción, Paraguay (see Figure 1): San Ignacio de Loyola (SIL), Secretaría Nacional de Deporte (SND), and Aeropuerto Internacional Silvio Pettirossi (AISP). Table I summarizes the coordinates of the sensors' locations. Each station records rainfall at ten-minute intervals, while SIL uniquely includes water level measurements from the Mburicao stream.

We analyzed two distinct temporal subsets:

July-August 2021 (all stations)

July 2021-May 2022 (SIL data only, including both water level and rainfall measurements)

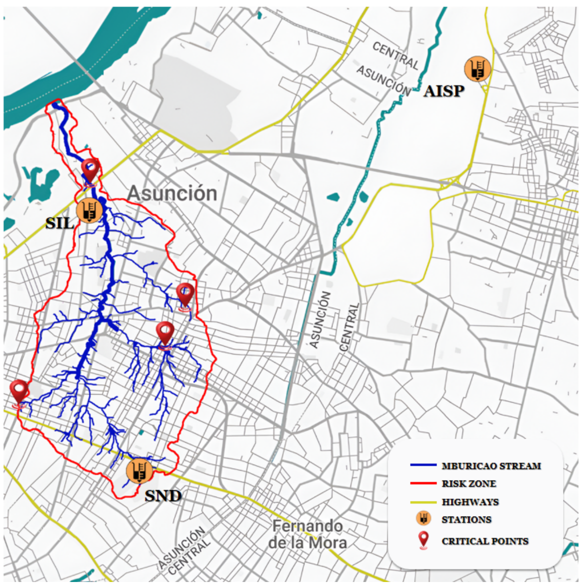

Fig. 1: Map showing the location of meteorological stations and critical points along the Mburicao stream.

TABLE I: Summary of measurement stations.

Station	City	(Lat., Long.)	Variables
SIL	Asunción	-25.27, -57.59	Precipitation, Stream level
SND	Asunción	-25.32, -57.58	Precipitation
AISP	Luque	-25.24, -57.51	Precipitation

The dataset captures a wide range of rainfall conditions, from light showers to intense storm events. Figure 2 illustrates the temporal variations in stream levels (target variable) and rainfall across the study period. For both regression and classification tasks, we used the SIL station's stream level measurements as the target variable. Before modeling, we aligned the raw data across all stations and addressed missing values through context-aware interpolation or removal, depending on data completeness and neighboring observations.

III. Feature Engineering

The quality and relevance of input features play a significant role in the performance of machine learning models, particularly in time-sensitive applications like real-time flood

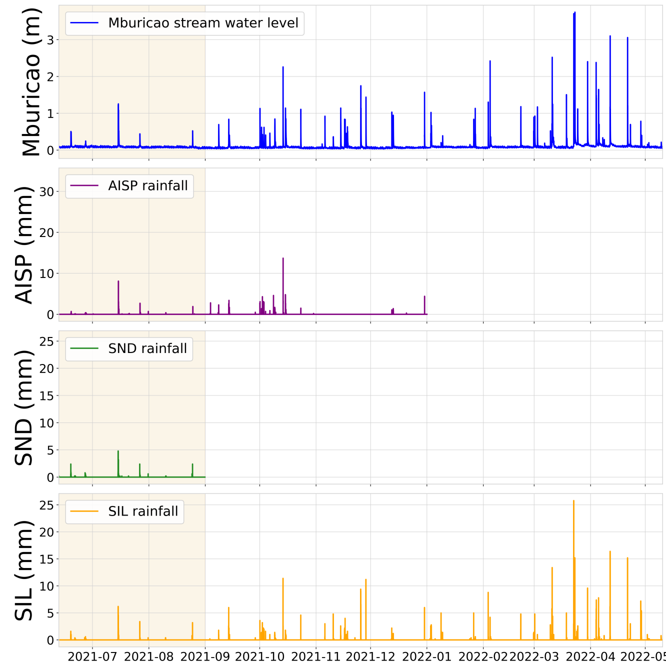

Fig. 2: Time series of rainfall and water level measurements at different stations.

forecasting. This study utilizes rainfall measurements from three stations to develop predictive features that reflect the relationship between precipitation and changes in stream levels in the Mburicao basin. Figure 3 shows an event that took place

