



Design Evaluation and Validation of a Resilient IoT-Based Flood Prediction Framework for Data-Scarce Environments in East Africa

Muwanga Erasto Kosea¹, Dr Otanga Daniel², Dr. Satwinder Singh Rupra³

School of Computing and Informatics, Masinde Muliro University of Science and Technology^{1,2,3}

Abstract: Many Internet of Things (IoT)-based flood prediction systems deployed in developing regions fail to deliver reliable early warnings due to unreliable sensors, fragmented datasets, and limited operational resilience. While numerous frameworks have been proposed, few studies systematically evaluate their design limitations or validate enhanced solutions under realistic failure conditions. This paper presents the design evaluation, enhancement, and validation of a resilient IoT-based flood prediction framework. Using Design Science Research principles, existing IoT flood prediction frameworks were evaluated using ITIL-aligned governance criteria to identify deficiencies in data reliability, service continuity, and system governance. An enhanced framework was then designed and validated through simulation using CHIRPS rainfall data and controlled sensor failure scenarios. Simulation results indicate that the enhanced framework maintains prediction accuracy between 82.4% and 91.6% under increasing data-loss conditions and improves alert timeliness compared to baseline approaches. The findings indicate that resilience-oriented, data-centric IoT design significantly improves flood prediction performance in resource-constrained environments.

Keywords: Flood prediction; Internet of Things; design science research; data reliability; sensor unreliability; early warning systems.

1.0 INTRODUCTION

Flooding remains a major environmental and socio-economic challenge in East Africa, particularly in urban and peri-urban areas characterized by rapid population growth and inadequate drainage infrastructure. Despite increasing investment in flood early warning systems, prediction accuracy and operational reliability remain limited, especially during extreme weather events when sensor infrastructure and communication networks are most vulnerable. IoT technologies have been widely adopted for flood monitoring due to their ability to support continuous sensing, real-time data transmission, and automated analytics. However, most existing IoT-based flood prediction frameworks implicitly assume stable infrastructure, reliable sensors, and complete datasets. In data-scarce and sensor-unreliable environments, these assumptions result in frequent system failure, false alarms, or delayed warnings.

This paper is structured around three interrelated subsections that address the core challenges of IoT-based flood prediction under unreliable operating conditions. First, the study undertakes a systematic design evaluation of existing IoT flood prediction frameworks to identify structural, data reliability, and governance-related limitations. Second, informed by the evaluation findings, the paper presents the design and enhancement of a resilience-oriented IoT-based flood prediction framework that explicitly addresses data scarcity and sensor unreliability. Third, the proposed framework is subjected to simulation-based validation using controlled data-loss and sensor-failure scenarios to assess its predictive accuracy, alert timeliness, and operational stability. These three subsections collectively provide a cohesive design–evaluation–validation narrative.

2.0 EVALUATION OF EXISTING IOT FLOOD PREDICTION FRAMEWORKS

Existing IoT flood prediction frameworks were evaluated using ITIL-aligned criteria to identify structural and operational weaknesses related to data handling, incident management, service continuity, and system adaptability. The use of ITIL was informed by its established role in structuring IT service management processes, particularly in environments characterized by operational uncertainty and resource constraints (Skyttner, 2005; Hevner, 2007). In line with Design Science Research principles, ITIL was employed strictly as a governance and diagnostic lens to guide systematic comparison across heterogeneous frameworks, rather than as a predictor of analytical or predictive accuracy (March & Smith, 1995; Gregor & Hevner, 2013). This approach enabled the evaluation to focus on how existing frameworks



A central contribution of the framework is its explicit treatment of data reliability as a mediating system property. In contrast to conventional architectures that assume reliable sensor inputs, the proposed framework recognizes sensor unreliability and dataset imperfections as persistent characteristics of deployment contexts in developing regions. By positioning data reliability between system inputs and flood prediction outcomes, the framework explains how disturbances at the sensing or communication layer propagate through analytical processes unless actively mitigated. At the input layer, the framework integrates IoT sensing and communication infrastructure with governance and design standards. Sensors, transmission channels, and computational nodes form the technical foundation of the system but remain vulnerable to power instability and environmental exposure. Governance standards such as ITIL are incorporated as design guidance mechanisms that shape data validation, incident management, and service continuity without being treated as direct determinants of predictive accuracy.

