

ANN–HEC-RAS Integrated Modeling for Flood Inundation Mapping in the Bagmati River at Khokana, Nepal

Aasish Shrestha

Department of Geomatics Engineering,
School of Engineering, Kathmandu University,
Dhulikhel, Kavre, Nepal
sthaaasish4@gmail.com

Arbin Chaudhary

Department of Geomatics Engineering,
School of Engineering, Kathmandu University,
Dhulikhel, Kavre, Nepal
arbinchaudhary.0006@gmail.com

Ganesh Saud

Department of Geomatics Engineering,
School of Engineering, Kathmandu University,
Dhulikhel, Kavre, Nepal
saudganesh2058@gmail.com

Rajesh Rokaya

Department of Geomatics Engineering,
School of Engineering, Kathmandu University,
Dhulikhel, Kavre, Nepal
rajeshrokaya758@gmail.com

Sahadev Sharma

Department of Geomatics Engineering,
School of Engineering, Kathmandu University,
Dhulikhel, Kavre, Nepal
sharmasahadev447@gmail.com

Abstract— Floods in the Bagmati river pose serious risks to the residents of Khokana, with their floodplain gradually being encroached on due to urbanization. Modeling floods in areas that lack sufficient hydrological measurements and topographical information can be difficult. In this paper, an Artificial Neural Network (ANN) rainfall–runoff model was combined with a 2-dimensional (2D) hydraulic HEC-RAS model to provide high-resolution flood inundation maps of the Bagmati River reach in Khokana. Seven-year daily hydro-meteorological data (2019–2025) were applied to train (2019–2023), validate (2024), and test (2025) chronologically without using early stopping and/or model selection on any other period except the validation year in order to avoid overfitting. A feedforward multilayer perceptron (MLP) neural network, trained to minimize a percentile-targeted, peak-weighted loss function, scored a Nash-Sutcliffe Efficiency (NSE) of 0.86 and a peak-flow percent bias (PBIAS) of 2.0% for the independent year 2025 test. A recurrent Long Short-Term Memory (LSTM) neural network model was further compared to the MLP model under the same chronological approach. A 14.9 cm high-resolution Digital Terrain Model (DTM), obtained from a UAV, was employed for capturing micro-topography of the embankments and floodplains, and the flood extents delineated using this model were compared against the satellite-derived Normalized Difference Water Index (NDWI) flood extents obtained from Sentinel-2 at the 10 m native resolution. Furthermore, discharge sensitivity analysis helps to quantify the propagation of residual peak flow errors from ANN to inundation extent and depth, thereby making a clear connection between the hydrological model performance and the consequence of the flood mapping. The results demonstrate that an ANN–HEC-RAS framework, rigorously validated with a held-out test year, can support disaster risk reduction and land-use planning in data-scarce, rapidly urbanizing floodplains such as Khokana.

Keywords—Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), HEC-RAS, UAV photogrammetry, Digital Terrain Model (DTM), Sentinel-2, NDWI, peak-flow sensitivity analysis, discharge sensitivity analysis

I. INTRODUCTION

A flood happens when a river's stream goes beyond its banks, causing the waters to overflow into the surrounding floodplains

[1]. Although such overflow is a normal hydrological process, it becomes a disaster when riverside ecological systems, man-made structures, and human activities are negatively affected. Human activities such as rapid urbanization, deforestation, and riverbank encroachment significantly increase flood risk by reducing a river's natural capacity to manage excess water [2]. For Geomatics engineers, flood studies are essential because they translate complex numerical water data into visual spatial maps used for urban planning and early-warning systems [3]. Globally, floods account for nearly half of all weather-related disasters [4], and Asia, heavily influenced by the South Asian Monsoon and large river systems supporting dense populations, is the most flood-prone continent on Earth [5].

In Nepal, it is not uncommon to have floods occur every year. This nation has some of the highest mountains on earth which gradually descend to low plains of Terai resulting in rapid flowing rivers. During the rainy season (June–September), this nation gets around 80% of its annual precipitation resulting in floods and landslides.

In the past, incidents like the flood of 1993 in central Nepal have caused damages to crucial infrastructures including bridges and dams [6]. Kathmandu valley has undergone tremendous urbanization, and the Bagmati River, which flows through the valley, has been subjected to extensive floodplain encroachment. This paper focuses on the village of Khokana, located where the river passes through a narrow gorge. It examines flooding during the peak monsoon season, when discharge exceeds the channel's flow capacity, inundating agricultural fields and settlements. Traditional hydrological models require extensive physical data that is often unavailable in Nepal. To overcome this, an Artificial Neural Network (ANN) is applied: a data-driven technique capable of learning complex, non-linear patterns from historical rainfall and temperature data to predict river discharge with high accuracy even when physical watershed parameters are incomplete [7]. This predicted discharge drives HEC-RAS, a hydraulic modeling software developed by the US Army Corps of Engineers, which routes discharge through the river channel and predicts flood extent using a high-resolution Digital Terrain Model reconstructed from UAV photogrammetry and DGPS

surveying [8]. This integration allows the ANN to act as the hydrological “brain” while HEC-RAS serves as the hydraulic “visualizer,” together producing markedly more detailed inundation maps than those derived from coarse satellite elevation data.

Despite recurrent flood disasters, the Khokana area lacks detailed, locally validated flood inundation maps. Most existing flood prediction models rely on low-resolution, global-scale digital elevation models that do not capture local riverbank and floodplain variation, making inundation predictions inaccurate. Traditional hydrological models also struggle to capture the complex, non-linear, and “flashy” nature of the Bagmati River, and typically require extensive calibration datasets that are unavailable in data-scarce regions. Furthermore, a rainfall–runoff model that is only validated against a single training period, without an independent, held-out year used exclusively for model selection, risks overstating its own generalization ability. Without a reliable method linking accurate, properly validated rainfall–runoff prediction with high-resolution terrain data, local authorities face significant barriers to effective disaster preparedness and land-use planning.