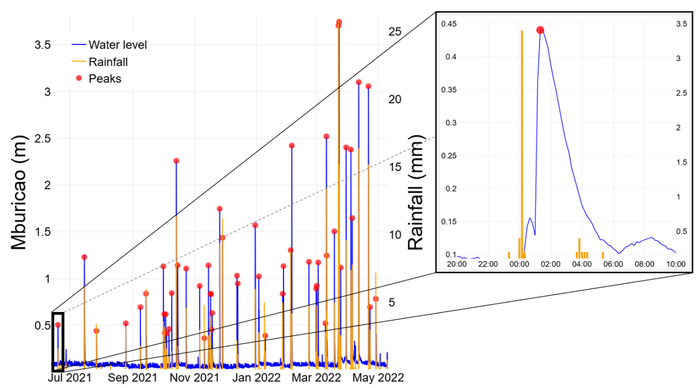

Fig. 3: Time series of stream level and rainfall for the event of July 27, 2021. The inset highlights the rapid increase in stream level following local precipitation measured at the SIL station.

on July 27, 2021, when the stream level reached about 0.45 m around midnight. This event is observed in which the rainfall recorded at the SIL station came before a quick increase in water level, characteristic of flash response behavior in the Mburicao stream. The rainfall began just before midnight and peaked in intensity, leading to a rapid rise in level, which was then followed by a slow decline over the next few hours. This event shows the brief delay between rainfall and stream response, highlighting the importance of using detailed rainfall data in predictive models.

A. Temporal Aggregation of Rainfall

To explore the relationship between rainfall and stream levels, accumulated precipitation was calculated over sliding windows of different lengths before each prediction timestamp t_0. The duration of these windows varies from 10 to 80 minutes and is calculated separately for each meteorological station. The features obtained, which include both raw rainfall and accumulated rainfall over various time periods, serve as input variables for regression and classification tasks.

Each accumulation window is defined as:

$$R_w^{\text{acc}} = \sum_{t=t_0-w_{\text{end}}}^{t_0-w_{\text{start}}} R(t), \tag{1}$$

where $R(t)$ represents the rainfall intensity at time t, and $w_{\text{start}}, w_{\text{end}}$ indicate the start and end of the window in minutes prior to t_0. This formulation enables the models to consider the impact of recent rainfall events over specific intervals.

B. Feature Design by Task

In the regression task focused on estimating peak water levels, features are gathered from rainfall values noted between 80 and 10 minutes before the observed maximum level. This selection aims to illustrate the delay between precipitation and the response of streams.

The classification task involves determining if a future event surpasses a set water level threshold. In this situation, features are created based on rainfall values gathered from the moment the rainfall threshold is surpassed until 50 minutes before the peak occurs. This approach focuses on the early indicators of significant events.

C. Feature Selection Strategy

Three methods are examined for reducing feature dimensionality and enhancing model interpretability:

- Mutual Information (MI): This method assesses the relationship between each feature and the target variable, organizing features according to their estimated contribution. The mutual information is calculated as:

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \tag{2}$$

  Figure 9 presents the mutual information scores for a selection of rainfall features.
- **Cross-Correlation**: This method evaluates the timing relationship between rainfall and stream level data. Cross-correlation assists in determining which lags provide useful information for predicting hydrological response. The findings are presented in Figure 5.
- **Exhaustive subset evaluation**: For each accumulation window, various feature subsets are evaluated through 5-fold cross-validation. The subset with the least validation error is kept for model training.

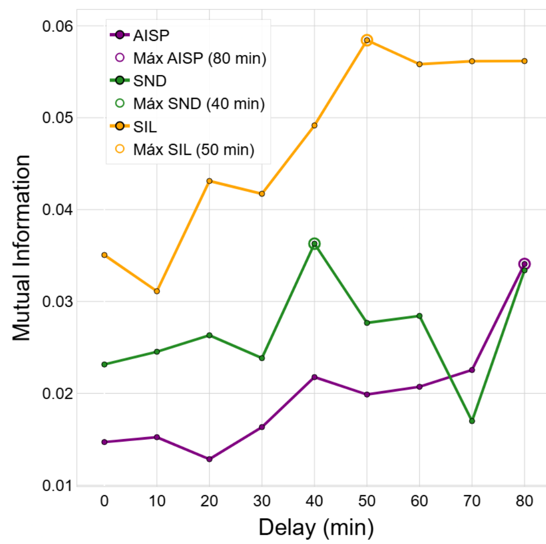

Fig. 4: Mutual information between rainfall-derived features and the target variable.