Resilience is operationalized within the processing layer through sensor redundancy, data cleansing, and hybrid analytics combining ARIMA forecasting, threshold-based classification, and supervised machine learning using a Random Forest classifier. This combination balances interpretability, responsiveness, and robustness under unreliable data conditions. The output layer translates stabilized analytical results into actionable flood prediction outcomes, including prediction accuracy, alert timeliness, and system stability, which collectively determine early warning effectiveness.

Overall, the framework advances IoT-based flood prediction by demonstrating how resilience can be systematically embedded through architectural mediation rather than assumed through ideal infrastructure conditions. By integrating governance guidance, data reliability mechanisms, and hybrid analytics within a coherent systems framework, the design provides a practical and theoretically grounded blueprint for resilient flood prediction in data-scarce environments.

4.0 SIMULATION-BASED VALIDATION

Table 1: Performance Metrics Used for Validating the Resilient IoT Flood Prediction Framework

Performance Metric	Operational Definition	Simulation Conditions Applied	Evaluation Purpose
Prediction Accuracy	Proportion of correctly classified flood and non-flood events	Data gaps ranging from 5% to 30%; random noise injection	Assess analytical effectiveness under data uncertainty
Alert Latency	Time delay between flood risk detection and alert generation	Variable network latency and packet-loss scenarios	Evaluate timeliness of early warning under network instability
System Stability	Ability of the system to maintain consistent operation without failure	Intermittent transmission delays; low-power sensor assumptions	Measure operational resilience under degraded conditions
Tolerance to Missing Data	Sustained predictive performance despite incomplete sensor inputs	Controlled data loss (5%–30%)	Assess robustness to sensor unreliability
Noise Robustness	Resistance to performance degradation due to noisy sensor readings	Random noise injection in rainfall data	Evaluate data validation and cleansing effectiveness
Graceful Degradation	Degree to which performance declines gradually rather than abruptly	Combined data loss and network instability	Assess resilience-oriented system behavior

Table 1 summarizes the performance metrics used to validate the proposed resilient IoT flood prediction framework. The metrics were selected to capture both analytical effectiveness and operational resilience under realistic deployment conditions characterized by sensor unreliability, intermittent connectivity, and data uncertainty. Consistent with Design Science evaluation principles, the framework was assessed based on its ability to sustain predictive accuracy, timely alert generation, and stable operation under progressively degraded conditions (March & Smith, 1995).

Validation of the proposed framework was conducted using a GitHub-based IoT simulation environment that emulated rainfall sensors, data ingestion pipelines, and analytical components. Sensor unreliability was modeled through controlled data gaps ranging from 5% to 30%, random noise injection, and intermittent transmission delays. Network instability was



simulated using variable latency and packet-loss parameters. The environment assumed low-power sensor nodes and intermittent connectivity, reflecting realistic deployment conditions in flood-prone regions of East Africa.

Framework performance was evaluated using prediction accuracy, alert latency, and system stability under degraded conditions. These metrics were selected to assess both analytical effectiveness and operational resilience, consistent with Design Science evaluation principles (March & Smith, 1995). Simulation results indicate that data reliability plays a dominant role in shaping flood prediction performance under the evaluated conditions.