The main objective of this study is to simulate flood inundation by integrating a machine-learning (ANN) rainfall–runoff model with HEC-RAS, under a validation methodology that supports the resulting predictions. The specific sub-objectives are:

- 1) To generate a high-resolution DTM of the Bagmati reach at Khokana using UAV photogrammetry and DGPS-based Ground Control Points (GCPs), and to develop a rainfall–runoff ANN, benchmarked against a recurrent (LSTM) architecture, trained and validated under a chronological train/validation/test split to predict river discharge.
- 2) To develop precise spatial flood inundation maps for the Bagmati River–Khokana area and validate model accuracy using remote-sensing indicators (NDWI), and to quantify how residual peak-discharge error propagates into inundation extent and depth.

The present research concerns the southern portion of the Bagmati River at Khokana, focusing on the technical integration of geomatics and machine-learning tools for flood simulation. It uses a high-resolution DTM generated from UAV photogrammetry and DGPS ground control data, an ANN to predict river discharge, and HEC-RAS to simulate hydraulic flow. The final output is confined to flood inundation mapping (a spatial visualization of flood depth and extent), together with a quantitative sensitivity analysis of how discharge prediction error affects that mapping. Future work could integrate climate and hydrological variability, advanced hydrodynamic simulation, and coupling with socio-economic risk assessment.

II. LITERATURE REVIEW

Traditionally, flood mapping has relied on process-based hydraulic models, including HEC-RAS, which uses numerical solutions of either the Saint-Venant or Shallow Water equations to model flood extent and characteristics such as velocity. The

accuracy of such models depends largely on the hydrological boundary conditions used, and in data-poor areas, the lack of discharge records is usually the weakest link in the chain. This has driven a paradigm shift toward hybrid, data-driven frameworks in which machine-learning (ML) techniques for the hydrological component are combined with detailed hydraulic simulation from physical models [9].

A wide array of ML algorithms, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forests (RFs), and Long Short-Term Memory (LSTM) networks, have been successfully applied to rainfall–runoff prediction [10]. Tamiru and Wagari demonstrated that an ANN could reliably predict daily river discharge from meteorological inputs alone in the Baro River Basin, Ethiopia, effectively bridging critical hydrological data gaps [11]. Complementary work in the Awash Basin emphasized data-driven feature selection to identify the most influential flood-causative factors prior to model training [12]. Despite the theoretical suitability of such models as LSTM for modeling the sequence-dependent process of rainfall–runoff transformation [10], this superiority over simpler models is only evident when the training set is large, long-term and multi-basin. The small single-station, single-basin daily dataset that might be obtained from mountainous regions with poor data availability will not be enough to prove this superiority. Therefore, model selection should be based on comparative evaluation under the same data-splitting and validation protocol rather than assuming the superiority of LSTM a priori.

Within the hybrid framework, HEC-RAS remains the dominant hydraulic simulation engine due to its robust 1D and 2D capabilities. 2D modeling is essential for simulating complex overland flow in urban and peri-urban environments such as Khokana, where flow paths are strongly influenced by structures and micro-topography. HEC-RAS requires geometric data (channel and floodplain shape, derived from a DEM/DTM), hydraulic parameters (Manning's roughness coefficients), and boundary conditions (a flow hydrograph), the last of which is the critical point of integration with the ML hydrological output, and therefore the point at which hydrological model error most directly propagates into hydraulic model error.

The credibility of such integrated models is established through standardized statistical validation. As established by Tamiru and Wagari, NSE, PBIAS, and R^2 form the standard hydrological validation suite, while spatial inundation accuracy is assessed through overlapping-percentage agreement between simulated and remotely sensed flood extents, with over 85% overlap considered good agreement [11]. Beyond aggregate-period statistics, however, few studies explicitly quantify how a model's peak-flow-specific error, as opposed to its overall-period error, propagates into inundation extent and depth, even though peak flows are what determine maximum flood extent in practice.

The accuracy of any inundation model is fundamentally constrained by topographic data quality. Publicly available satellite-derived DEMs (12.5–30 m resolution) are insufficient for complex urban or peri-urban flood dynamics [13]. UAV-

based photogrammetry has emerged as a best practice for generating centimeter-level DEMs, enabling explicit representation of anthropogenic features such as embankments, road crowns, and retaining walls that control flood pathways [14].

Flood research in Nepal and the Bagmati Basin has predominantly relied on conventional HEC-RAS modeling using DHM discharge records combined with coarse global DEMs such as SRTM or ASTER [15], [16]. Statistical and GIS-based approaches, including Flood Frequency Analysis, have also been used to estimate return periods [17]. While foundational, these studies are limited by sparse gauging networks and coarse topographic resolution, and detailed, site-specific inundation information for densely populated areas like Khokana remains largely absent from the literature.

A. Research Gap

Three gaps motivate this study. The Contextual Gap concerns the study area: benchmark ANN–HEC-RAS studies such as [11] were conducted in large, predominantly rural river basins, whereas Khokana is a densely populated, peri-urban landscape whose flood dynamics are heavily influenced by channel encroachment and impervious surfaces. The Technological Gap concerns data resolution: prior studies relied on satellite DEMs of meter-level resolution, unable to resolve fine-scale peri-urban features. The Methodological Gap concerns model validation and benchmarking: prior ANN–HEC-RAS studies in this basin have not (i) used a held-out validation year disjoint from both training and test data for model selection, (ii) benchmarked the feedforward ANN against a recurrent architecture under an identical split, or (iii) quantified how the ANN's residual peak-discharge error translates into inundation-mapping error. This work addresses all three gaps using an ultra-high-resolution (14.9 cm) UAV-derived DTM in a human-modified, peri-urban floodplain.