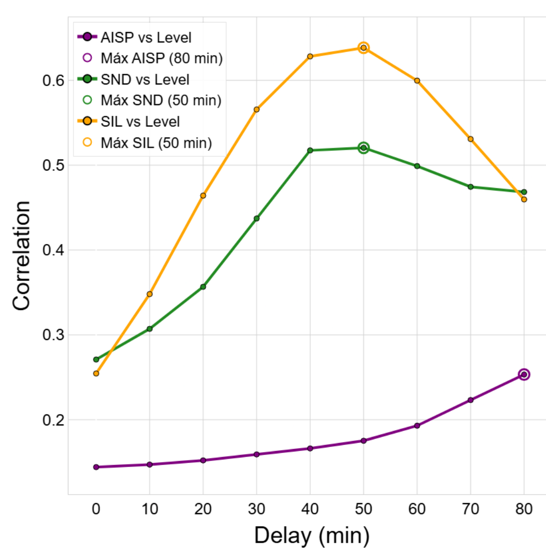

Fig. 5: Cross-correlation between rainfall time series and stream level response.

IV. Regression Model for Peak Prediction

We model the relationship between rainfall and streamflow peaks by defining a set of input features based on accumulated rainfall over sliding windows. Figure 6 shows the layout of the feature extraction process. Each window w relates to a specific accumulation interval, beginning at $t_0 - 80$ minutes and concluding at different moments leading up to t_0, the time of peak level.

For instance, the window $w = 1$ indicates the rainfall collected during the 10 minutes right before the peak, whereas

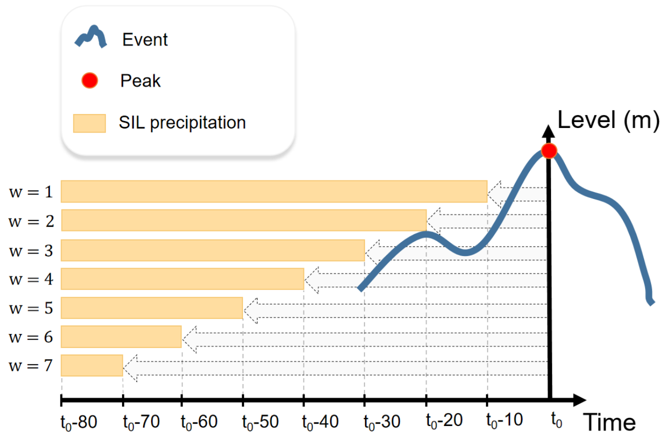

Fig. 6: Sliding rainfall accumulation windows used to construct features for the regression model. The time of peak water level t_0 serves as the reference point, and accumulated rainfall from the SIL station is computed over backward-looking intervals.

$w = 7$ encompasses rainfall from 80 to 70 minutes earlier. The accumulation windows serve as the foundation for the feature vectors that are utilized to train the linear regression model. This strategy helps us understand the role of recent and older precipitation in estimating peak levels, as well as the balance between lead time and predictive accuracy.

A. Linear Regression

Given a feature vector $\mathbf{X} \in \mathbb{R}^n$, which represents accumulated rainfall over selected time windows, the linear regression model predicts the corresponding peak stream level $y \in \mathbb{R}$ based on the following form:

$$\hat{y} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i, \tag{3}$$

where β_0 is the intercept and β_i are the coefficients estimated for each feature x_i. The model was implemented using the `LinearRegression` class from the `scikit-learn` library, and evaluated through 5-fold cross-validation to reduce the risk of overfitting. In each fold, 80% of the data is used for training and 20% for validation.

