PAPER 2

COST-EFFECTIVE FLOOD FORECASTING FOR MUNICIPAL DISASTER MANAGEMENT SYSTEMS

Shawn Johnson
Obscape B.V.

ABSTRACT

This paper presents a cost-effective approach to flood forecasting using simple linear regression trained on single-point rainfall and water level observations. Research conducted on the Palmiet River in Durban demonstrated that a model using only three input features (the previous hour's water level, current rainfall, and previous rainfall) could explain over 85% of the variance in observed water levels when trained on as little as three months of data including a flood event. The critical finding was that including significant rainfall events in the training data dramatically improved model accuracy, with R^2 values increasing from 0.63–0.74 to 0.85 and mean absolute errors (MAE) decreasing from 0.067–0.069 m to 0.048–0.056 m.

Building on these research findings, a pilot phase was initiated to test the approach under real-time operational conditions, where forecast rainfall must be substituted for observed measurements. A multi-site platform (wxi.info) now delivers hourly water level forecasts for the Mzinyathi and Palmiet rivers in KwaZulu-Natal, using rainfall forecasts from eleven global numerical weather prediction models accessed freely via the Open-Meteo API. An automated grid search selects the optimal lag combination per site, adapting to each catchment's hydrological response without manual calibration.

Initial results from over three months of pilot operation confirm that the regression approach delivers operationally useful water level forecasts, with MAE remaining within two centimetres over the first six hours and under seven centimetres even at extended horizons beyond 72 hours. As anticipated, forecast accuracy at longer horizons is governed primarily by the quality of rainfall forecast inputs, which was the central variable the pilot phase was designed to investigate. The multi-model ensemble provides robustness and a practical measure of forecast uncertainty. This work demonstrates that real-time flood forecasting is achievable affordably using single-point data where traditional hydrological modelling is not feasible.

