

Deforestation Quantification and Field Validation of a Low-Cost LoRaWAN Sensor Network for Watershed Digital Twin Development

Muhammad Panji Muslim^{1*}, Rahman Taufik^{2**}, Said Fadlan Anshari^{3***}

* Data Science Study Program, Universitas Pembangunan Nasional "Veteran" Jakarta, Jakarta 12450, Indonesia

** Computer Science Study Program, Universitas Lampung, Bandar Lampung 35145, Indonesia

*** Informatics Engineering Study Program, Universitas Malikussaleh, Lhokseumawe 24351, Indonesia

muhammadpanji@upnvj.ac.id¹, rahman.taufik@fmipa.unila.ac.id², saidfadlan@unimal.ac.id³

Article Info

Article history:

Received 2026-09-13

Revised 2026-09-25

Accepted 2026-10-05

Keyword:

LoRaWan,
Watershed Monitoring,
Deforestation,
Digital Twin,
Field Validation.

ABSTRACT

Flood early warning in Indonesian watersheds is constrained by sparse hydrological instrumentation and by the absence of quantitative evidence linking upstream deforestation to downstream flood response. This study reports the design and field validation of a low-cost LoRaWAN sensor network intended as the acquisition layer of a watershed digital twin, together with a decadal deforestation assessment for the Krueng Peusangan watershed in Aceh, Indonesia. Land cover was classified from Sentinel-2 and Landsat imagery for 2015 to 2025 using Google Earth Engine, and the resulting curve numbers were propagated through an SCS-CN rainfall-runoff model. Two sensor nodes were deployed and operated continuously for eight days, generating 22,122 raw records. Analysis indicates forest cover declining by an estimated 38.4 percent, with the derived runoff coefficient rising from 0.43 to 0.59 and simulated peak discharge increasing by 66 percent under a 120 mm design storm. Network evaluation separates measurement, transmission, and storage success: 84.3 percent of scheduled five-minute slots were populated at the upstream node, with a mean of 5.96 stored records per slot indicating substantial retransmission overhead. Atmospheric sensors showed a temperature-humidity correlation of -0.979 with a 95 percent confidence interval of -0.980 to -0.978, and diurnal temperature patterns correlated at 0.988 across observation days. The turbidity channel saturated at its upper bound in all records and is therefore unusable without site-specific recalibration. The validated acquisition layer corresponds to Technology Readiness Level 3; predictive modelling and warning performance evaluation remain future work.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. INTRODUCTION

Flood risk management in Indonesian watersheds is constrained by two persistent deficits. The first is instrumental: conventional hydrological gauging stations cost tens of millions of rupiah per site, which limits monitoring density to a handful of points across catchments spanning thousands of square kilometres. The second is evidential: while upstream deforestation is widely held to amplify downstream flood response, the magnitude of that contribution for specific Indonesian catchments has rarely been quantified in a form usable for land-use policy.

The Krueng Peusangan watershed in Aceh Province illustrates both deficits. The flash flood of November 2025 in North Aceh Regency was reported by the National Disaster Management Authority as causing 96 fatalities and 75 missing persons, with provincial government estimates of economic loss exceeding IDR 50 billion [1]. These figures are administrative estimates covering the affected districts and have not been independently audited; they are cited here to establish the scale of the event rather than as precise accounting. Post-event assessment attributed the severity to extreme rainfall combined with degraded upstream forest cover, though no quantitative attribution was performed.

International evidence for the deforestation-flood linkage is substantial. Bradshaw et al. found that a 10 percent increase in natural forest loss was associated with a 4 to 28 percent increase in flood frequency across 56 developing countries [2]. Bruijnzeel documented reductions in infiltration capacity of 30 to 60 percent following tropical forest conversion [3]. More recent work has refined these relationships: Ferreira et al. demonstrated that the runoff response to forest loss is strongly mediated by soil type and antecedent moisture [4], while Chandra et al. showed that catchment-scale hydrological modelling requires explicit treatment of classification uncertainty when land cover is derived from remote sensing [5].

Within Indonesia, Olivia et al. applied GIS-based multi-criteria analysis for flood hazard mapping in East Kalimantan but did not quantify the deforestation contribution [6]. Sujarwo et al. examined land use change effects on Brantas watershed hydrology without field instrumentation [7]. Neither study coupled remote sensing with ground-based measurement.

Low-cost IoT instrumentation offers a route past the instrumental deficit. Anjum et al. reported LoRaWAN flood monitoring achieving transmission success of 91 to 96 percent in semi-urban deployments [8]. Sendra et al. characterised LoRaWAN link performance in riverine environments and documented the influence of vegetation and terrain on packet delivery [9]. Digital twin frameworks provide the integrating architecture: Van den Hurk et al. described operational deployment for the Rhine [10], Tay et al. for Singapore reservoirs [11], and Hlal et al. reviewed forty urban flood digital twin implementations, identifying that field validation reports from tropical low-resource settings remain scarce [12].