III. METHODOLOGY

This section describes the study area, datasets, and the integrated ANN–HEC-RAS methodology used to generate and validate flood inundation maps for the Bagmati River at Khokana, summarized in Fig. 1.

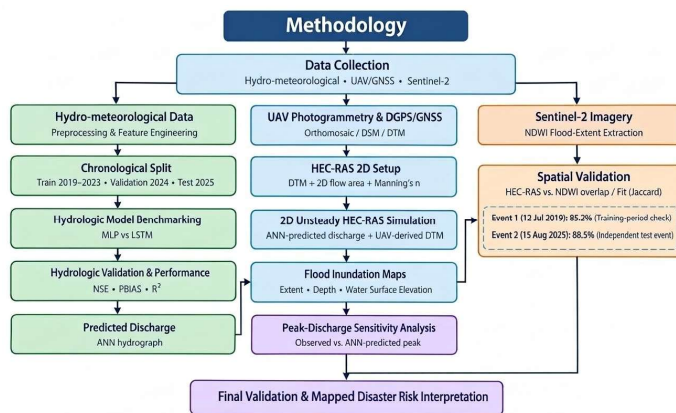


Fig. 1. Overall methodology workflow integrating ANN-based discharge prediction

A. Study Area

The study site is situated in the Khokana village of the Bagmati River basin, extending between 27°37'00"N–27°38'15"N and 85°17'30"E–85°18'00"E. It covers an approximate area of 1.48 km² and a perimeter of 5.49 km (Fig. 2). It encompasses parts of Wards 1, 2, and 4 of Dakshinkali Municipality and Wards 21 and 22 of Lalitpur Sub-Metropolitan City. The study area features sloping topography and Bagmati River floodplains originating from the Shivapuri hills. A DHM-operated gauging station, Khokana Station (550.05), was established within the study area, from which discharge data was extracted.

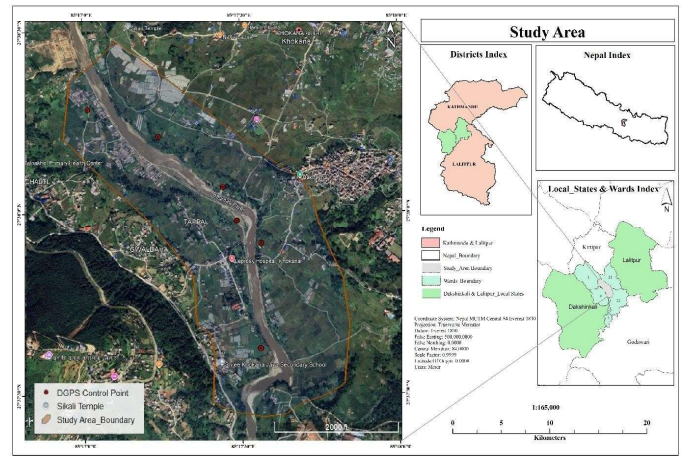


Fig. 2. Study area map of the Bagmati River reach at Khokana, showing DGPS control points and the study area boundary.

B. Datasets and Software

Daily discharge and hydro-meteorological data spanning seven years (January 2019 to December 2025) were recorded using the Khokana Gauging station and the seven meteorological stations of the Department of Hydrology and Meteorology (DHM). The DTM was created at 14.9 cm resolution from UAV surveys, and NDWI validation was conducted using Sentinel-2 MultiSpectral Instrument (MSI) Level-2A surface-reflectance imagery at native 10 m resolution (Green Band 3, 560 nm, and Near-Infrared Band 8, 842 nm), a threefold improvement in spatial detail over the 30 m multispectral bands of Landsat 8 commonly used in earlier flood-validation studies. Instruments used were the DJI Mavic 2 Pro UAV and the Spectra Precision SP60 GNSS receiver. Software used includes Pix4Dmapper (UAV processing), ArcGIS 10.8 (spatial analysis), Python with PyTorch and scikit-learn (ANN/LSTM development), HEC-RAS 6.7 (hydraulic modeling), Google Earth Engine (Sentinel-2 acquisition, cloud masking, and NDWI computation), and Microsoft Excel (data management).

C. Data Collection

Hydro-meteorological data spanning seven years (January 2019 to December 2025) were collected from DHM. Raw meteorological data were quality checked and aggregated across stations. Aerial images were collected using a DJI Mavic 2 Pro via DroneDeploy and DJI GO 4, with 75% front overlap and 65% side overlap. A network of DGPS-surveyed Ground Control Points (GCPs) was established, with the base station coordinate resolved via AUSPOS and remaining stations computed in Trimble Business Center. UAV imagery was processed in Pix4Dmapper using Structure-from-Motion algorithms to generate an Orthomosaic, Digital Surface Model (DSM), and Digital Terrain Model (DTM) (Fig. 3); GCP-based georeferencing achieved a horizontal RMSE of approximately 2 mm and a vertical RMSE of approximately 9 mm, well within the tolerance required to resolve embankment and floodplain micro-topography at the DTM's 14.9 cm resolution.

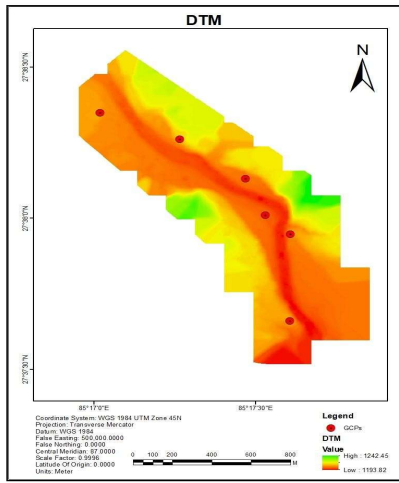


Fig. 3. Ultra-high-resolution (14.9 cm) Digital Terrain Model (DTM) generated from UAV photogrammetry and DGPS ground control points.