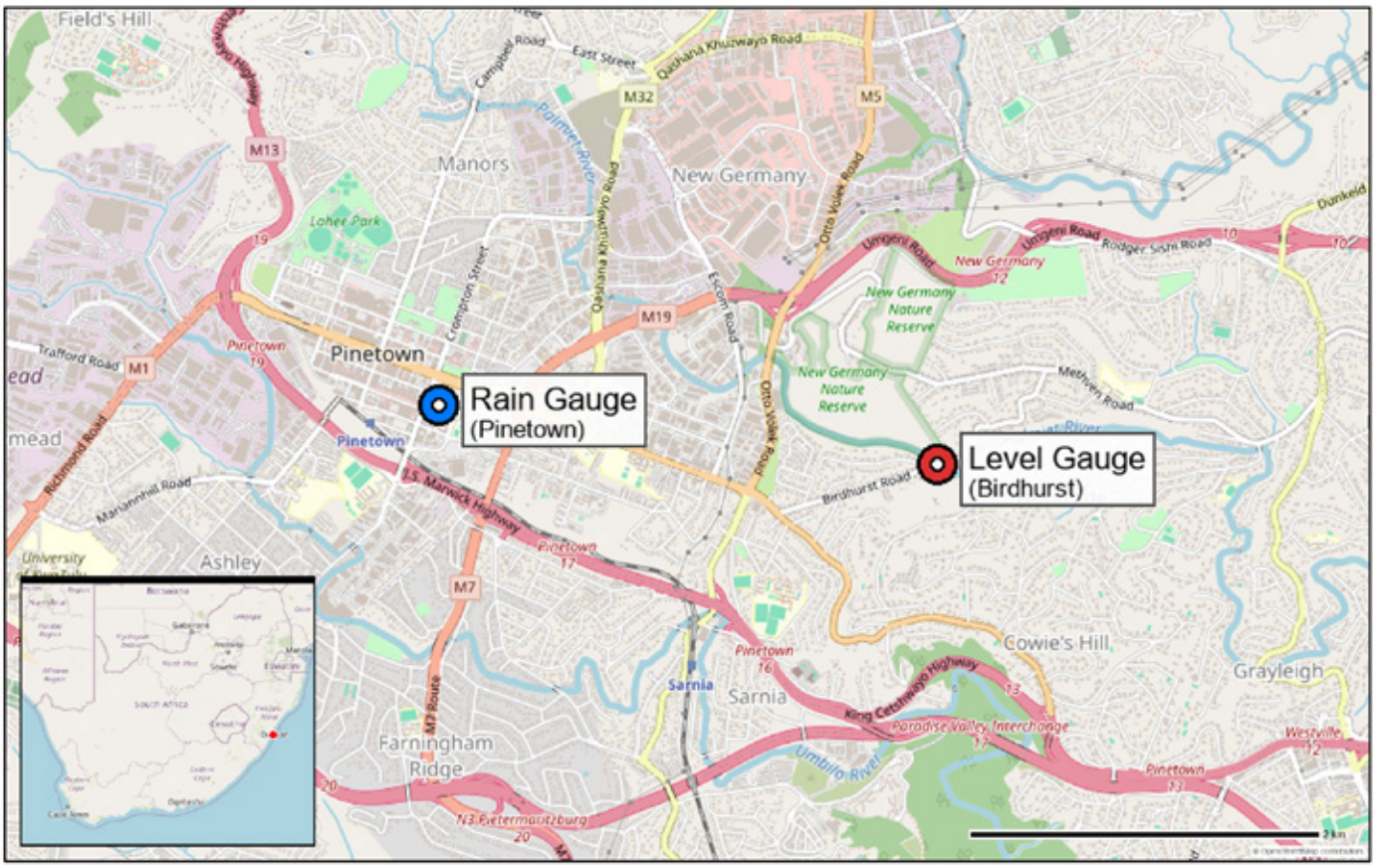


FIGURE 1: *Positioning of the water level and rain gauge on the Palmiet River catchment*

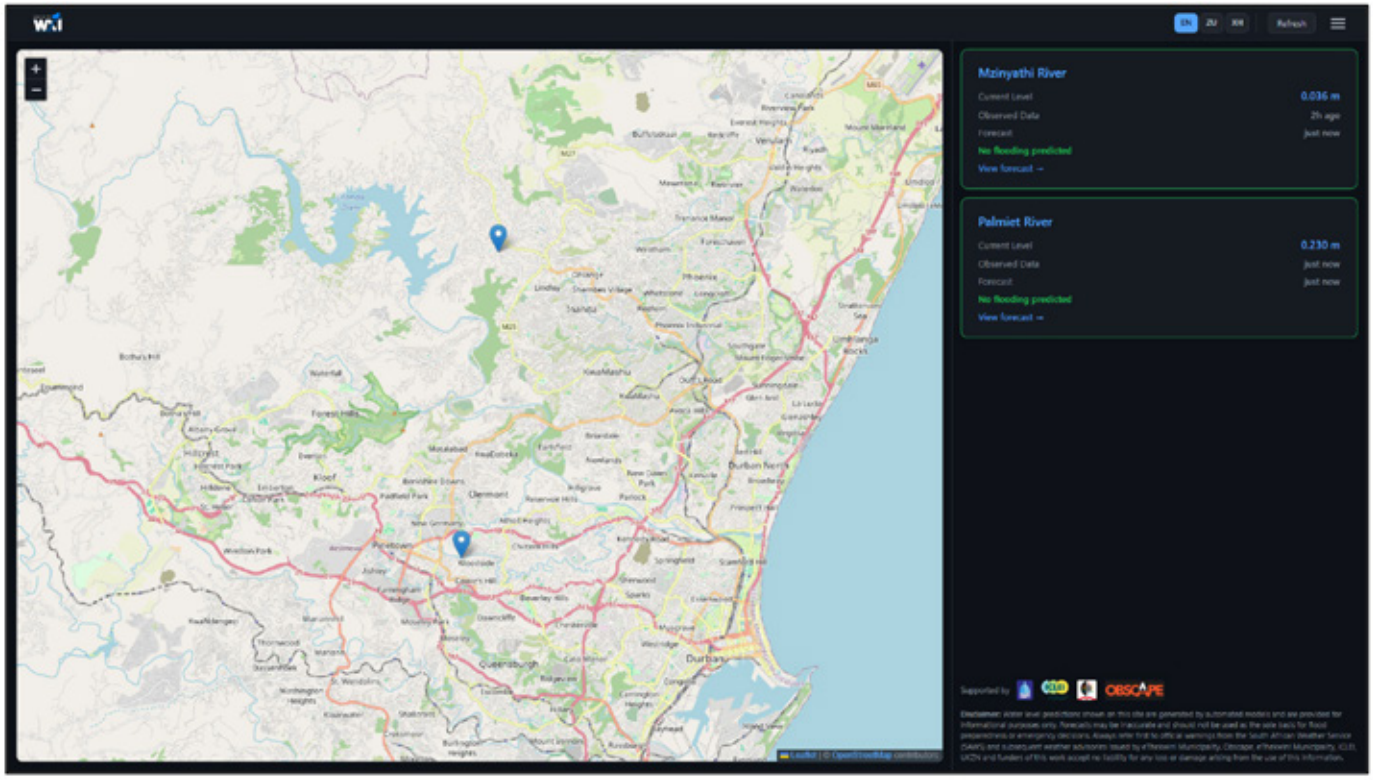


FIGURE 2: *Location of active monitoring sites within the eThekwini Metropolitan Municipality*

INTRODUCTION

Flooding is one of the most devastating natural hazards affecting South African municipalities. The KwaZulu-Natal floods of April 2022 resulted in over 400 fatalities and billions of rands in infrastructure damage, exposing critical gaps in municipal early warning capabilities (Dlamini *et al.* 2023). The eThekwini Metropolitan Municipality, encompassing Durban and surrounding areas, is particularly vulnerable due to its steep topography, rapid urbanisation, and subtropical climate producing intense convective rainfall events.

