FORUM GEOGRAFI
ISSN: 0852-0682 | E-ISSN: 2460-3945
HTTPS://JOURNALS.UMS.AC.ID/INDEX.PHP/FG/

Research article

Landscape Associations with Flood Response to Extreme Rainfall: A Remote-Sensing, Multi-Event Analysis Across Sumatra, Indonesia

Kuswaji Dwi Priyono[1,*], Siti Azizah Susilawati[2], Dewi Novita Sari[3], M. Iqbal Sunariya[1], Yuli Priyana[1], Andhika Argya Pragata[1], Arif Rohman[3], Muhammad Yusuf[4], Mohd Hairy Ibrahim[5], Farha Sattar[6], Muhammad Nawaz[7]

[1] Department of Geography, Faculty of Geography, Universitas Muhammadiyah Surakarta, Indonesia; [2] Department of Geography Education, Faculty of Teacher Education and Educational Sciences, Muhammadiyah University of Surakarta, Indonesia; [3] Faculty of Environment, University of Leeds, United Kingdom; [4] Interdisciplinary Research Center for Industrial Nuclear Energy (IRC-INE), King Fahd University of Petroleum and Minerals (KFUPM); [5] Faculty of Humanity, Universiti Sultan Idris, Tanjung Malim, Malaysia; [6] Faculty of Arts and Society, Charles Darwin University, Australia; [7] Department of Geography, Faculty of Art and Social Sciences, National University of Singapore, Singapore.

[*)] Correspondence: Kuswaji.Priyono@ums.ac.id

Citation:
Priyono, K. D., Susilawati, S. A., Sari, D. N., Sunariya, M. I., Priyana, Y., Pragata, A. A., Rohman, A., Yusuf, M., Ibrahim, M. H., Sattara, F., & Nawaz, M. (2027). Landscape Associations with Flood Response to Extreme Rainfall: A Remote-Sensing, Multi-Event Analysis Across Sumatra, Indonesia. Forum Geografi. 41(1), 91-114.

Article history:
Received: 6 June 2026
Revised: 6 September 2026
Accepted: 13 September 2026
Published: 14 September 2026

Abstract

Tropical flooding emerges from interactions among rainfall forcing, drainage structure, land-surface condition, and evolving land cover, yet these components are often analyzed in separate remote-sensing workflows. This study develops a multi-event, sub-basin-scale association framework linking extreme-rainfall coverage, vegetation cover, built-up cover, and Sentinel-1-derived flood response across HydroBASINS Level 9 units in Sumatra, Indonesia. Nine annual area-mean daily rainfall maxima from 2017–2025 were selected using CHIRPS. Vegetation fraction was defined as the proportion of valid Sentinel-2 pixels with NDVI $\geq$ 0.40, built-up fraction as the proportion with NDBI $\geq$ 0.00, flood fraction as the proportion of Sentinel-1 pixels with Random-Forest flood probability $\geq$ 0.50, and extreme fraction as the proportion of CHIRPS pixels whose 7-day accumulated rainfall exceeded the local monthly 95th percentile. The full panel contained 14,409 polygon-event observations, of which 11,818 were complete across the four core parameters. Fractional logit models with event fixed effects showed a negative association between vegetation fraction and flood fraction ($\beta = -0.073$, $p = 0.004$) and a positive association for built-up fraction ($\beta = 0.082$, $p < 0.001$), whereas extreme-rainfall coverage showed no direct association after event-level heterogeneity was absorbed. The corresponding landscape effects were modest in magnitude and should be interpreted as associations rather than causal effects. Cross-scale analysis at HydroBASINS Level 8 preserved the direction of the vegetation and built-up effects. The contribution therefore lies not in a new satellite sensor or classifier, but in combining multi-sensor fractions, hydrologically aligned sub-basin units, and repeated-event statistical inference. Because the Sentinel-1 flood product is pseudo-label based and was not independently validated for all nine events, flood_fraction is treated as a SAR-derived flood indicator rather than a fully validated inundation product, and no operational warning threshold is inferred.

Keywords: tropical flooding; HydroBASINS; CHIRPS; Sentinel-1 SAR; Sentinel-2; extreme rainfall; NDVI; NDBI; Sumatra; flood fraction; fractional logit; multi-event analysis.

1. Introduction

Floods remain among the most damaging hydrometeorological hazards worldwide, and satellite observations show that exposure has increased as settlements and economic assets expand into flood-prone areas (Tellman *et al.,* 2021; Jumadi *et al.,* 2024). Climate-driven intensification of the hydrological cycle further increases the probability of heavy rainfall and compound inundation, but rainfall magnitude alone does not determine where floodwater accumulates. Runoff generation, storage, topography, drainage connectivity, surface permeability, and land cover jointly shape the translation from atmospheric forcing to inundation (Dharmarathne *et al.,* 2024; Huning *et al.,* 2024; Pizzorni *et al.,* 2024). This distinction is especially important in humid tropical regions, where intense convection and rapid land-cover change coexist with sparse hydrometeorological observations.