D. Hydrologic Modeling using ANN

Feature engineering. In addition to same-day rainfall, minimum temperature, and maximum temperature, the discharge series was expanded with lagged discharge terms (1, 2, 3, 4, 7, 10, 14, 21, and 30 days); rolling rainfall statistics (3- and 7-day means, 14- and 30-day sums, 3-day maximum, and 3- and 7-day exponential moving averages); an antecedent precipitation index (API) computed as a weighted sum of the previous three days' rainfall; 7- and 14-day rolling discharge means; diurnal temperature range; and cyclical day-of-year encodings (sin, cos) to represent seasonality. This yielded a 25-feature input vector per day.

Chronological train/validation/test split. To prevent the tuning-leakage risk inherent in a two-way split, the seven years of history were divided into a training set (2019 to 2023, $n=1,796$), validation set (2024, $n=366$), and test set (2025, $n=365$) along chronological lines (without any randomization, as that would leak future information into the training process). The validation set consists of data points which are disjoint in time

from the training and test sets and was only used to implement early stopping and model architecture/hyperparameter selection. Feature scaling (StandardScaler) and the log-transform of the target ($\log(1+Q)$, used to stabilize variance across the wide dynamic range of daily discharge) were fit only on the training set and applied unchanged to validation and test.

Feedforward MLP architecture. The main architecture is a feedforward neural network having three hidden layers (with 256, 128, and 64 units, respectively) along with ReLU activation, batch normalization, and dropout (for the first two layers at 0.20 and 0.15, respectively) to prevent overfitting given the small dataset. Training was done using the Adam optimizer and learning rate scheduling in which the learning rate was reduced to half in case of plateauing validation loss, and early stopping criterion (patience = 40 epochs) based on validation loss was used.

Peak-focused loss function. Because flood-relevant extremes are rare relative to the bulk of the daily record, a naive mean-squared-error loss under-penalizes error on high-flow days. Training therefore used a peak-weighted loss:

$$L = \left(\frac{1}{N}\right) \sum_{i=1}^N w_i^2 (\hat{y}_i - y_i)^2, w_i = 1 + 4 \cdot I[y_i > \tau] \quad (1)$$

where y_i is the observed transformed discharge, $\hat{y}_i$ is the ANN prediction, $I(\cdot)$ is the indicator function, and τ is the 90th percentile of the training-set target. This concentrates additional gradient weight specifically on genuine high-flow extremes, rather than on the broader, less-informative set of merely above-average days.

Recurrent (LSTM) benchmark. To empirically test whether a recurrent architecture better exploits the sequential, memory-dependent structure of rainfall-runoff response, a single-layer LSTM (32 hidden units, dropout 0.10) was trained on 10-day sliding windows of the same 25-feature set, under an identical chronological split, loss function, and early-stopping criterion, with gradient-norm clipping (1.0) and a reduced learning rate (10^{-4}) for training stability.

Validation-derived peak bias correction. As a final, optional post-processing step, a multiplicative correction factor is computed as the ratio of observed to predicted discharge summed over validation-year days exceeding the 90th-percentile (unscaled) training threshold. This factor is applied to test-year predictions only if the raw validation-year peak PBIAS exceeds $\pm 10\%$ (the acceptability band used throughout this study); otherwise, the raw predictions are retained, since a correction derived from one year's bias direction can overshoot if that bias has already been reduced by the training procedure itself.

E. HEC-RAS Model Setup

The hydraulic model was built in HEC-RAS 6.7, with the high-resolution DTM imported as topography. A 2D Flow Area was defined over the channel and adjacent floodplain using RAS Mapper, with a computational mesh of uniform $2 \text{ m} \times 2 \text{ m}$ cells across the flow area, a resolution chosen to resolve channel and embankment geometry while keeping the $\sim 370,000$ -cell domain

computationally tractable given the 1.48 km² study extent. Cell-center elevations were derived from the 14.9 cm UAV-DTM using RAS Mapper's default terrain-to-mesh sampling, so that fine-scale topographic variation from the underlying DTM is incorporated into the hydraulic computation even though the mesh itself is coarser than the DTM's native resolution.

Manning's roughness coefficients were assigned by land-cover type following established parameterization values from the literature [18]: river = 0.03, forest = 0.05, grass = 0.03, flooded vegetation = 0.045, cropland = 0.025, shrubs = 0.035, built-up = 0.013, and bare soil = 0.025. Given the absence of surveyed high-water marks or a rated calibration event at Khokana, these values were adopted directly from published parameterization rather than site-specific calibration; this is treated as a limitation and discussed alongside other sources of model uncertainty in Section IV-F. The upstream boundary was set according to the ANN-predicted discharge hydrograph for the selected flood events; the downstream boundary used the normal-depth option following the channel slope.

F. Flood Inundation Mapping and Validation

The 2D unsteady flow solver computed water surface elevation, depth, and velocity across the 2D flow area for the selected flood events, producing flood depth and water-surface-elevation grids in RAS Mapper and, ultimately, an inundation-extent polygon. To validate this extent independently, surface-water masks were derived from Sentinel-2 MSI Level-2A (atmospherically corrected, Sen2Cor-processed) imagery in Google Earth Engine, using the cloud-least scene available within the shortest practicable window of each flood peak. For each scene, the Normalized Difference Water Index was computed as $NDWI = (Green - NIR)/(Green + NIR)$ using the native 10 m Green (B3) and Near-Infrared (B8) bands, with cloud and cloud-shadow pixels masked using the Scene Classification (SCL) band prior to computation. Pixels with NDWI greater than zero were classified as open water, consistent with the conventional McFeeters formulation [19]; the resulting raster mask was filtered to remove sub-pixel noise (a 3×3 majority filter) and converted to vector polygon form in ArcGIS Pro, following established practice for Sentinel-2-based surface-water delineation at sub-30 m spatial fidelity [20]. Agreement between HEC-RAS-predicted and NDWI-derived flood extents was evaluated using the Fit/Jaccard statistic, applied uniformly to both the training-period consistency event and the independent test-period event.

$$F = \frac{A_{sim} \cap A_{obs}}{A_{sim} + A_{obs} - (A_{sim} \cap A_{obs})} \quad (2)$$

where $A(sim)$ is the inundation area simulated using HEC-RAS and $A(obs)$ is the NDWI-derived reference area. The Jaccard/Intersection-over-Union (IoU) metric divides the intersection by the sum of the areas of the simulated and reference extents, thus addressing both the issues of over- and underestimation of the inundation area. It is, thus, a more cautious measure of spatial agreement compared to an overlap ratio that is normalized with respect to the reference extent only. An 85% level of agreement is considered as the agreement threshold based on the criteria outlined in [11].