Effective flood early warning systems are widely recognised as among the most cost-effective measures for reducing flood losses (UNDRR 2022). However, the implementation of such systems in South African municipalities faces significant barriers. Traditional hydrological models, such as HEC-HMS, MIKE SHE, or SWAT, require detailed catchment data, expert calibration, and ongoing maintenance that many local authorities cannot sustain (Smithers & Schulze 2002). The shortage of qualified hydrologists in municipal service further limits adoption. Even where models exist, they are often research tools that remain disconnected from real-time operational decision-making.

Recent advances in low-cost environmental monitoring and open-access weather forecast data create an opportunity to approach flood forecasting differently. Internet-of-Things (IoT) sensor networks now provide continuous, high-frequency water level and rainfall observations at a fraction of the cost of traditional gauging infrastructure (Mao *et al.* 2019). Meanwhile, services such as Open-Meteo aggregate outputs from major global numerical weather prediction (NWP) centres, including the European Centre for Medium-Range Weather Forecasts (ECMWF), the United States National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS), and the United Kingdom Met Office (UKMO), and make them freely available via standardised application programming interfaces (APIs) (Open-Meteo 2025), eliminating a key data barrier for municipal applications.

This paper presents research into the use of simple linear regression for water level forecasting in small urban catchments, using only single-point observations from one water level sensor and one rain gauge. The research investigates model accuracy across different training strategies and demonstrates the critical importance of including flood events in training data. It then describes the deployment of these findings into a live pilot system, publicly accessible at wxi.info, that tests the approach under real-time operational conditions using forecast rainfall from multiple NWP models. The objectives are to: (1) present the research methodology and findings; (2) describe the transition to operational pilot deployment; and (3) report initial operational results and their implications for municipal flood early warning.

STUDY AREA AND DATA SOURCES

Study area

The research was conducted on the Palmiet River in Durban, a fast-responding catchment serving a densely populated area including several informal settlements at elevated flood risk. Water level data was recorded at the Birdhurst site, with rainfall measured at a single upstream gauge in Pinetown (Figure 1). The proximity of these stations minimises lag between rainfall and river response, reinforcing the suitability of simplified modelling approaches.

For the pilot phase, the system was expanded to two river sites within the eThekwini Metropolitan Municipality (Figure 2). The **Mzinyathi River**