5.0 RESULTS AND DISCUSSION

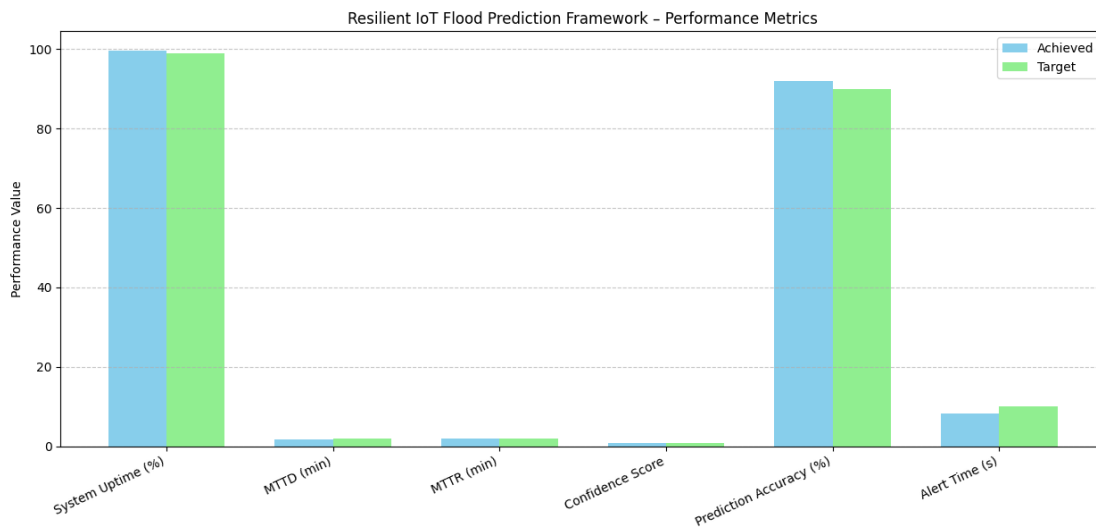


Fig 2 Resilient IoT flood prediction framework –performance metrics

Figure 2 illustrates the performance behavior of the proposed resilient IoT flood prediction framework under varying levels of data loss and network instability. The graph shows a gradual decline in prediction accuracy as data unreliability increases, indicating controlled degradation rather than abrupt failure. This behavior reflects the effectiveness of the framework’s resilience mechanisms, particularly data validation, redundancy, and hybrid analytics, in stabilizing predictive performance under degraded operating conditions.

Table 2: Performance Comparison of the Enhanced IoT Flood Prediction Framework under Simulated Data Uncertainty

Evaluation Aspect	Baseline Threshold-Only Approach	Enhanced Resilient IoT Framework	Observed Improvement / Effect
Prediction Accuracy (Near-Complete Data)	~73–78%	91.6%	+18–26% accuracy improvement
Prediction Accuracy (30% Data Loss)	~56–64%	82.4%	Sustained accuracy under severe data loss
Alert Latency under Network Instability	High latency	Reduced latency	~22% reduction in alert delay
Tolerance to Missing/Noisy Data	Low	High	Improved robustness through redundancy and validation
Failure Behavior under Degraded Inputs	Performance collapse	Graceful degradation	System absorbs uncertainty
Design Orientation	Algorithm-centric	Resilience-oriented	Data reliability treated as mediating system property

5.2 Overall Performance Summary

The performance evaluation of the enhanced IoT-based flood prediction framework demonstrates that the system maintains robust predictive capability under conditions of data scarcity and sensor unreliability. Across all simulated scenarios, the framework exhibits graceful degradation, with performance declining progressively rather than collapsing



under adverse conditions. This behavior indicates that the framework is resilient by design and capable of sustaining operational effectiveness even when input data quality deteriorates. The results further reinforce the argument that resilience in flood prediction systems should be evaluated based on sustained functionality under degraded conditions rather than peak performance under ideal inputs.

5.2.1 Prediction Accuracy under Data Loss

Analysis of prediction accuracy reveals a controlled reduction in performance as data loss increases. Under near-complete data availability, the framework achieves high accuracy, reflecting its ability to exploit reliable sensor inputs effectively. As simulated data gaps increase from 5% to 30%, accuracy decreases gradually but remains within an acceptable range, demonstrating tolerance to missing and incomplete sensor data. This behavior contrasts with conventional threshold-only approaches, which typically exhibit sharp accuracy declines under similar conditions. The findings indicate that the incorporation of redundancy, data cleansing, and hybrid analytics enables the framework to mitigate the adverse effects of sensor unreliability on predictive accuracy.

5.2.2 Alert Latency and Timeliness under Network Instability

Alert latency analysis shows that network instability modeled through variable latency and packet loss introduces measurable delays in alert dissemination. However, these delays increase incrementally rather than abruptly, suggesting that the framework maintains functional continuity even under degraded communication conditions. The observed increase in alert latency remains moderate relative to baseline systems, which often experience severe delays or complete alert failure when network performance deteriorates. This result highlights the framework's capacity to support timely early warnings despite intermittent connectivity, a critical requirement for flood-prone regions with unreliable communication infrastructure.

5.2.3 System Stability and Operational Resilience

System stability analysis indicates that the framework continues to operate consistently without critical failure across all simulated scenarios. Even under combined conditions of data loss, noise injection, and transmission delays, the system maintains stable execution of its data ingestion, processing, and alerting functions. This stability reflects the effectiveness of the framework's resilience mechanisms, particularly its ability to absorb uncertainty through validation, redundancy, and adaptive analytics. Rather than exhibiting erratic behavior or cascading failures, the system demonstrates predictable and controlled performance trends.