Model performance was additionally evaluated using NSE, PBIAS [21], [22], and R^2 , with acceptability thresholds of $NSE \geq 0.75$, $PBIAS \in [-10,10]$, and $R^2 > 0.85$ for "Very Good" performance.

G. Peak-Discharge Sensitivity Analysis

To directly connect residual ANN discharge error to its hydraulic consequence, HEC-RAS was additionally run twice for the largest discharge event in the independent 2025 test year, holding the DTM, Manning's roughness values, and downstream boundary fixed: once using the observed discharge as the upstream boundary, and once using the ANN-predicted discharge for the same event. The resulting inundation extent (km²) and maximum depth at fixed reference points near the Khokana settlement and heritage site were compared between the two runs, isolating the inundation-mapping impact attributable purely to the ANN's discharge error, independent of any other source of model uncertainty.

IV. RESULTS AND DISCUSSION

A. ANN Hydrological Model Performance

Table I summarizes MLP performance across all three chronologically disjoint splits. Training-set performance ($NSE = 0.78$) is intentionally more conservative than the near-perfect fits often reported for two-way (train/test-only) splits, because the peak-focused loss and dropout regularization used here trade a small amount of training-set fit for improved generalization and reduced overfitting risk, the very concern motivating the validation split. Validation-year performance ($NSE = 0.87$, $PBIAS = -4.21\%$) confirms the model generalizes to a wholly unseen year, and the untouched, independent 2025 test year achieves $NSE = 0.86$ with an overall PBIAS of -5.25% , within the $[-10,10]$ "Good" acceptability band used throughout this study.

TABLE I. ANN (MLP) Model Performance — Chronological Splits

Split	Yrs	NSE	R^2	PBIAS(%)	PkPB(%)
Train	19–23	0.78	0.78	-7.59	-8.76
Valid.	2024	0.87	0.87	-4.21	4.80
Test	2025	0.86	0.86	-5.25	2.00

Critically for flood-mapping purposes, peak-flow PBIAS (computed over the top 5% of observed discharge days) on the independent test year is only 2.00%, within the overall acceptability band used throughout this study. The single largest observed 2025 event (102.9 m³/s) was predicted at 130.6 m³/s, a moderate overestimation rather than the underestimation that is more typical of peak-flow prediction in rainfall-runoff modeling; this event is carried forward into the sensitivity analysis of Section IV-D. Because the validation-year peak PBIAS (4.80%, Table I) already falls within the $\pm 10\%$ band specified in Section III-D, the validation-derived correction was not triggered, and raw test-year predictions were retained unmodified.

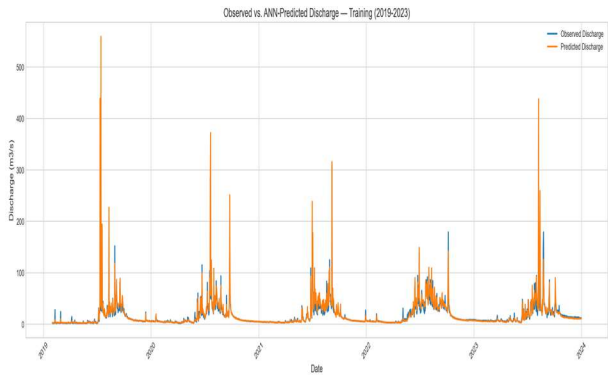


Fig. 4. Observed vs. ANN-predicted discharge, training period (2019–2023).

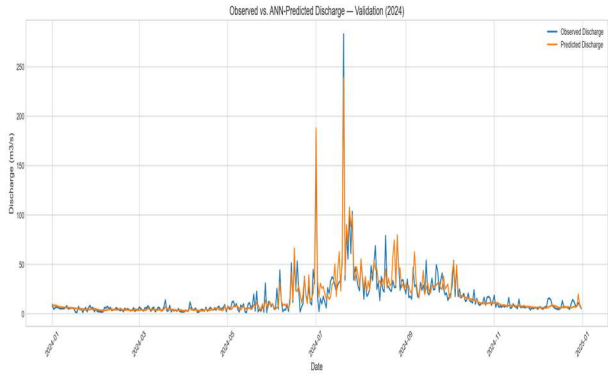


Fig. 5. Observed vs. ANN-predicted discharge, validation period (2024)

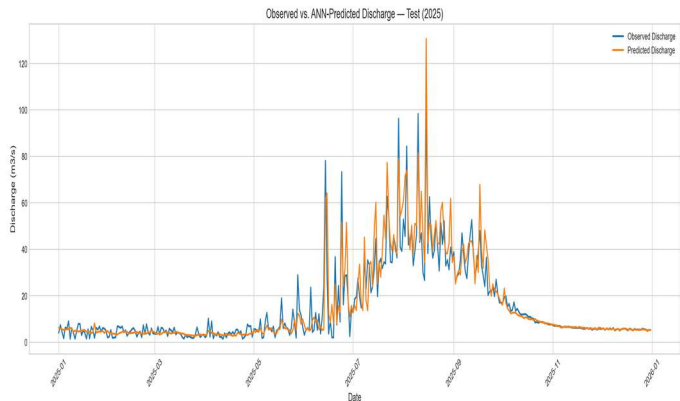


Fig. 6. Observed vs. ANN-predicted discharge, independent test period (2025).