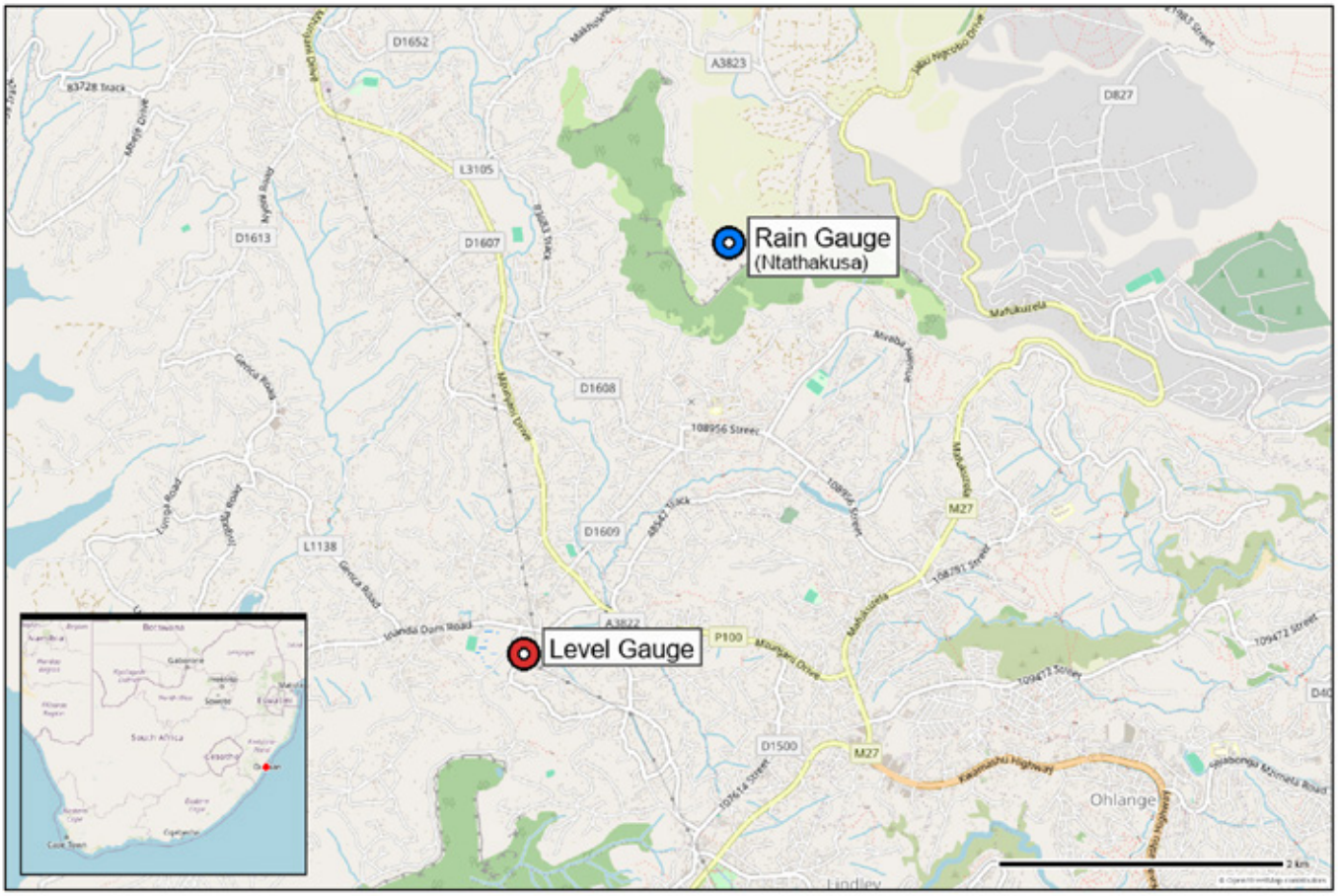


FIGURE 3: *Positioning of the water level and rain gauge on the Mzinyathi River catchment*

is located in the northern eThekwini region and is characterised by a relatively responsive catchment with rapid runoff during rainfall events (Figure 3). The **Palmiet River** site continues from the research phase. Both sites are instrumented with single-point IoT monitoring stations, with one water level sensor and one rain gauge per site, recording at high temporal resolution.

Observed data

Water level and rainfall observations are sourced from single-point IoT monitoring stations operated within the eThekwini municipal monitoring network. Each site is equipped with one water level sensor (typically pressure transducer) and one tipping-bucket rain gauge, both logging at sub-hourly intervals. No spatially distributed sensor networks or catchment-averaged data are required.

Raw observations are resampled to hourly resolution. Water level values are aggregated as the **maximum** reading per hour, capturing peak conditions relevant to flood risk. Rainfall values are aggregated as the **sum** per hour, preserving total precipitation volumes. Invalid readings and sensor sentinel values are filtered during resampling.

Rainfall data

For the research phase, observed (measured) rainfall was used as input, allowing model accuracy to be assessed independently of rainfall forecast quality.

For the pilot phase, forecast rainfall inputs are obtained from the Open-Meteo API (Open-Meteo 2025), a free, open-access service that aggregates outputs from major global NWP centres. The system ingests hourly precipitation forecasts from eleven models, listed in Table 1. Each model provides forecasts for approximately seven days ahead. Forecasts are used as rainfall inputs to the water level prediction model, which is run independently for each weather model. An ensemble mean prediction is then computed as the simple average across all eleven model outputs. No API key is required and no cost is incurred.

TABLE 1: *Numerical weather prediction models used for rainfall forecasting (pilot phase)*

Model	Source centre
Best Match	Open-Meteo (auto-selected best)
CMA GRAPES Global	China Meteorological Administration
ECMWF AIFS	ECMWF (AI-enhanced)
ECMWF IFS 0.25°	ECMWF IFS (0.25° resolution)
ECMWF IFS	ECMWF IFS
GFS Global	NCEP GFS
GFS Seamless	NCEP GFS Seamless
MET Norway Seamless	Norwegian Meteorological Institute
NCEP AI-GFS	NCEP AI-GFS
NCEP HGEFS Ensemble Mean	NCEP HGEFS ensemble mean
UKMO Global	UK Met Office

RESEARCH METHODOLOGY AND RESULTS

Model design

The forecasting approach uses ordinary least squares linear regression with lagged input features. The predicted water level at time *t* is given by:

$$\hat{H}(t) = \beta_0 + \Sigma \beta_i \cdot x_i \ldots (1)$$

where β_0 is the intercept, β_i are the learned coefficients, and x_i are the input features. For the research phase, three input features were used: the previous hour's water level (H_lag1), the current hour's rainfall (R_t), and the previous hour's rainfall (R_lag1). This minimal feature set was selected based on the physical reasoning that in small, steep catchments, the hydrological response to rainfall is near-immediate, and the water level at any given hour is primarily determined by the recent water level (representing baseline conditions and drainage) and the recent rainfall (representing inflow). Predicted values are clamped to a minimum of zero.

The model produces multi-step forecasts through an iterative (recursive) procedure. Starting from the most recent observed water level and rainfall, the model predicts one hour ahead. This predicted value then becomes an input for the next prediction step. The process repeats for each hour of the forecast horizon, with each step's output becoming the next step's input. This simulates real-world operational conditions where only the initial water level and subsequent rainfall measurements are available.

Training strategy investigation

To assess the effect of training data composition on model performance, eight models were trained across four training durations (3, 6, 9, and 12 months) and two strategies: one excluding the April 2022 KwaZulu-Natal flood event from the training data, and one including it. All models were tested on an unseen four-week period (February–March 2025) using the iterative prediction approach described above.

Research results

Table 2 presents the research results across all training configurations. The effect of including the April 2022 flood event in the training data was substantial and consistent.