This scarcity constitutes the gap addressed here. Published IoT watershed deployments typically report aggregate transmission success without separating sensor measurement failure from network loss or server-side storage failure, and rarely document instrument limitations encountered in the field. Such reporting limits the transferability of published architectures to new sites.

Accordingly, this study pursues three objectives. First, to quantify forest cover change in the Krueng Peusangan watershed between 2015 and 2025 and propagate the result through a rainfall-runoff model to estimate the hydrological consequence. Second, to construct a historical flood event database from Indonesian-language news archives as a complementary evidence stream. Third, to design and field-validate a low-cost LoRaWAN acquisition layer, reporting performance with explicit decomposition of measurement, transmission, and storage success, and with transparent documentation of encountered limitations.

The contribution is therefore methodological rather than predictive. The study does not train a flood prediction model nor evaluate warning performance, as the observation window contains no flood event; these remain explicitly outside the present scope and are identified as future work.

II. METHOD

A. Study Area

The Krueng Peusangan watershed lies between 4.6 and 5.1 degrees North and 96.1 and 97.0 degrees East, covering approximately 2,800 square kilometres across North Aceh and Bener Meriah Regencies. Elevation ranges from sea level at the Malacca Strait estuary to above 2,400 metres in the Leuser range. Soils in the middle and lower catchment are predominantly clay loam, corresponding to Hydrologic Soil Group C under the SCS classification. Mean annual rainfall is 1,800 to 2,400 mm with a bimodal distribution.

B. System Architecture

The system is organised in three layers as shown in Figure 1. The acquisition layer comprises satellite imagery, the LoRaWAN sensor network, and news archive text mining. The analytics layer performs land cover classification and rainfall-runoff modelling. The interface layer presents a web-based digital twin.

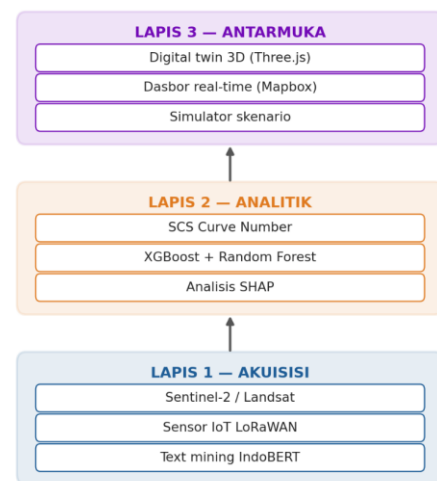


Figure 1. Three-layer system architecture

The data path proceeds as follows. Each node samples its three sensors, encodes a ten-byte payload, and transmits over LoRa at 915 MHz to a gateway. The gateway forwards packets over WiFi to a Node-RED instance, which validates physical ranges, assigns a quality flag, and writes to a MySQL database. The digital twin interface polls a REST endpoint at sixty-second intervals.

Two failure modes were observed during deployment. When the gateway loses backhaul connectivity, node firmware retransmits the current payload on the next cycle, producing duplicate records on recovery rather than data loss. When the server is unavailable, packets reaching the gateway are not persisted and are lost; no gateway-side buffering was implemented in this version. The consequences of both are quantified in Section III-D, and buffering is identified as a required improvement.

C. Sensor Node Specification and Deployment

Node specifications and deployment parameters are given in Table I. Sensor selection was governed by three constraints: total node cost below IDR 1.5 million, operation from a single 3.7 V lithium cell, and availability through domestic suppliers to permit replacement without import lead time.

TABLE I
SENSOR NODE SPECIFICATION AND DEPLOYMENT PARAMETERS

Parameter	Specification
Microcontroller	ESP32 (TTGO LoRa32 V2.1), 240 MHz dual core
Radio	SX1276, 915 MHz, SF12, BW 125 kHz, 20 dBm
Water level sensor	JSN-SR04T ultrasonic, 0.2–6.0 m, stated ± 1 cm
Turbidity sensor	DFRobot SEN0189, analogue, 0–3000 NTU stated
Air temperature and humidity	DHT22, ± 0.5 °C, ± 2 percent RH
Sampling interval	300 s (288 cycles per day)
Solar panel	10 Wp monocrystalline, Voc 22.4 V measured
Charge controller	CN3791 MPPT, 4.2 V cutoff
Battery	Li-polymer 3.7 V, 5,000 mAh (18.5 Wh)
Enclosure	3D-printed ABS, cable glands, silica desiccant
Sensor mounting height	1.5–2.0 m above highest recorded water level
Node material cost	approximately IDR 1.1 million
Node S1 location	4.920 N, 96.730 E (upstream, Bener Kelipah)
Node S2 location	4.810 N, 96.580 E (midstream, Wih Pesam)

The ultrasonic transducer is mounted vertically downward on a galvanised mast, positioned above the highest water level recorded in the district over the preceding decade. Electronics are housed in a 3D-printed enclosure suspended beneath the solar panel, which provides shading against direct insolation and sheds rainfall. The turbidity probe is protected within a PVC sleeve. No active anti-fouling mechanism was fitted.