B. Benchmarking Against a Recurrent (LSTM) Architecture

Table II compares the feedforward MLP against the LSTM benchmark under an identical chronological split, loss function, and early-stopping criterion. Despite hyperparameter tuning (reduced learning rate, gradient clipping, and a reduced hidden-layer size to match the available sample size), the LSTM did not outperform the MLP on any split, and its training-set NSE (0.39) indicates it did not fully fit the training signal itself, rather than merely generalizing poor.

TABLE II. MLP vs. LSTM Comparison Across Chronological Splits

Split	MLP NSE	LSTM NSE	MLP PBIAS%	LSTM PBIAS%	MLP PkPBIAS%	LSTM PkPBIAS%
Training	0.78	0.39	−7.59	−3.64	−8.76	53.07
Validation	0.87	0.32	−4.21	−11.66	4.80	49.09
Test	0.86	0.59	−5.25	−15.00	2.00	33.19

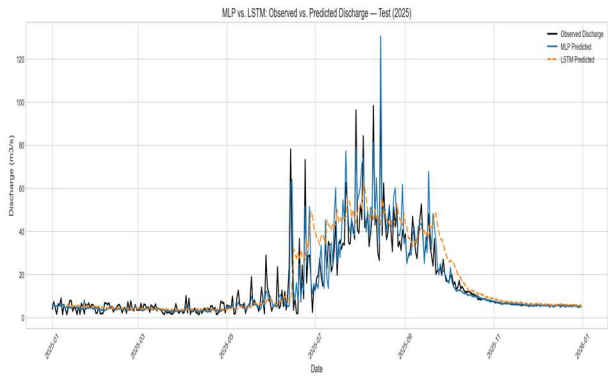


Fig. 7. MLP vs. LSTM: observed vs. predicted discharge, test period (2025).

A plausible explanation is feature redundancy across the sequence window: the 25-feature vector already includes discharge lags of 1–30 days computed globally, so consecutive time steps within a 10-day LSTM window are highly collinear, diluting the sequential signal a recurrent architecture is designed to exploit. More generally, recurrent architectures are typically reported to outperform feedforward networks on rainfall–runoff tasks when trained on substantially larger, often multi-basin datasets [10]; the ~1,800-day single-station record available here may simply be insufficient to realize that advantage. We report this comparison in full rather than omitting the LSTM, since an honest negative result, obtained under a fairly tuned, identical evaluation protocol, is itself informative for future work in similarly data-scarce basins.

C. HEC-RAS Validation against NDWI

HEC-RAS inundation maps based on ANN-simulated discharge for the 12 July 2019 event, included as a training-period consistency check, were compared with the NDWI-derived flood extent from Sentinel-2 imagery (Fig. 8). The HEC-RAS-simulated inundation extent was 127,201.64 m², while the Sentinel-2 NDWI-derived reference extent was 114,982.16 m². Their spatial intersection was 111,444.86 m², and the union area was 130,738.94 m². Using the Fit/Jaccard statistic defined in Eq. (2), the resulting spatial agreement was ($F=85.2\%$), slightly exceeding the adopted 85% agreement benchmark.

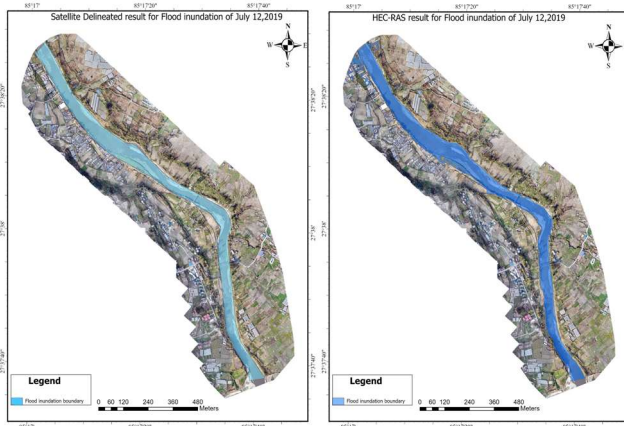


Fig. 8. Comparison of NDWI-delineated satellite and HEC-RAS-simulated flood inundation extents for the 12 July 2019 event (training-period consistency check).

Because the revised chronological split uses 2019–2023 for training and 2025 for independent testing, an out-of-sample inundation assessment was performed using the largest flood event available in the 2025 test year, 15 August 2025. The ANN-predicted peak discharge for this event was 130.61 m³/s. The HEC-RAS-simulated inundation extent was 107,070.42 m², while the Sentinel-2 NDWI-derived reference inundation extent was 100,105.71 m². Their spatial intersection was 97,237.35 m², and the union area was 109,938.77 m². Using the Fit/Jaccard statistic defined in Eq. (2), the resulting spatial agreement was ($F=88.5\%$). This result indicates substantial spatial agreement between the HEC-RAS-simulated inundation extent and the independently derived Sentinel-2 flood extent for the 15 August 2025 event. Because this independent spatial validation is based on a single 2025 event, it provides event-level evidence of spatial consistency rather than definitive evidence of generalization across all flood events.