Models trained with the April 2022 flood achieved R^2 values of 0.845–0.854 and mean absolute errors (MAE) of 0.048–0.056 m, compared to R^2 values of 0.634–0.736 and MAE of 0.067–0.069 m for models trained without the flood. The best-performing configuration, a three-month training window including the flood event, achieved an R^2 of 0.854 and an MAE of 0.048 m.

TABLE 2: *Research results: model performance across training configurations*

Training strategy	Period	R^2	MAE (m)	Bias (m)
Excluding April 2022 flood	3 months	0.634	0.069	–0.015
	6 months	0.694	0.066	–0.021
	9 months	0.689	0.070	–0.023
	12 months	0.736	0.067	–0.026
Including April 2022 flood	3 months	0.854	0.048	–0.030
	6 months	0.848	0.050	–0.033
	9 months	0.845	0.054	–0.039
	12 months	0.845	0.056	–0.042

For models trained without flood data, performance improved steadily with longer training durations (R^2 from 0.634 at 3 months to 0.736 at 12 months), as longer windows provided greater exposure to the variability of normal flow conditions. For models trained with flood data, performance was strong across all durations, with the shorter windows slightly outperforming, likely because the flood event represented a larger proportion of the training data.

Figure 4 illustrates the model performance during a representative rain event. Models trained with flood data closely track the observed water level rise and recession, while all four training durations produce similar predictions, confirming the robustness of the approach. Performance was also assessed during active rainfall periods only, since these are the conditions most relevant to flood warning. Flood-trained models maintained R^2 values above 0.80 even when evaluated exclusively on rainy hours.

Bias analysis

All models exhibited a consistent slight negative bias, indicating minor systematic underprediction. Models trained with flood data showed bias values of –0.030 to –0.042 m, while those without showed –0.015 to –0.026 m. During rainfall periods specifically, models trained without the flood exhibited positive bias (overprediction), while flood-trained models

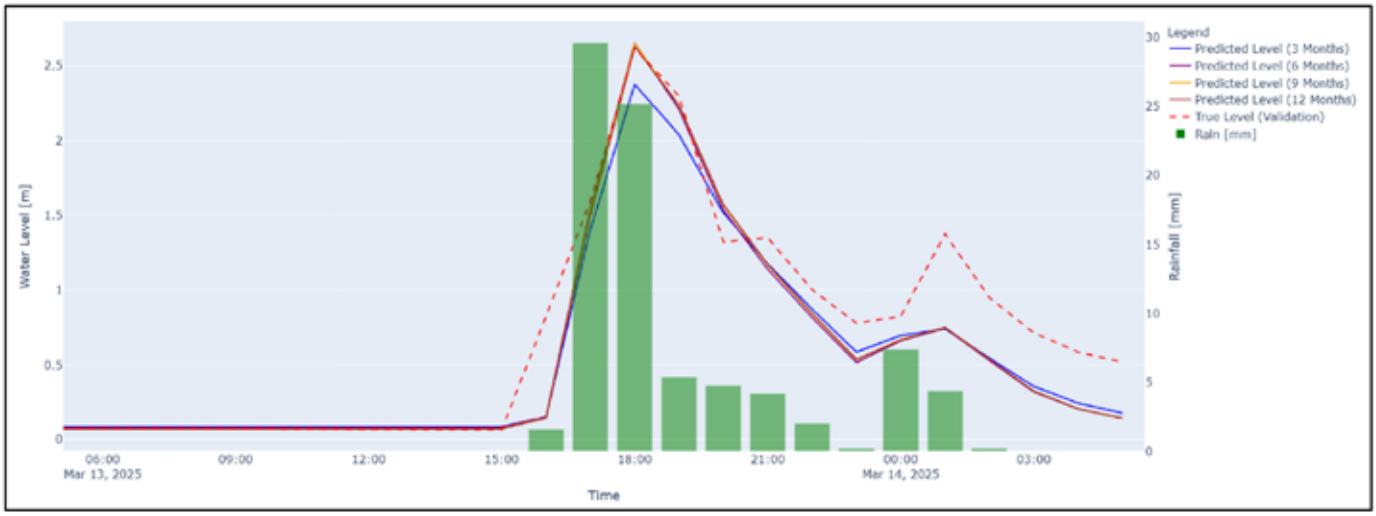


FIGURE 4: *Close-up of modelled versus observed water levels during a rain event (models trained with April 2022 flood data)*

maintained a mild negative bias, suggesting more stable and consistent performance under wet conditions.

Key findings

The research established three principal findings:

1. Simple linear regression using only three single-point input features can explain over 85% of the variance in water levels for small urban catchments, achieving MAE below 0.05 m over a four-week iterative test period.
2. Including significant flood events in the training data is critical. Without such events, models cannot adequately learn the rainfall-runoff response at higher water levels, resulting in substantially lower accuracy.
3. As little as three months of training data is sufficient to produce an accurate model, provided a flood event is captured, making the approach deployable within a single wet season of instrument installation.