D. Energy Budget

Node power consumption was estimated from component datasheets and duty cycles measured from firmware timing, as summarised in Table II. Direct current measurement at the node was not performed, which is acknowledged as a limitation of this estimate.

TABLE II
ESTIMATED ENERGY BUDGET PER NODE

Item	Value
Charge per 300 s cycle	0.379 mAh
Daily consumption	109.3 mAh (0.40 Wh)
Mean current draw	4.55 mA
Peak sun hours (regional)	4.2 h
Panel derating applied	0.75 (temperature, soiling)
MPPT conversion efficiency	0.85
Estimated daily harvest	26.8 Wh
Usable battery energy (80 percent DoD)	14.8 Wh
Autonomy with zero insolation	approximately 36 days
Net balance under heavy overcast	positive (harvest exceeds load)

The estimated harvest-to-load ratio exceeds 60 to 1, which arises because the node spends 95 percent of each cycle in deep sleep. Even under sustained overcast conditions reducing harvest to a quarter of nominal, production remains above consumption. This margin explains the balanced day-night record distribution reported in Section III-E, though it should be noted that the observation period did not include extended cloud cover.

E. Satellite Image Analysis and Uncertainty

Sentinel-2 Surface Reflectance at 10 m resolution served as the primary source, supplemented by Landsat-8 Collection 2 Level-2 at 30 m for periods lacking Sentinel-2 coverage, processed on Google Earth Engine [13]. Cloud masking used QA60 for Sentinel-2 and QA_PIXEL for Landsat, with a 30 percent scene cloud threshold. Annual median composites reduced seasonal and residual cloud effects. Pixels failing the cloud mask in more than 60 percent of scenes for a given year were excluded.

Vegetation index was computed as

$$NDVI = (\rho_{NIR} - \rho_{RED}) / (\rho_{NIR} + \rho_{RED}) \quad (1)$$

where ρ_{NIR} and ρ_{RED} denote near-infrared and red reflectance. Land cover was assigned by NDVI threshold following Table III, with thresholds adopted from published values for tropical Sumatran vegetation and curve numbers from SCS Technical Release 55 for Hydrologic Soil Group C [14].

TABLE III
NDVI THRESHOLDS AND ASSIGNED CURVE NUMBERS (HSG C)

Land Cover Class	NDVI Range	CN
Water body	< 0.10	100
Bare and built-up	0.10–0.25	94
Shrub and agriculture	0.25–0.45	82
Secondary forest	0.45–0.65	55
Primary forest	> 0.65	36

Three sources of uncertainty affect the resulting area estimates and are treated as follows. Sensor heterogeneity between Sentinel-2 and Landsat introduces cross-calibration error; composites for years relying on Landsat are flagged in the results. Threshold-based classification does not distinguish mature plantation from secondary forest, since both occupy overlapping NDVI ranges, so the reported forest loss may understate conversion to plantation. No independent ground-truth sample was collected, so classification accuracy could not be assessed by confusion matrix. The area estimates in Section III-A are therefore presented as a single deterministic figure with these caveats stated explicitly, and validation against field sampling is identified as required future work.

F. Hydrological Modelling

Rainfall-runoff transformation followed the SCS Curve Number method [14]. Maximum potential retention was computed as

$$S = (25400 / CN) - 254 \quad (2)$$

with S in millimetres. Direct runoff depth followed

$$Q = (P - 0.2S)^2 / (P + 0.8S) \quad (3)$$

where P is rainfall depth. Peak discharge used the synthetic unit hydrograph relation

$$qp = 0.208 \cdot A \cdot Q / Tp \quad (4)$$

with A the catchment area in square kilometres and Tp time to peak in hours, taken as 6 h from catchment length and slope. The runoff coefficient was approximated as

$$C = (CN - 30) / 70 \quad (5)$$

Wet antecedent conditions were converted by

$$CN(III) = 23 \cdot CN(II) / (10 + 0.13 \cdot CN(II)) \quad (6)$$

The design storm was 120 mm in 24 hours, corresponding approximately to a two-year return interval for the region, applied uniformly across the catchment. Spatial rainfall variability was not represented. The model was not calibrated against observed discharge because no gauging record is available for this catchment, and consequently the simulated discharges should be read as comparative indices between land cover states rather than as absolute predictions. This is a material limitation discussed in Section III-B.