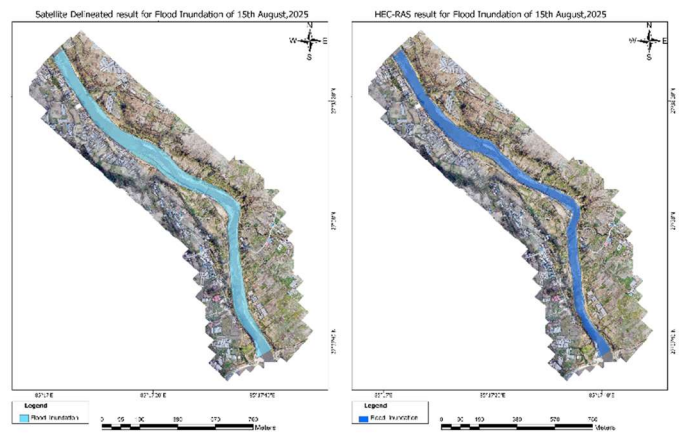


Fig. 9. Comparison of NDWI-delineated and HEC-RAS-simulated flood inundation extents for the 15 August 2025 event (independent test-period spatial validation)

The independent test-period event achieved a Fit/Jaccard value of 88.5%, exceeding the adopted 85% agreement benchmark and indicating substantial spatial agreement between the HEC-RAS-simulated and Sentinel-2 NDWI-derived inundation extents. However, because the spatial validation is based on a single 2025 event, this result should be interpreted as event-level evidence of spatial consistency rather than as definitive evidence of generalization across all flood events. Together, the two events provide consistency between the simulated and remotely sensed inundation extents, while the 2025 event provides the independent test-period spatial assessment.

TABLE III. HEC-RAS vs. NDWI Overlap Agreement

Event	Period type	Date	Fit/Jaccard(%)
Event 1	Training	12 Jul 2019	85.2
Event 2	Independent Test	15 Aug 2025	88.5

D. Peak-Discharge Sensitivity Analysis

Table IV reports the outcome of the paired HEC-RAS run described in Section III-F, isolating the inundation-mapping consequence of the ANN's discharge error on the single largest 2025 test-year event (observed 102.9 m³/s vs. ANN-predicted 130.6 m³/s, a 27.7 m³/s / 26.9% overestimation).

TABLE IV. Discharge Sensitivity: Observed vs. ANN-Predicted Peak Event (15 Aug 2025)

Scenario	Inflow(m ³ /s)	Area(km ²)	Depth(m)
Observed	102.9	0.1039	1.064
ANN-pred	130.6	0.1071	1.190
Difference	+27.7 (+26.9%)	+0.0032 (+3.04%)	+0.126 (+11.83%)

On the 15 August 2025 peak event, the 26.9% discharge overestimation by the ANN (130.6 vs. 102.9 m³/s observed) translated into only a 3.04% increase in inundated area (0.1071 vs. 0.1039 km²) but an 11.83% increase in depth at the fixed reference site (1.190 vs. 1.064 m). For this event, the relatively

small change in inundated area compared with the change in peak discharge suggests that the simulated flood extent was less sensitive than the depth at the reference site to the tested discharge error. This behavior may be associated with local channel and floodplain geometry, which can constrain lateral expansion while allowing additional discharge to increase water depth within already inundated areas. A practical implication for disaster planning is that extent-only flood maps may understate the consequence of peak-discharge error at structures and settlements within the existing flood boundary, even when the mapped boundary itself changes only modestly; depth-based risk indicators should therefore be reported alongside extent-overlap statistics such as those in Section IV-C.

E. Comparative Analysis with Reference Study

As compared to the benchmark study conducted in Ethiopia [11], regarding the Baro River Basin, having an NSE range of 0.86-0.88, R^2 range of 0.91-0.93, and 94.6-96% spatial overlap using a 30 m DEM at basin scale level, the Khokana model performed equally well during the test year period in terms of NSE (value=0.86) and R^2 (value=0.86). The lower R^2 value in the current study could be partially due to the fact that a shorter independent test period was used in this study. Nevertheless, there is a substantial difference between the two studies in terms of scale and terrain representation, as the Khokana framework makes use of a 14.9 cm UAV-derived DTM to represent floodplain morphology whereas [11] uses a 30 m DEM at basin scale level. Thus, the reported spatial overlap values cannot be treated as equal since the definitions and spatial resolutions are different in each case. As far as the spatial agreement measured in the current study using the union-normalized Jaccard/Fit metric is concerned, it was 88.5% for the independent test event of 15 August 2025.

F. Limitations

The spatial sparsity of meteorological stations limits the ANN's ability to capture full localized rainfall variability during extreme storms. The LSTM benchmark, while carefully tuned, may be constrained less by architecture than by the length of a single-station daily record; a multi-basin or sub-daily dataset may better reveal any recurrent-model advantage. NDWI-based validation is also time-dependent on satellite overpass timing. Sentinel-2's combined two-satellite constellation provides a revisit interval of approximately 5 days at the study latitude, an improvement over Landsat's 16-day repeat cycle, but still imperfect, so any mismatch between peak flood occurrence and image acquisition can result in underestimation of true flood extent if floodwaters recede before acquisition. Finally, the discharge-sensitivity analysis in Section IV-D isolates the effect of discharge error under a fixed DTM and roughness parameterization; it does not capture compounding uncertainty from terrain or Manning's n error, which would require a separate, multi-factor sensitivity study. Relatedly, the Manning's roughness values used in Section III-E were adopted from published land-cover-based parameterization [18] rather than calibrated against site-specific observations, since no surveyed high-water marks or rated flood event were available at Khokana

for calibration; the resulting roughness uncertainty is therefore unquantified and is a natural target for future work, ideally alongside a formal mesh-convergence study for the 2 m computational grid.