The critical open question remaining was whether these results, achieved with observed (measured) rainfall, would hold under operational conditions where forecast rainfall must be substituted.

PILOT PHASE: OPERATIONAL DEPLOYMENT

System design

To address the open question from the research phase, the model was deployed into a live operational system. Key adaptations for operational use included:

Automated feature selection: The research used a fixed three-feature model. The operational system replaces this with an exhaustive grid search over all non-empty subsets of an expanded candidate feature space (water level lags H_lag1 through H_lag3 and rainfall terms R_t through R_lag4). Each combination is evaluated using five-fold time series cross-validation scored by MAE, and the optimal set is selected independently per site.

The resulting feature sets are:

- **Mzinyathi River**: H_lag1, H_lag2, R_lag1 (3 features, cross-validated MAE = 0.008 m)
- **Palmiet River**: H_lag1, H_lag3, R_t, R_lag3, R_lag4 (5 features, cross-validated MAE = 0.012 m)

Multi-model ensemble: Rather than relying on a single rainfall forecast, the system runs the prediction model independently with each of eleven NWP models' rainfall forecasts, producing twelve forecast traces (eleven models plus an ensemble mean). This provides both a range of possible outcomes and a more robust central estimate.

Automated pipeline: An hourly scheduled task fetches observations, retrieves rainfall forecasts, generates predictions for each active site, and evaluates past forecast accuracy. If one site encounters a data source issue, the remaining sites continue unaffected.

Sensor outage gap-filling: The system maintains operational continuity through automated gap-filling: using the model to estimate missing water levels from observed rainfall, rearranging the equation to estimate missing rainfall from observed water levels, or falling back to NWP hindcast data when both sensors are offline. All estimated values are flagged and automatically replaced when sensors recover.

Public interface (Figure 5): The platform provides a community-facing view with a configurable flood threshold and warning banner. The flood threshold is user-adjustable, allowing community members to set it based on their local knowledge of river levels that begin to create challenges, such as inundation of low-lying areas or overtopping of bridges and crossings. These thresholds can be refined over time through observation and experience of flood events. When water level predictions from any model exceed the threshold, a warning banner is displayed. A technical dashboard provides per-model forecasts, ensemble averages, and accuracy statistics by forecast horizon. The platform is available in English, isiZulu, and isiXhosa, and is deployed as a progressive web application (PWA) that can be installed directly on mobile devices, improving accessibility for community users who may rely on smartphones as their primary means of accessing information.