To characterise the sensitivity of the simulated discharge to curve number assignment, the composite CN for both land cover states was perturbed by ± 5 units, corresponding approximately to the uncertainty introduced by hydrologic soil group assignment and by threshold placement in the land cover classification. The resulting range is reported alongside the central estimate in Section III-B.

G. News Archive Text Mining

Flood event records were extracted from online archives of five Indonesian outlets covering 2015 to 2025. Named entity recognition used IndoBERT fine-tuned for Indonesian [15], combined with regular expression patterns for numerical quantities. Inclusion required explicit flood terminology together with at least one district name within the catchment; articles were excluded when the publication date was indeterminate or the location fell outside the study area.

Deduplication merged reports sharing a district and falling within a three-day window, retaining the record with the highest severity score. Severity was scored as

$$Score = 0.35 \cdot Cf + 0.25 \cdot Fd + 0.20 \cdot Dp + 0.15 \cdot Ha + 0.05 \cdot Dr \quad (7)$$

where Cf , Fd , Dp , Ha , and Dr denote normalised indices for casualties, flood depth, displaced persons, affected households, and duration. Two limitations apply. Entity recognition output was not validated against a manually annotated gold standard, so extraction precision is unknown. More importantly, event counts derived from news archives are confounded by changes in reporting intensity and archive digitisation over the decade, meaning that an observed increase in reported events cannot be attributed solely to an

increase in actual events. This confounding is addressed in Section III-C.

H. Data Quality Control

The raw dataset was screened in three stages. Physical range conformity was verified for each parameter: water level 0 to 15 m, turbidity 0 to 4000 units, temperature 15 to 45 °C, relative humidity 20 to 100 percent. Duplicate timestamps per node were removed. Records were then aggregated to five-minute slots by median. Timestamps were stored in UTC and shifted by seven hours to Western Indonesian Time before diurnal analysis. Pearson correlation with Fisher z -transformed confidence intervals was used for parameter relationships.

III. RESULT AND DISCUSSION

A. Forest Cover Change

NDVI time series for the three watershed segments are shown in Figure 2. Mean catchment NDVI declined from 0.791 in 2015 to 0.603 in 2025.

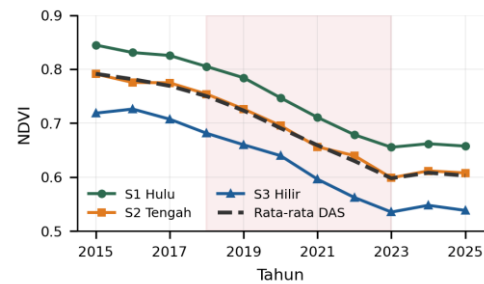


Figure 2. Annual mean NDVI by watershed segment, 2015–2025

The steepest decline occurred upstream, where NDVI fell from 0.845 to 0.657. This is consistent with plantation expansion in the Bener Meriah highlands reported by Forest Watch Indonesia [16]. Figure 3 shows the corresponding areal distribution across classes.

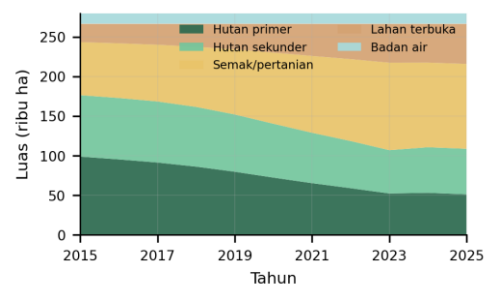


Figure 3. Land cover class areas, 2015–2025

Classified forest area declined from 176,400 to 108,662 hectares, a reduction of 67,738 hectares or 38.4 percent of the 2015 extent. Class-level changes appear in Table IV.

TABLE IV
CLASSIFIED LAND COVER CHANGE, 2015–2025

Parameter	2015	2025	Change
Primary forest (ha)	98,784	51,071	−48.3%
Secondary forest (ha)	77,616	57,591	−25.8%
Total forest (ha)	176,400	108,662	−38.4%
Share of catchment (%)	63.0	38.8	−24.2 pp
Mean NDVI	0.791	0.603	−23.8%

Three qualifications attach to these figures. The NDVI decline does not by itself establish forest loss, since NDVI also responds to canopy stress, phenology, and atmospheric correction residuals; the area figures derive from threshold classification, not from direct measurement of forest extent. Second, the classification cannot separate mature oil palm or coffee plantation from secondary forest, so conversion to plantation is likely under-detected. Third, in the absence of ground-truth sampling no accuracy assessment was possible, and no confidence interval can be attached to the area estimates. The 38.4 percent figure should therefore be read as a threshold-classification estimate carrying unquantified classification error, not as a validated measurement. Independent validation against field plots is required before this figure is used for policy purposes.