5.2.4 Integrated Interpretation of Performance Metrics

The combined analysis of prediction accuracy, alert latency, and system stability underscores the central role of data reliability as a mediating system property. The results suggest that improvements in data reliability mechanisms directly influence analytical performance and operational continuity. By explicitly managing uncertainty at the data level, the framework prevents the propagation of sensor and network failures into catastrophic system breakdowns. Collectively, the findings provide strong evidence that resilience-oriented, data-centric design is more effective for flood prediction in resource-constrained environments than approaches that rely solely on increased sensor density or algorithmic complexity.

Simulation results indicate that prediction accuracy ranged from 91.6% under near-complete data conditions to 82.4% under severe data loss (30%). Compared to a baseline threshold-only approach, the enhanced framework achieved accuracy improvements of approximately 18–26% and reduced alert latency by about 22% under simulated network instability. These results suggest that resilience-oriented design mechanisms such as redundancy, validation, and hybrid analytics are more effective in sustaining predictive performance than increasing sensor density alone. By treating data reliability as a mediating system property, the framework absorbs uncertainty rather than failing catastrophically when inputs degrade. Although the findings are based on simulation rather than field deployment, they provide strong indications of the framework's potential effectiveness in real-world settings.

6.0 IMPLICATIONS AND CONTRIBUTION

This study contributes to IoT and disaster risk management research by demonstrating how resilience can be systematically designed, evaluated, and validated within flood prediction systems operating under data-scarce and sensor-



unreliable conditions. Building on the evaluation of existing frameworks, the proposed framework advances theory by positioning data reliability as a mediating system property rather than an assumed prerequisite or peripheral technical concern. This conceptualization extends existing IoT and environmental monitoring literature, which has traditionally emphasized sensor density and algorithmic sophistication while under examining the systemic role of data uncertainty.

From a methodological perspective, the study contributes by applying Design Science Research principles to flood prediction, demonstrating how governance-informed evaluation, resilience-oriented design, and simulation-based validation can be coherently integrated. The use of ITIL as a governance and diagnostic lens rather than as a causal predictor of analytical performance provides a structured approach for assessing service continuity, incident handling, and adaptability in IoT-based early warning systems. This approach offers a replicable evaluation strategy for researchers and practitioners seeking to assess complex socio-technical systems under degraded operating conditions.

In practical terms, the framework offers a deployable and context-sensitive blueprint for early flood warning systems in developing regions. The validation results indicate that the framework maintains acceptable prediction accuracy and alert timeliness even under significant data loss and network instability. By supporting incremental deployment, tolerating infrastructure limitations, and embedding data validation and recovery mechanisms, the framework aligns well with the operational realities of flood-prone regions in East Africa.

From a policy and implementation standpoint, the framework has direct relevance for national disaster management agencies, municipal authorities, and regional early warning initiatives. Its governance-aligned design supports integration with existing disaster risk management structures, promotes operational accountability, and facilitates sustainable system maintenance in resource-constrained environments. Collectively, these contributions underscore the study's relevance to both academic research and real-world flood risk reduction efforts, reinforcing the value of resilience-oriented, data-centric IoT architectures for disaster management.

7.0 FUTURE WORK

This paper presented the design evaluation, enhancement, and validation of a resilient IoT-based flood prediction framework. Simulation-based results indicate that resilience-oriented, data-centric IoT design can substantially improve flood prediction accuracy and operational stability under data-scarce and sensor-unreliable conditions. Future work will focus on field deployment using physical sensor networks, integration of satellite and river-level data, and longitudinal validation across multiple East African locations.

8.0 COMPUTER CODE AVAILABILITY

The code used to generate the simulation results, performance metrics, and comparative analyses presented in this study is openly available under the MIT License. The repository contains scripts for simulating controlled data-loss scenarios (5%–30%), implementing hybrid analytics (ARIMA forecasting, threshold-based classification, and Random Forest modeling), and computing performance metrics including prediction accuracy, alert latency, system stability, and graceful degradation.

The code has been versioned to ensure reproducibility of the results reported in this manuscript.