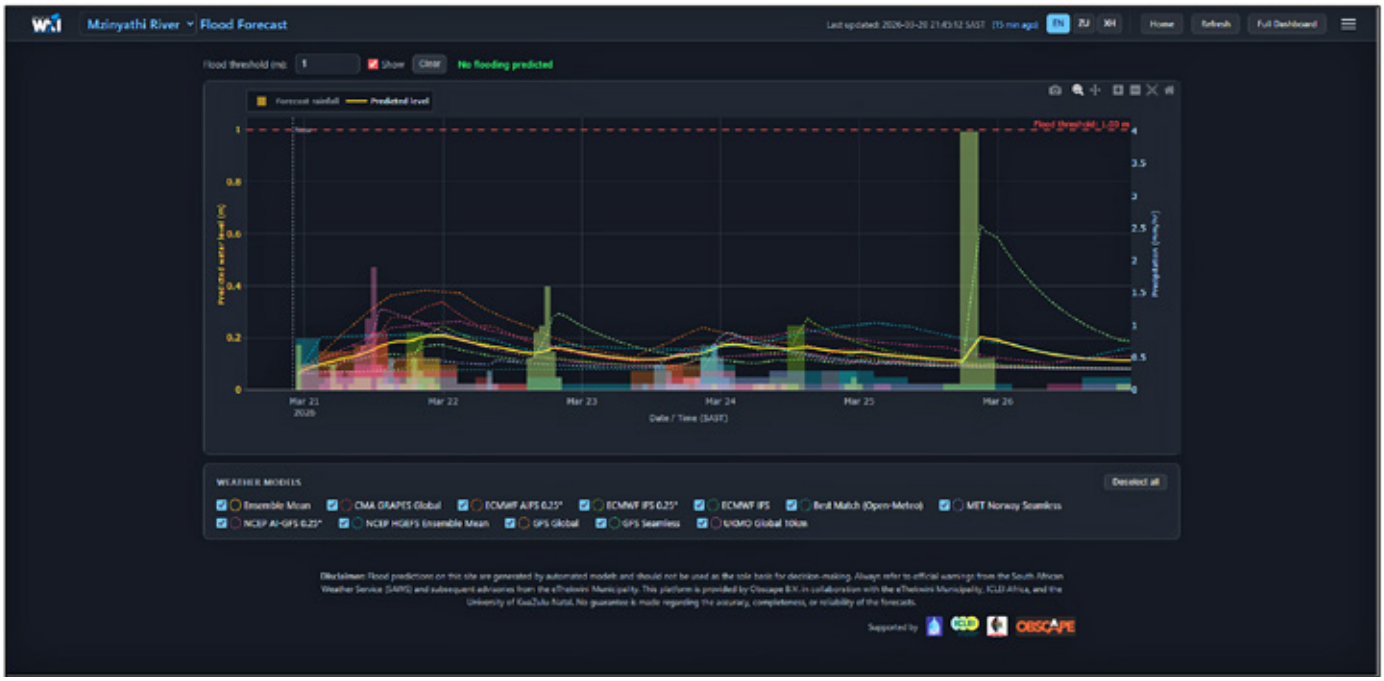


FIGURE 5: *Community-facing forecast view showing ensemble water level prediction, individual model traces, forecast rainfall, and configurable flood threshold*

TABLE 3: *Pilot phase water level prediction accuracy by weather model (all horizons combined)*

Weather model	MAE (m)	Bias (m)
Best Match	0.050	+0.041
GFS Global	0.055	+0.046
GFS Seamless	0.055	+0.046
ECMWF IFS	0.057	+0.049
MET Norway	0.057	+0.049
CMA GRAPES	0.058	+0.049
ECMWF IFS 0.25°	0.058	+0.049
Ensemble Mean	**0.059**	**+0.052**
NCEP AI-GFS	0.062	+0.055
NCEP HGEFS	0.063	+0.056
UKMO	0.068	+0.059
ECMWF AIFS	0.076	+0.069

Initial operational results

The results presented here are drawn from over three months of pilot operation, from March to June 2026. During this period, the system generated 3 930 hourly forecast runs across both sites (1 654 for the Mzinyathi River and 2 276 for the Palmiet River), producing over 2.8 million water level prediction-observation comparison pairs.

Table 3 summarises the water level prediction accuracy across all forecast horizons and both sites.

The overall MAE ranges from 0.050 m to 0.076 m, with the ensemble mean at 0.059 m. Table 4 presents the ensemble mean MAE by forecast horizon.

Short-range forecasts (1–6 h) achieve a MAE of less than 2 cm. Even at extended horizons beyond 72 hours, the MAE remains below 7 cm, a level of accuracy that is operationally meaningful for flood preparedness. While the pilot period included moderate rainfall events that tested the model's response characteristics, no extreme flood events comparable to the April 2022 disaster occurred. The evaluation therefore primarily reflects performance under normal to moderate conditions, and further evaluation during significant high-flow events is required to fully characterise operational accuracy.

The operational results show a positive bias (over-prediction), contrasting with the negative bias observed in the research phase. This shift is attributable to the use of forecast rather than observed rainfall. The bias varies by site: the Mzinyathi River shows a consistent positive bias (+0.057 m), while the Palmiet River exhibits a slight negative bias (–0.034 m), indicating that bias characteristics are site-specific.

Dependence on rainfall forecast quality

The research demonstrated model accuracy using observed rainfall. A central objective of the pilot phase was to evaluate whether that accuracy would be maintained when forecast rainfall is substituted.

TABLE 4: *Pilot phase ensemble mean water level MAE by forecast horizon*

Horizon band	MAE (m)	Bias (m)
1–6 h	0.019	+0.010
7–12 h	0.033	+0.023
13–24 h	0.044	+0.034
25–48 h	0.053	+0.045
49–72 h	0.060	+0.053
73 h+	0.067	+0.059

The results confirm that rainfall forecast quality is the dominant factor governing prediction accuracy at longer horizons. During dry periods, the autoregressive lag terms dominate the prediction, yielding low errors regardless of rainfall forecast quality. During rainfall events, discrepancies between predicted and actual rainfall propagate through the iterative prediction chain. This was anticipated, and it is the fundamental reason the system was designed around a multi-model ensemble rather than a single rainfall source.

DISCUSSION

The research phase established that simple linear regression, using only three single-point input features, can produce accurate water level forecasts for small urban catchments when trained on observed data that includes significant flood events. The pilot phase demonstrates that these findings translate to operational conditions, with forecast accuracy that is operationally useful for flood early warning, particularly over the critical short-range horizons (1–24 hours) where municipal action is most effective.

The multi-model ensemble approach provides practical advantages beyond raw accuracy. No single model consistently outperforms across all conditions, and model rankings shift between dry and wet periods and between sites. The ensemble averages out individual model biases, providing more stable performance over time. Divergence between model traces also provides an intuitive visual measure of forecast confidence.

The multi-site architecture has been validated with two operational sites. Adding the Palmiet River site required only model training and a historical data backfill, with no changes to the application or forecast pipeline. This configuration-only scalability is significant for municipal adoption.