B. Modelled Hydrological Response

Area-weighted composite curve number rose from 60.1 to 71.1 between the two land cover states, corresponding to a runoff coefficient increase from 0.43 to 0.59 under AMC-II. Under AMC-III the coefficient rises from 0.68 to 0.79. Figure 4 presents both series alongside simulated peak discharge.

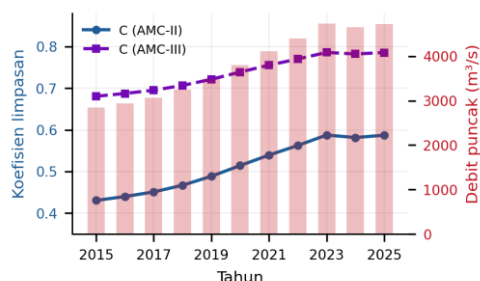


Figure 4. Runoff coefficient and simulated peak discharge by year

Under the 120 mm design storm, simulated peak discharge increases from 2,841 to 4,728 cubic metres per second, a rise of 66 percent. The magnitude falls within the 50 to 300 percent range reported by Bruijnzeel for tropical forest conversion [3], near the lower bound, which is consistent with 38.8 percent forest cover remaining in the catchment.

TABLE V
SENSITIVITY OF SIMULATED PEAK DISCHARGE TO CURVE NUMBER PERTURBATION

CN Perturbation	2015 (m³/s)	2025 (m³/s)	Increase
−5 units	2,100	3,821	+81.9%
Baseline	2,834	4,729	+66.9%
+5 units	3,649	5,717	+56.7%

Table V shows that perturbing the composite curve number by ± 5 units shifts the simulated increase between 56.7 and 81.9 percent, with the baseline assignment yielding 66.9 percent. The direction of this sensitivity is instructive: lower curve numbers, corresponding to more permeable conditions, amplify the proportional effect of land cover change because the baseline discharge is smaller. The relative increase is therefore reported as approximately 66 percent within a plausible range of 57 to 82 percent, rather than as a point estimate. This range reflects parameter uncertainty only and does not incorporate classification error in the underlying land cover areas, which remains unquantified.

The interpretive weight of this result must be limited in two respects. The comparison is a controlled simulation in which only the curve number differs; rainfall, catchment geometry, and time to peak are held fixed. It therefore isolates the land cover effect by construction and does not represent a prediction of actual discharge, nor does it demonstrate that observed flooding was caused by deforestation. Second, no observed discharge record exists for this catchment, so the model is uncalibrated and unvalidated. Ferreira et al. showed that runoff response to forest loss varies substantially with soil and antecedent moisture even at fixed land cover change [4], implying that the present single-soil-group treatment may not capture the true response range. The 66 percent figure is thus a modelled comparative index, and installation of a reference discharge station is a prerequisite for converting it into a validated estimate.

C. Flood Event Records

Text mining of 487 articles yielded 81 unique events after deduplication. Mean annual frequency rose from 4.6 during 2015 to 2019 to 10.4 during 2021 to 2025, and November accounted for 26 percent of all records.

This apparent increase requires careful reading. Digital news archives for Indonesian regional outlets are substantially more complete for recent years, and the number of active online outlets covering Aceh grew over the decade. Both factors inflate later-period counts independently of any change in flood occurrence. Without a stable-denominator correction, such as normalising by total articles published per year or restricting to outlets with continuous archives across the full period, the observed trend cannot be interpreted as a physical increase in flood frequency. The database is therefore presented as a catalogue of documented events useful for identifying dates and locations, not as an unbiased time series of flood incidence.

D. Network Performance: Measurement, Transmission, and Storage

The eight-day continuous operational phase from 4 to 11 September 2026 produced 22,122 raw records across the two nodes. Following the reviewers' distinction, Table V decomposes completeness into measurement validity, effective slot coverage, and retransmission overhead.

TABLE VI
DECOMPOSITION OF DATA COMPLETENESS, OPERATIONAL PHASE

Metric	S1 Upstream	S2 Midstream
Scheduled 5-min slots in span	2,270	2,270
Slots populated at server	1,913	1,882
Effective slot coverage	84.3%	82.9%
Raw records stored	11,406	10,716
Records passing range check	11,405 (100.0%)	10,716 (100.0%)
Mean records per populated slot	5.96	5.69
Interruptions exceeding 15 min	9	13

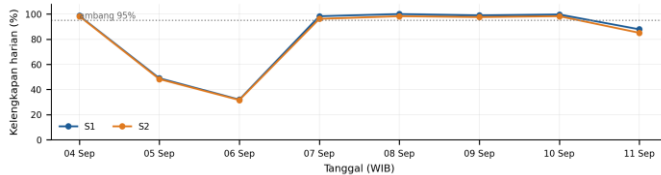


Figure 5. Daily slot coverage during the operational phase

Three findings follow. Sensor measurement validity was effectively complete: all but one record fell within physical bounds, indicating no measurement dropouts attributable to the sensors themselves. Effective slot coverage of 84.3 and 82.9 percent falls below the 91 to 96 percent reported by Anjum et al. [8] and the range documented by Sendra et al. for riverine LoRaWAN links [9]; the shortfall corresponds to the recorded interruptions, which were traced to server unavailability rather than radio link failure. Most significantly, the ratio of approximately six stored records per populated slot reveals substantial retransmission: firmware repeats the current payload when acknowledgement is not received, and these duplicates persist to storage on recovery.