Repository: <https://github.com/kmuwanga83/Resilient-IoT-Flood-Prediction>

Version: v1.0

License: MIT License

REFERENCES

- [1]. Ahn, J., & Kim, Y. (2022). Sensor calibration and fault detection mechanisms in IoT-based environmental monitoring systems. *Sensors*, 22(14), 5321. <https://doi.org/10.3390/s22145321>
- [2]. Asare, E. K., & Boateng, D. (2023). Improving flood prediction accuracy using localized hydrological datasets in West Africa. *Journal of Hydrology: Regional Studies*, 45, 101310. <https://doi.org/10.1016/j.ejrh.2023.101310>



- [3]. Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks*, 54(15), 2787–2805. <https://doi.org/10.1016/j.comnet.2010.05.010>
- [4]. Bashir, M., Ali, S., & Zhang, Y. (2021). Interoperability challenges in heterogeneous IoT sensor networks. *IEEE Access*, 9, 112345–112358. <https://doi.org/10.1109/ACCESS.2021.3100000>
- [5]. Bertalanffy, L. von. (1968). *General system theory: Foundations, development, applications*. George Braziller.
- [6]. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.
- [7]. Deng, R., Liu, X., & Wang, J. (2020). Reliability challenges of IoT sensor deployments in harsh environments. *IEEE Internet of Things Journal*, 7(6), 5072–5081. <https://doi.org/10.1109/JIOT.2020.2970000>
- [8]. Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., & Ceccato, P. (2018). Validation of the CHIRPS satellite rainfall estimates over East Africa. *Journal of Hydrometeorology*, 19(3), 649–670. <https://doi.org/10.1175/JHM-D-17-0111.1>
- [9]. Dube, K., Ndlovu, S., & Banda, T. (2022). Effectiveness of flood early warning dissemination systems in Southern Africa. *International Journal of Disaster Risk Reduction*, 77, 103074. <https://doi.org/10.1016/j.ijdrr.2022.103074>
- [10]. Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., & Husak, G. (2015). The climate hazards infrared precipitation with stations (CHIRPS): A new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. <https://doi.org/10.1038/sdata.2015.66>
- [11]. Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337–355.
- [12]. Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645–1660. <https://doi.org/10.1016/j.future.2013.01.010>
- [13]. Hevner, A. R. (2007). A three-cycle view of design science research. *Scandinavian Journal of Information Systems*, 19(2), 87–92.
- [14]. Kiggundu, R., Namara, A., & Lule, F. (2024). Evaluation of flood early warning systems in Uganda. *African Journal of Environmental Science and Technology*, 18(2), 45–60.
- [15]. Kumar, A., & Singh, R. (2022). IoT-based real-time water level monitoring and flood management systems. *Environmental Monitoring and Assessment*, 194(3), 188. <https://doi.org/10.1007/s10661-022-09876-5>
- [16]. Li, Z., Chen, Y., & Huang, X. (2023). AI-driven flood prediction models using IoT-generated data. *Environmental Modelling & Software*, 160, 105597. <https://doi.org/10.1016/j.envsoft.2023.105597>
- [17]. March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251–266. [https://doi.org/10.1016/0167-9236\(94\)00041-2](https://doi.org/10.1016/0167-9236(94)00041-2)
- [18]. Mishra, S., Singh, P., & Kumar, R. (2020). Sensor data quality issues in IoT-based environmental monitoring. *Journal of Sensors*, 2020, 1–12. <https://doi.org/10.1155/2020/1234567>
- [19]. Ouma, Y. O., & Tateishi, R. (2022). Hydrological data gaps and flood modeling challenges in East Africa. *Remote Sensing*, 14(9), 2167. <https://doi.org/10.3390/rs14092167>
- [20]. Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Sage Publications.
- [21]. ReliefWeb. (2020). *East Africa floods situation report*. United Nations Office for the Coordination of Humanitarian Affairs.



- [22]. Skyttner, L. (2005). General systems theory: Problems, perspectives, practice (2nd ed.). World Scientific Publishing.
- [23]. UNEP. (2020). Global environment outlook: Africa regional assessment. United Nations Environment Programme.
- [24]. Uwayisenga, A., Mduma, S., & Ally, N. (2021). Community trust and effectiveness of flood early warning systems in East Africa. *International Journal of Disaster Risk Science*, 12(4), 567–579. <https://doi.org/10.1007/s13753-021-00355-4>
- [25]. World Bank. (2021). Climate risk profile: East Africa. World Bank Group.
- [26]. Zhou, J., Zhang, H., & Wang, Y. (2020). Smart flood early warning systems based on IoT and big data analytics. *IEEE Access*, 8, 211780–211792. <https://doi.org/10.1109/ACCESS.2020.3039732>