V. CONCLUSION

This study demonstrates an integrated ANN-HEC-RAS framework for high-resolution flood inundation mapping in the data-scarce, rapidly urbanizing Bagmati River Basin at Khokana, Nepal, validated under a chronological train (2019–2023), validation (2024), and test (2025) split that keeps model-selection decisions strictly separate from final reported performance. The feedforward MLP achieved NSE = 0.86 and peak-flow PBIAS of 2.0% on the fully independent 2025 test year, outperforming a comparably tuned LSTM benchmark under an identical evaluation protocol, a result attributable to feature redundancy across the sequence window and the modest scale of a single-station daily record. The centimeter-resolution UAV-based DTM continues to enable accurate representation of floodplain morphology not captured by SRTM or ASTER GDEM elevation models, and HEC-RAS inundation mapping driven by ANN-predicted discharge shows good agreement with NDWI-derived flood extents. A discharge-sensitivity analysis further connects the ANN's residual peak-flow error to its practical consequence for inundation extent and depth, directly addressing the link between hydrological model performance and flood-mapping reliability that is essential for disaster risk reduction and land-use planning at Khokana. On the largest 2025 test-year event, a 26.9% discharge overestimation produced only a 3.04% change in inundated area but an 11.83% change in peak depth at a fixed reference site, indicating that flood extent at this reach is comparatively robust to discharge-prediction error while depth-based risk at existing structures is more sensitive, a distinction with direct relevance for how ANN-HEC-RAS flood maps should be interpreted and communicated in operational disaster-planning use.

REFERENCES

- [1] UNESCO, "The United Nations World Water Development Report 2023: Partnerships and cooperation for water," UNESCO, 2023.
- [2] S. Thapa, R. Talchabhadel, J. Kanel, and Y. Iseri, "Impact of land use and land cover change on discharge and sediment yield of the Bagmati River Basin, Nepal," *J. Hydrol.: Reg. Stud.*, vol. 30, 2020.
- [3] C. C. Sampson et al., "A high-resolution global flood hazard model," *Water Resour. Res.*, vol. 51, no. 9, pp. 7358–7381, 2015.
- [4] Centre for Research on the Epidemiology of Disasters, "2023 disasters in numbers," UCLouvain, 2024.
- [5] H. Tamiru and M. Wagari, "Machine-learning and HEC-RAS integrated models for flood inundation mapping in Baro River Basin, Ethiopia," *Model. Earth Syst. Environ.*, vol. 8, no. 2, pp. 2291–2303, 2022.
- [6] D. K. Gautam and M. L. Kharbuja, "Flood management in Nepal," in *Proc. Int. Workshop Flash Flood Manage. Sustain. Develop. Himalayas*, 2006.

- [7] K. Abhishek, A. Kumar, R. Ranjan, and S. Kumar, "A data-driven model for rainfall-runoff forecasting using artificial neural networks," *Int. J. Comput. Sci. Issues*, vol. 9, no. 3, p. 582, 2012.
- [8] United States Army Corps of Engineers, "HEC-RAS, River Analysis System Hydraulic Reference Manual," Version 6.0, Hydrologic Engineering Center, 2020.
- [9] V. Sharma and R. Gupta, "Hybrid modeling in hydrology: Integrating machine learning and physical models for flood forecasting," *J. Hydroinform.*, vol. 23, no. 4, pp. 712–729, 2021.
- [10] A. Mosavi, P. Ozturk, and K.-W. Chau, "Flood prediction using machine learning models: Literature review," *Water*, vol. 10, no. 11, Art. no. 1536, 2018.
- [11] H. Tamiru and M. Wagari, "Flood inundation mapping using integrated HEC-RAS and Artificial Neural Network: A case study of Baro River Basin, Ethiopia," *Appl. Water Sci.*, vol. 12, no. 10, p. 234, 2022.
- [12] G. K. Wedajo, M. Ayana, and O. Megersa, "Data-driven feature selection for flood causative factors: A case study of the Awash River Basin, Ethiopia," *Environ. Earth Sci.*, vol. 83, no. 2, pp. 115–130, 2024.
- [13] N. Bhattacharya, R. K. Misra, and S. Das, "Impact of DEM resolution on flood inundation modeling in complex urban landscapes," *Int. J. Geoinformatics*, vol. 16, no. 2, pp. 45–58, 2020.
- [14] A. Horstmann and K. Wenzel, "UAV-based photogrammetry for high-fidelity DEM generation in local-scale hydraulic studies," in *Proc. Int. Conf. Geomatics Flood Manage.*, 2019.
- [15] A. Adhikari and S. Sharma, "Flood hazard mapping of the Bagmati River, Nepal, using HEC-RAS and GIS," *J. Hydrol. Sci.*, vol. 63, no. 11, pp. 1589–1602, 2018.
- [16] B. K. Pandey, R. Kumar, and L. P. Devkota, "Inundation mapping and hazard assessment of the Bagmati River Basin using HEC-RAS," *Nepal J. Sci. Technol.*, vol. 21, no. 1, pp. 88–97, 2020.
- [17] M. S. Shrestha, "Flood Frequency Analysis and risk assessment of the Bagmati River, Nepal," *J. Flood Risk Manage.*, vol. 12, no. S1, p. e12456, 2019.
- [18] P. K. Parhi, "HEC-RAS model for Manning's roughness: A case study," *Open J. Mod. Hydrol.*, vol. 3, no. 3, pp. 97–101, 2013.
- [19] S. K. McFeeters, "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features," *Int. J. Remote Sens.*, vol. 17, no. 7, pp. 1425–1432, 1996.
- [20] Y. Du, Y. Zhang, F. Ling, Q. Wang, W. Li, and X. Li, "Water bodies' mapping from Sentinel-2 imagery with modified Normalized Difference Water Index at 10-m spatial resolution produced by sharpening the SWIR band," *Remote Sens.*, vol. 8, no. 4, Art. no. 354, 2016.
- [21] D. Ouali and A. J. Cannon, "Estimation of rainfall intensity-duration-frequency curves at ungauged locations using quantile regression methods," *Stoch. Environ. Res. Risk Assess.*, vol. 32, no. 10, pp. 2821–2836, 2018.
- [22] J. Pérez-Sánchez, J. Senent-Aparicio, F. Segura-Méndez, D. Pulido-Velazquez, and R. Srinivasan, "Evaluating hydrological models for deriving water resources in peninsular Spain," *Sustainability*, vol. 11, no. 10, pp. 1–36, 2019.