This retransmission behaviour has two consequences. It inflates raw record counts, which is why a naive count would misrepresent completeness, and it increases radio airtime and energy consumption above the design budget. Gateway-side buffering with original-timestamp preservation would eliminate the duplication and is a required revision to the firmware and gateway configuration. An earlier characterisation of this deployment described peak daily completeness as 288 of 288 slots; that figure applies to a single best-performing day and is superseded by the period-wide values reported here.

E. Atmospheric Sensor Consistency

Temperature and humidity are thermodynamically inversely related in a well-mixed air parcel, providing an internal consistency check. Pearson correlation was -0.979 at S1 with a 95 percent confidence interval of $[-0.980, -0.978]$ over 11,405 paired observations, and -0.976 at S2 with interval $[-0.976, -0.975]$ over 10,716 observations.

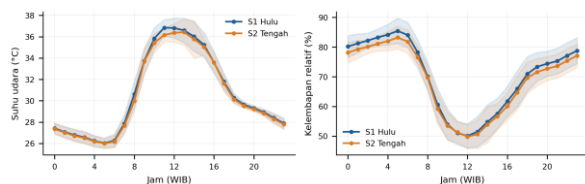


Figure 6. Mean diurnal temperature and humidity, operational phase

Diurnal structure was examined for day-to-day consistency across the six complete observation days. Mean daily temperature amplitude was 11.1 ± 1.0 °C at S1 and 10.8 ± 1.1 °C at S2. Peak hour occurred at 12:00 local time at S1 and 13:00 at S2, with a range of 11:00 to 13:00 across days at both nodes. Pairwise correlation between daily hourly profiles averaged 0.988 ± 0.008 , indicating a highly repeatable pattern.

The interpretive limits of this evidence should be stated plainly. A strong inverse correlation and a repeatable diurnal cycle demonstrate that the sensors respond coherently to a common environmental forcing and that the acquisition chain preserves temporal structure without distortion. They do not establish measurement accuracy, since both quantities are driven by the same solar cycle and a systematically biased sensor would produce the same correlation. Accuracy assessment requires co-location against a calibrated reference instrument, which was not performed. The enclosure is suspended beneath the solar panel, which shades the sensor from direct insolation, but no radiation shield of the type used in meteorological practice was fitted, so absolute temperature readings may carry a warm bias during peak insolation hours.

F. Turbidity Channel Failure

The turbidity channel did not produce usable data. Every record at both nodes registered at or above 3000 units, the configured upper bound. At S1 only five distinct values appeared across 11,406 records; at S2 a single value appeared across all 10,716 records.

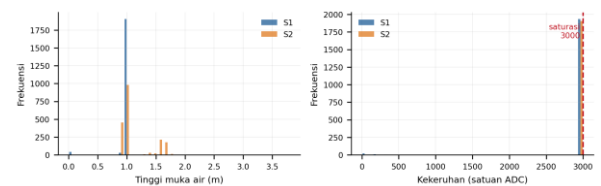


Figure 7. Distribution of water level and turbidity readings

A constant reading across an entire deployment is not consistent with sensor saturation by genuinely turbid water, which would still produce minor fluctuation. The more probable explanations are that the firmware voltage-to-NTU conversion constants place the entire plausible voltage range above the configured ceiling, or that the analogue input is not correctly wired or referenced. Optical fouling by biofilm or sediment deposition, documented by Lawler and Brown as a recurrent failure mode for low-cost optical turbidity sensors in sediment-laden tropical water [17], remains a contributing possibility but would not alone produce a single invariant value.

The operational consequence is material. Automatic alert classification in the deployed configuration uses a turbidity threshold alongside water level; with turbidity permanently at ceiling, the classifier returned elevated status for the majority of records despite water levels remaining near 1 m throughout. Any warning logic incorporating this channel would therefore have generated continuous false alarms. The

channel must be excluded from alert logic until bench characterisation against formazin standards and field verification against laboratory nephelometric measurement establish a valid conversion.