Three key limitations should be acknowledged:

1. **Training data requirements**: Model accuracy is highly dependent on the variance present in the training data, specifically the inclusion of significant rainfall events. For new sites, sufficient data capturing the range of hydrological conditions must be accumulated before the model can be considered reliable.
2. **Spatial resolution of rainfall forecasts**: Global NWP models operate at grid resolutions of 10–25 km, while convective rainfall in KwaZulu-Natal is highly localised. This spatial mismatch is a fundamental source of error for catchment-scale applications. The forecasting approach is likely most valid for small catchments where the contributing area falls within a single NWP grid cell, such that the grid-scale rainfall estimate reasonably approximates the catchment-average input. For larger or more spatially complex catchments, distributed rainfall inputs or downscaling techniques may be required to maintain forecast accuracy.
3. **Early-stage pilot testing**: The three-month pilot period represents the initial phase of operational evaluation. No events comparable in magnitude to the April 2022 floods have occurred during this period. The operational results should be considered preliminary, and continued evaluation across a wider range of conditions is required to fully characterise system performance (Hauswirth *et al.* 2023).

CONCLUSIONS

This paper has presented research into cost-effective flood forecasting using simple linear regression and the initial results from an operational pilot deployment.

The key conclusions are:

1. Simple linear regression using only three single-point input features (previous water level, current rainfall, and previous rainfall) can explain over 85% of the variance in water levels for small urban catchments,

achieving MAE below 0.05 m over a four-week iterative test period using observed rainfall.

2. Including significant flood events in the training data is critical to model accuracy. Models trained without the April 2022 flood achieved R^2 values of only 0.63–0.74, compared to 0.85 when the flood was included. As little as three months of data is sufficient if a flood event is captured.
3. Initial results from the pilot phase confirm that the approach delivers operationally useful forecasts under real-time conditions using NWP rainfall inputs. MAE remains below 2 cm for the first six hours and below 7 cm at extended horizons beyond 72 hours.
4. As anticipated, the quality of rainfall forecast inputs is the dominant factor governing prediction accuracy at longer horizons, confirming that the transition from observed to forecast rainfall is the critical challenge in operationalising this approach.
5. The system can be scaled to additional monitoring sites through configuration and model training alone, requiring only single-point instrumentation (one water level sensor and one rain gauge per site), making it accessible to resource-constrained municipalities. This work demonstrates that real-time flood forecasting is affordably achievable using single-point data with a minimum of three months' worth of data during the rain season. Over time and with further testing and upscaling, this approach could also support community-level monitoring, extending its value beyond municipal forecasting functions. This will enable municipalities to meet their mandates with regard to the Disaster Management Act and to improve the resilience of their citizens.

RECOMMENDATIONS

Based on the research and pilot phase experience, the following recommendations are made for municipalities considering similar systems:

1. **Prioritise data collection that includes flood events**: Model accuracy depends critically on capturing significant rainfall events in the training data. Municipalities should begin data collection as early as possible, ideally before the wet season, and retain all data from high-flow events.
2. **Invest in observation quality**: Well-maintained single-point monitoring stations with redundant communication pathways are the foundation of any operational forecasting system. Only one water level sensor and one rain gauge per site are required.
3. **Adopt multi-model ensembles**: No single weather model consistently outperforms across all conditions. Municipalities should leverage the free availability of multiple NWP outputs rather than relying on any single rainfall forecast source.
4. **Communicate uncertainty transparently**: Public-facing flood forecasting systems should present accuracy metrics alongside predictions, allowing users to calibrate their trust in forecasts at different lead times.
5. **Plan for continued validation**: Systems should be evaluated across a wide range of conditions, particularly significant flood events, before being relied upon for critical decision-making.

ACKNOWLEDGEMENTS

This research was commissioned by ICLEI Africa as part of a collaborative project with the eThekwini Municipality and UKZN, under the CLARE-funded 'INACCT Resilience' project. The eThekwini Municipality is gratefully acknowledged for providing access to their IoT monitoring network data, which made this work possible. Weather forecast data was provided free of charge by Open-Meteo (open-meteo.com).

REFERENCES

Dlamini T, Mohanty S & Worku Z 2023. The April 2022 KwaZulu-Natal floods: A hydrometeorological assessment and implications for disaster risk reduction. *South African Journal of Science* 119(5/6): 1–9.

Hauswirth SM, van der Wiel K, Bierkens MFP, Beijk V & Wanders N 2023. Simulating hydrological extremes for different warming levels–combining large scale climate ensembles with local observation based machine learning models. *Frontiers in Water* 5: 1108108.

Mao F, Khamis K, Krause S, Clark J & Hannah DM 2019. Low-cost environmental sensor networks: Recent advances and future directions. *Frontiers in Earth Science* 7: 221.

Open-Meteo 2025. Open-Meteo Free Weather API Documentation. Available at: https://open-meteo.com/en/docs (accessed 18 June 2026).

Smithers JC & Schulze RE 2002. The role of hydrology in water resources management in South Africa. In: Schulze RE (ed.). *Hydrology and Water Resources of KwaZulu-Natal*. WRC Report TT 154/01, Water Research Commission, Pretoria.

UNDRR 2022. *Global Assessment Report on Disaster Risk Reduction 2022: Our World at Risk — Transforming Governance for a Resilient Future*. United Nations Office for Disaster Risk Reduction, Geneva.