B. Performance Metrics

Model performance was evaluated using three standard metrics:

- R^2: the coefficient of determination, showing the proportion of variance accounted for by the model.
- RMSE: root mean squared error, measuring the average size of prediction errors.
- MAPE: mean absolute percentage error, providing a straightforward measure of prediction accuracy in relation to observed values.

These metrics offer a comprehensive perspective on the models' performance across various temporal aggregation windows.

C. Results by Time Window

Table II summarizes the linear regression results across various rainfall accumulation windows. Performance tends to degrade as the accumulation window extends further into the past, which may reflect reduced influence of older rainfall data on stream level peaks.

TABLE II: Linear regression results across different time windows.

Window (w)	R^2	RMSE (m)	MAPE (%)
1	0.8412	0.3304	27.61
2	0.8443	0.3272	28.07
3	0.8470	0.3243	28.19
4	0.8209	0.3509	31.27
5	0.6139	0.5152	36.72
6	0.0478	0.8090	73.72
7	0.0194	0.8210	76.39

As shown in Table II, windows covering the 10 to 30 minutes before peak events yield the highest R^2 values and the lowest RMSE and MAPE, suggesting these intervals carry the most relevant predictive signal.

Figure 7 presents a scatter plot comparing predicted and actual peak levels using the best-performing window configuration. Although some variance remains, the trend follows the ideal 1:1 line, indicating reasonable approximation for a first-order predictive system.

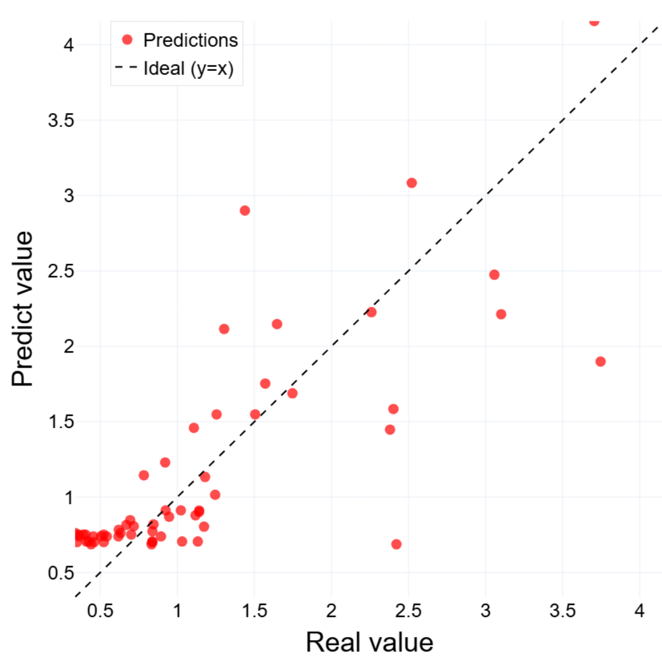

Fig. 7: Predicted vs. observed peak water levels using the best regression window (10–40 min before peak).

The results indicate a balance between early prediction and accuracy. Shorter windows near the peak offer improved performance, probably because they are closer to the hydrological response. However, longer windows not only extend the lead time but may also bring in additional noise, potentially affecting the quality of the predictions. The linear regression model shows enough consistency to function as a baseline. Its straightforward nature and clarity allow for effective integration into operational systems in real-time.

V. Event Classification Model (SVM)

Beyond estimating streamflow magnitudes, it is equally important to determine whether an incoming event should trigger a flood alert. In this section, we describe a binary classification approach using a Support Vector Machine (SVM) to predict whether the observed stream level will exceed a predefined threshold.

A. Problem Definition

The classification task involves a binary target variable $y \in \{0, 1\}$, where $y = 1$ signifies that the peak water level during the event meets or surpasses a critical threshold h_{crit}, and $y = 0$ indicates that it does not. The model receives a feature vector $\mathbf{X} \in \mathbb{R}^n$ created from rainfall data collected in sliding time windows before the event.