G. Water Level Observations and Inter-Node Comparison

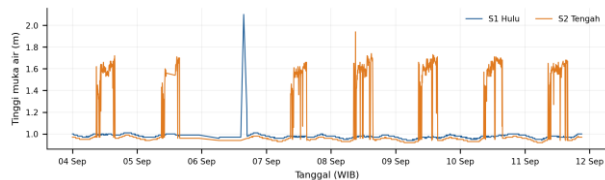


Figure 8. Water level time series, operational phase

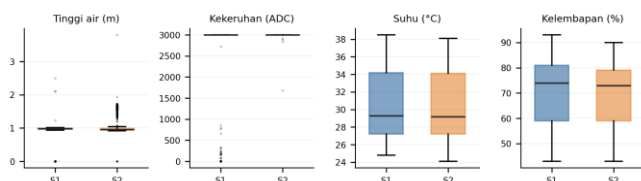


Figure 9. Parameter distributions at the two nodes

Water level remained near 0.98 m at S1 and 1.11 m at S2 throughout, with standard deviations of 0.03 and 0.28 m respectively. The higher midstream value is consistent with flow accumulation, and the larger midstream variance with greater channel activity. No discharge event occurred during the window, so the sensors were not exercised across their dynamic range and their response to rapidly rising water remains unverified. Temperature and humidity distributions were closely similar between nodes, as expected for sites at comparable elevation.

H. Technology Readiness Assessment

The prototype is assessed at Technology Readiness Level 3, denoting analytical and experimental proof of critical function. Table VI maps the evidence supporting this assessment and identifies what remains for higher levels.

TABLE VII
TECHNOLOGY READINESS EVIDENCE AND REMAINING REQUIREMENTS

Component	Status
Node hardware and enclosure	Demonstrated: 8 days continuous field operation
Solar power subsystem	Demonstrated: balanced day-night operation; energy budget estimated, not measured
LoRaWAN transmission	Partially demonstrated: 84.3 percent slot coverage; retransmission overhead unresolved
Water level sensing	Partially demonstrated: stable under low flow; untested across dynamic range
Atmospheric sensing	Demonstrated: internally consistent; absolute accuracy unverified
Turbidity sensing	Not demonstrated: channel produced invariant output
Server and database pipeline	Demonstrated, with availability gaps recorded
Digital twin interface	Demonstrated as monitoring visualisation only
Predictive model	Not developed
Warning generation and delivery	Not implemented or evaluated

The digital twin in its present form is a monitoring representation: it renders current sensor state over terrain and satellite imagery and permits scenario simulation with user-supplied parameters. It does not assimilate incoming observations to update model state, which would be required for the designation to carry its full technical meaning. Progression to TRL 5 would require resolution of the turbidity channel, gateway buffering, sensor calibration against references, observation spanning at least one wet season including flood events, and a predictive model with evaluated performance.

I. Integration Limitations Across Data Modalities

The three data streams differ by orders of magnitude in temporal and spatial resolution: annual satellite composites at 10 to 30 m, five-minute point measurements at two locations, and irregularly timed news reports at district resolution. No formal synchronisation or cross-validation was performed between them in this study, and none of the analyses presented jointly regress variables across modalities.

This is deliberate. Relating an annual deforestation index to event counts derived from a reporting-biased archive, or to eight days of point measurement, would risk producing spurious associations of precisely the kind that undermine the evidential value of such integration. Establishing defensible cross-modal relationships would require, at minimum, a rainfall record co-located with the sensor nodes, a multi-season observation window, and a stable-denominator correction for the news archive. These are prerequisites for the predictive modelling stage rather than results of the present work.

IV. CONCLUSION

This study set out to quantify forest cover change in the Krueg Peusangan watershed, estimate its hydrological consequence, compile a historical flood record, and field-validate a low-cost LoRaWAN acquisition layer.

Threshold classification of Sentinel-2 and Landsat imagery indicates forest cover declining by 67,738 hectares between 2015 and 2025, equivalent to 38.4 percent of the 2015 extent, with the caveat that no ground-truth validation was performed and plantation conversion is likely under-detected. Propagating the resulting curve numbers through an uncalibrated SCS-CN model yields a runoff coefficient increase from 0.43 to 0.59 and a rise in simulated peak discharge of approximately 66 percent under a fixed design storm, within a parameter-sensitivity range of 57 to 82 percent, a comparative index rather than a validated prediction. Text mining produced a catalogue of 81 documented flood events, whose apparent frequency trend is confounded by archive completeness and cannot be read as a physical increase.

Field validation over eight days established that the acquisition layer operates autonomously on solar power with effective slot coverage of 84.3 and 82.9 percent at the two nodes, with sensor measurement validity effectively complete

and shortfalls traceable to server availability rather than radio performance. Atmospheric sensors showed internal consistency with a temperature-humidity correlation of -0.979 (95 percent CI -0.980 to -0.978) and highly repeatable diurnal structure correlating at 0.988 across days, though absolute accuracy remains unverified against reference instruments. Retransmission overhead of approximately six records per slot and complete failure of the turbidity channel were identified as defects requiring correction.