We assess model robustness by examining different severity levels and the associated thresholds.

$$h_{\text{crit}} \in \{0.5\,\text{m},\ 1.0\,\text{m},\ 1.5\,\text{m},\ 2.0\,\text{m}\}$$

Every threshold establishes a unique classification situation. The experiments discussed here concentrate on $h_{\text{crit}} = 1.0\,\text{m}$, a typical operational threshold often applied in flood monitoring. These definitions are illustrated in Figure 8, which outlines the temporal dynamics that guide both labeling and input feature construction. This representation helps visualize the relationship between rainfall timing and classification output, which is particularly relevant in early warning scenarios.

B. SVM Formulation

The SVM classifier transforms the input features into a higher-dimensional space through a nonlinear process and identifies a decision boundary that aims to increase the margin between classes. The optimization problem for the soft-margin SVM with an RBF kernel can be expressed as follows:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n} \xi_i \tag{4}$$

subject to:

$$y_i(\mathbf{w}^\top \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \tag{5}$$

where $\phi(\cdot)$ indicates the feature mapping, C denotes the penalty parameter, and ξ_i are slack variables that permit certain misclassifications. The radial basis function (RBF) kernel helps in understanding non-linear relationships between rainfall features and event severity.

Models were implemented using the `SVC` class from the `scikit-learn` library, incorporating cross-validation for hyperparameter tuning.

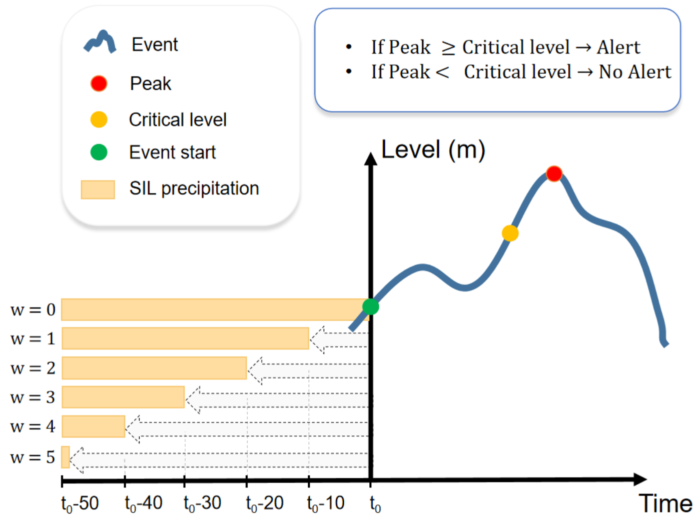

Fig. 8: Illustration of an event classification setup, including the event start threshold, the critical peak level h_{crit}, and the rainfall accumulation window used to construct features for the classifier.

C. Evaluation Metrics

The evaluation of model performance involves four classification metrics:

- **Precision**: the proportion of predicted positives that are true positives.
- **Recall**: the proportion of actual positives that are correctly predicted.
- **F1-Score**: the harmonic mean of precision and recall, providing a balanced performance indicator.
- **False Negatives**: the number of missed critical events, relevant for early warning systems.

D. Experimental Results

The SVM classifier was tested using feature sets generated from rainfall accumulation windows of increasing length. Results for the 1.0 m threshold are summarized in Table III. Each row corresponds to a different configuration of the sliding window index w, which controls the lead time and information density.

TABLE III: SVM classification performance at threshold $h_{\text{crit}} = 1.0\,\text{m}$.

Window (w)	F1-Score	Precision	Recall	False Negatives
0	0.8235	0.8400	0.8077	5
1	0.8235	0.8400	0.8077	5
2	0.7755	0.8261	0.7308	7
3	0.6000	0.8571	0.4615	14
4	0.6364	0.5250	0.8077	5
5	0.6567	0.5366	0.8462	4

Table III shows that shorter to medium-sized windows (e.g., $w = 1$ and $w = 2$) provide a more balanced approach between precision and recall. These windows probably reflect rainfall patterns that are closely linked to sudden increases in stream levels.

To complement the summary metrics, we analyze the confusion matrix for the SVM model applied with a critical threshold of 1.0 m and a sliding window of $w = 1$. This configuration corresponds to one of the more balanced results in terms of classification performance.

TABLE IV: Observed and predicted outcomes for the classification model ($h_{\text{crit}} = 1.0\,\text{m}$, $w = 1$).

		Predicted	
		No Event	**Event**
Actual	**No Event**	28	4
	Event	5	21

The model identified 21 out of 26 critical events, while 5 were not detected, as indicated in Table IV. The count of false negatives is significant, highlighting the difficulty in detecting events that have less noticeable rainfall patterns or shorter lead times. Conversely, there were only 4 false positives, indicating that the classifier tends to minimize misclassification in low-risk scenarios. A possible approach is to adjust the decision threshold or investigate ensemble methods that take into account temporal persistence and rainfall dynamics.

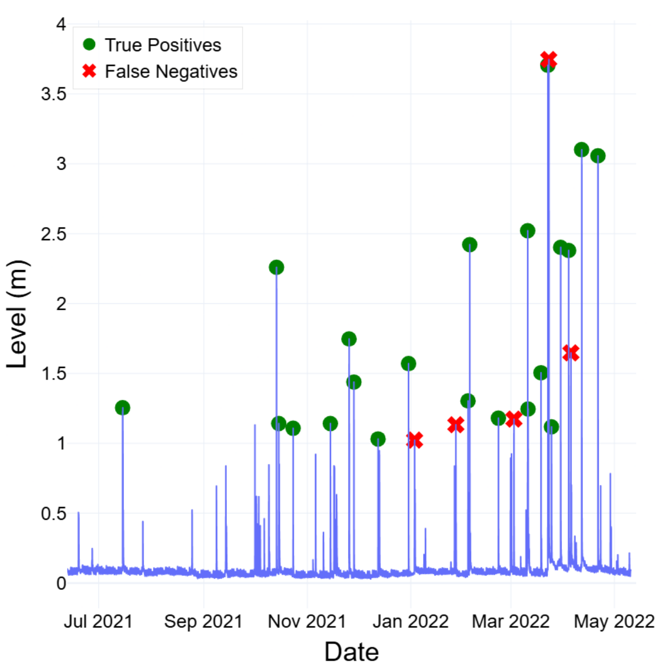

Fig. 9: True/False predict

The increase in false negatives for larger w values suggests that using older rainfall data could introduce noise rather than improve signal detection. This behavior holds significance for real-time applications, as missed alerts can affect the system's reliability in high-risk situations.

The SVM classifier is easy to implement and demonstrates consistent performance when trained on certain rainfall intervals. The performance can be modified to emphasize recall or precision according to the operational requirements of the flood alert system. Future work may explore hybrid classifiers or ensemble methods to improve sensitivity to rare but important events.

VI. Conclusion

This study presented a data-driven approach for predicting peak water levels and identifying critical streamflow events in the Mburicao urban basin. Using rainfall observations from three meteorological stations, two predictive models were developed: a linear regression model to estimate the magnitude of peak levels, and a support vector machine classifier to determine whether an event exceeds predefined thresholds.

The linear regression model achieved a peak performance of $R^2 = 0.8470$, with a root mean squared error (RMSE) of approximately 0.32 meters and a mean absolute percentage error (MAPE) near 28% when using rainfall windows between 10 and 30 minutes before peak events. These results suggest that even simple models can provide meaningful short-term estimates of stream response in urban basins.

The SVM classifier, designed to identify events that exceed a critical level of 1.0 meters, reached an F1-score of 0.82, with a precision of 0.84 and recall of 0.81. These metrics reflect a balanced detection capacity, capturing most relevant events while maintaining a manageable false positive rate.

Feature selection played a key role in refining model inputs. Strategies such as mutual information analysis and exhaustive subset evaluation helped identify the most informative rainfall windows, reducing input dimensionality without compromising performance.

A critical limitation identified during the study is the lack of an operational real-time monitoring system in the basin. Although radar sensors and pluviometers had been previously installed, they are currently offline. Work is ongoing to restore these systems. In parallel, we are collaborating with Castier et al. [1] to improve level–flow characterization and deploy new measurements at critical points throughout the basin.

The models developed here offer a practical baseline for early warning systems in urban settings with rapid hydrological response. Integrating these models into a real-time monitoring platform would support more timely alerts for local authorities and communities. The combination of timely data, lightweight predictive models, and distributed sensors could enhance flood resilience in urban environments.