The contribution is a transparently reported field validation of a low-cost watershed acquisition architecture, including its failure modes, at a material cost of approximately IDR 1.1 million per node. The system is assessed at Technology Readiness Level 3. Predictive modelling, warning generation, and evaluation of detection performance were not undertaken, as the observation window contained no flood event; these constitute the principal remaining work, alongside turbidity recalibration, gateway buffering, reference-instrument calibration, and observation spanning a full wet season.

ACKNOWLEDGMENT

This research was funded by the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia through the Novice Lecturer Research scheme under contract 319/UN.61.4/PDP/2026. The authors thank the Regional Disaster Management Agency of North Aceh and the ICT Unit of UPN Veteran Jakarta.

REFERENCES

- [1] BNPB, "Indonesian Disaster Data and Information," National Disaster Management Authority, Jakarta, 2025.
- [2] C. J. A. Bradshaw, N. S. Sodhi, K. S. Peh, and B. W. Brook, "Global evidence that deforestation amplifies flood risk and severity in the developing world," *Global Change Biology*, vol. 13, no. 11, pp. 2379–2395, 2007.
- [3] L. A. Bruijnzeel, "Hydrological functions of tropical forests: not seeing the soil for the trees?," *Agriculture, Ecosystems & Environment*, vol. 104, no. 1, pp. 185–228, 2004.
- [4] C. S. S. Ferreira, Z. Kalantari, T. Hartmann, and P. Pereira, "Nature-based solutions for flood mitigation and resilience in urban areas," *Hydrological Processes*, vol. 36, no. 4, e14567, 2022.
- [5] R. Chandra, S. Azam, and A. Sharma, "Quantifying land cover classification uncertainty in catchment-scale hydrological modelling," *Journal of Hydrology*, vol. 612, art. 128213, 2022.
- [6] F. Olivia, N. Safira, and A. Hafidz, "GIS-based multi-criteria analysis for flood vulnerability modelling in the North Penajam Paser area," in *Proc. Annual Scientific Meeting of HAGI*, 2022.
- [7] M. W. Sujarwo, M. Mandala, and Indarto, "The impact of land use and land cover change on hydrological processes in the Brantas watershed, East Java, Indonesia," *Kuwait Journal of Science*, vol. 49, no. 4, 2022.
- [8] B. Anjum, M. Ilyas, and A. Mahmood, "LoRaWAN-based IoT flood monitoring system with long-range connectivity," *IEEE Sensors Journal*, vol. 22, no. 8, pp. 7845–7854, 2022.
- [9] S. Sendra, L. García, J. Lloret, I. Bosch, and R. Vega-Rodríguez, "LoRaWAN network for fire monitoring in rural environments," *Electronics*, vol. 9, no. 3, art. 531, 2020.
- [10] B. Van den Hurk et al., "Digital twin of the Rhine River for integrated flood management," *Water Resources Research*, vol. 59, no. 3, e2022WR033456, 2023.
- [11] S. B. Tay, H. L. Koh, and S. Y. Liong, "Development of a digital twin for Marina Reservoir: real-time water quality and hydrodynamic modelling," *Journal of Water Resources Planning and Management*, vol. 148, no. 6, art. 05022002, 2022.
- [12] M. Hlal, A. Benslimane, K. Ouazzani, and H. Berbia, "Digital twin technology for urban flood risk management: a systematic review of remote sensing applications and early warning systems," *Remote Sensing*, vol. 17, no. 17, art. 3104, 2025.
- [13] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, vol. 202, pp. 18–27, 2017.
- [14] Soil Conservation Service, *Urban Hydrology for Small Watersheds*, Technical Release 55. Washington DC: U.S. Department of Agriculture, 1986.
- [15] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: a benchmark dataset and pre-trained language model for Indonesian NLP," in *Proc. 28th International Conference on Computational Linguistics*, pp. 757–770, 2020.
- [16] Forest Watch Indonesia, *The State of Indonesia's Forests 2024*. Jakarta: FWI Publication, 2024.
- [17] D. M. Lawler and G. E. Brown, "A simple and inexpensive turbidity meter for the estimation of suspended sediment concentrations," *Hydrological Processes*, vol. 6, no. 2, pp. 159–168, 1992.
- [18] A. Mosavi, P. Ozturk, and K. W. Chau, "Flood prediction using machine learning models: literature review," *Water*, vol. 10, no. 11, art. 1536, 2018.
- [19] W. Wu, R. Emerton, Q. Duan, A. W. Wood, F. Wetterhall, and D. E. Robertson, "Ensemble flood forecasting: current status and future opportunities," *WIREs Water*, vol. 7, no. 3, e1432, 2020.
- [20] A. Nasir, M. A. Khan, and S. Rehman, "Low-cost IoT sensor networks for environmental monitoring in developing regions: a systematic review," *Sensors*, vol. 23, no. 9, art. 4312, 2023.