References

[1] A. Thiessen-Anttila, M. Castier, and P. Barreto, "Simulating the effect of rainwater harvesting on flood mitigation: the case of asunción, paraguay," *Natural Hazards*, vol. 121, pp. 4335–4358, 10 2024.

[2] G. K. Devia, B. Ganasri, and G. Dwarakish, "A review on hydrological models," *Aquatic Procedia*, vol. 4, pp. 1001–1007, 2015, iNTERNATIONAL CONFERENCE ON WATER RESOURCES, COASTAL AND OCEAN ENGINEERING (ICWRCOE'15). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214241X15001273

[3] J. Hou, N. Zhou, G. Chen, M. Huang, and G. Bai, "Rapid forecasting of urban flood inundation using multiple machine learning models," *Natural Hazards*, vol. 108, no. 2, pp. 2335–2356, Sep 2021. [Online]. Available: https://doi.org/10.1007/s11069-021-04782-x

[4] W. J. Wee, N. B. Zaini, A. N. Ahmed, and A. El-Shafie, "A review of models for water level forecasting based on machine learning," *Earth Science Informatics*, vol. 14, no. 4, pp. 1707–1728, Dec 2021. [Online]. Available: https://doi.org/10.1007/s12145-021-00664-9

[5] A. N. Ahmed, A. Yafouz, A. H. Birima, O. Kisi, Y. F. Huang, M. Sherif, A. Sefelnasr, and A. El-Shafie, "Water level prediction using various machine learning algorithms: a case study of durian tunggal river, malaysia," *Engineering Applications of Computational Fluid Mechanics*, vol. 16, no. 1, pp. 422–440, 2022. [Online]. Available: https://doi.org/10.1080/19942060.2021.2019128

[6] S. S. Rudolf Scitovski, Siniša Maričić, "Short-term and long-term water level prediction at one river measurement location," *Croatian Operational Research Review*, vol. 3, 2012. [Online]. Available: https://hrcak.srce.hr/file/142442

[7] M. Ehteram, A. Ferdowsi, M. Faramarzpour, A. M. S. Al-Janabi, N. Al-Ansari, N. D. Bokde, and Z. M. Yaseen, "Hybridization of artificial intelligence models with nature inspired optimization algorithms for lake water level prediction and uncertainty analysis," *Alexandria Engineering Journal*, vol. 60, no. 2, pp. 2193–2208, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1110016820306840

[8] V.-H. Nhu, H. Shahabi, E. Nohani, A. Shirzadi, N. Al-Ansari, S. Bahrami, S. Miraki, M. Geertsema, and H. Nguyen, "Daily water level prediction of zrebar lake (iran): A comparison between m5p, random forest, random tree and reduced error pruning trees algorithms," *ISPRS International Journal of Geo-Information*, vol. 9, no. 8, 2020. [Online]. Available: https://www.mdpi.com/2220-9964/9/8/479

[9] J. A. J. P. Soares, L. C. S. M. Ozelim, L. Bacelar, D. B. Ribeiro, S. Stephany, and L. B. L. Santos, "Ml4ff: A machine-learning framework for flash flood forecasting applied to a brazilian watershed," *Journal of Hydrology*, vol. 652, p. 132674, 2025. [Online]. Available: https://www.mindat.org/reference.php?id=17914228

[10] A. Faruq, S. S. Abdullah, A. Marto, M. A. A. Bakar, Samin, and A. Mubin, "River water level forecasting for flood warning system using deep learning long short-term memory network," *IOP Conference Series: Materials Science and Engineering*, vol. 821, no. 1, p. 012026, apr 2020. [Online]. Available: https://dx.doi.org/10.1088/1757-899X/821/1/012026

[11] N. Imam Suprayogi1, Joleha2, "Applying adaptive neuro fuzzy inference system approach to river level forecasting," 2011. [Online]. Available: https://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/view/2964

[12] J. Duque, L. B. d. L. Santos, R. Santos, R. Oyarzabal, and J. Arteaga, "Nonlinear hydrological time series modeling to forecast river level dynamics in the rio negro uruguay basin," *Chaos*, vol. 34, no. 5, p. 053132, 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/38780437/