Remote sensing provides an increasingly important basis for analyzing these coupled processes. Sentinel-1 SAR supports flood detection under cloud and at night, whereas Sentinel-2 provides spatial indicators of vegetation and built-up surfaces; CHIRPS supplies a long and spatially continuous rainfall record (Funk *et al.,* 2015; Chen *et al.,* 2024; Riazi *et al.,* 2023). These data sources are well established individually, and their use is not itself novel. Their limitations are also well documented: SAR flood signals vary with acquisition timing, flooded vegetation, urban backscatter, radar geometry, and shallow-water conditions, while optical indices remain sensitive to cloud and compositing choices (San Jose *et al.,* 2026; Tavus *et al.,* 2022).

The more specific gap addressed here concerns how these established observations are brought into a common inferential framework. Many flood studies end with an inundation, susceptibility, or hazard map; others use regular raster cells or administrative units that facilitate data integration but are not designed to follow drainage structure. Fewer studies evaluate repeated flood events using proportional measures of rainfall coverage, landscape composition, and inundation within the same hydrological units, and then test their associations statistically across events. The present study therefore does not claim a new CHIRPS–Sentinel or machine-learning combination. Its contribution is narrower: a repeated-event, fraction-based analysis in hydrologically aligned sub-basins, with cross-scale evaluation of whether the inferred associations persist when the aggregation level changes.

HydroBASINS provides a globally consistent hierarchy of sub-basins derived from hydrography and network routing (Lehner & Grill, 2013). We use Level 9 polygons as the primary analytical units because they retain substantially more hydrological structure than arbitrary raster partitions while remaining sufficiently detailed for multi-sensor zonal aggregation. This choice should not be interpreted as evidence that HydroBASINS is universally superior to regular grids or administrative units; the present study does not perform such a benchmark. Instead, HydroBASINS is used as a process-aligned spatial support, and the analysis is repeated at Level 8 to test sensitivity to hydrological aggregation scale and the Modifiable Areal Unit Problem.

Landscape composition is expected to modify runoff response, although its effects are context dependent. Vegetation may increase interception, infiltration, roughness, and storage, whereas built-up surfaces can increase imperviousness and accelerate runoff. In tropical lowlands, however, high NDVI can also represent riparian vegetation, wetlands, plantations, and naturally flood-prone areas, while built-up effects depend on topographic position and drainage capacity (Lin *et al.,* 2026; Pal *et al.,* 2022; Hoang & Liou, 2024; Wang *et al.,* 2025; Maketa *et al.,* 2026). Consequently, simple bivariate correlations between rainfall, vegetation, urbanization, and flooding may be weak even when these processes are hydrologically relevant.

Sumatra provides a demanding setting for evaluating these relationships. The island combines strong rainfall variability, a mountainous western spine, broad eastern lowlands, peatlands, major river systems, plantations, forest change, and expanding urban corridors (Aso *et al.,* 2024; Kartika *et al.,* 2022; Thoha *et al.,* 2024; Candraningrum, 2026). Extreme rainfall is further modulated by regional and large-scale atmospheric variability, including the MJO, ENSO, and Indian Ocean processes (Chrysanti & Son, 2025; Saufina *et al.,* 2025; Syamsudin *et al.,* 2026). These interacting controls make Sumatra appropriate for testing whether local extreme-rainfall coverage is sufficient to explain sub-basin-scale flood response once event heterogeneity and landscape composition are considered.

Accordingly, this study evaluates associations among extreme_fraction, vegetation_fraction, builtup_fraction, and flood_fraction across nine annual extreme-rainfall events from 2017 to 2025. The objectives are to (1) characterize the spatial and inter-event distributions of the four fractions; (2) quantify bivariate and multivariate associations with event fixed effects; (3) assess the practical magnitude and robustness of the landscape coefficients; and (4) test whether the main associations persist at HydroBASINS Level 8. The nine-event design is an event-conditioned explanatory analysis rather than a comprehensive flood-frequency or early-warning model. It therefore does not estimate a population warning threshold, and conclusions are restricted to the variables and events evaluated.

2. Methods

The methods are structured around a final research design based on hydrological units. The unit of analysis is the HydroBASINS Level 9 polygon, derived from the HydroSHEDS framework. HydroBASINS provides globally consistent basin and sub-basin delineations developed from hydrography and river-network routing, well suited for analyzing river systems and hydrological response across scales (Lehner & Grill, 2013).

Analytically, the study integrates CHIRPS daily rainfall, Sentinel-1 SAR, Sentinel-2 Surface Reflectance, HydroBASINS Level 9, and multi-event panel statistical modeling. Each observation represents a combination of a sub-basin polygon and an annual extreme rainfall event. Four principal variables are constructed as fractional parameters on the [0, 1] scale: vegetation fraction, built-up fraction, flood fraction, and extreme fraction. The construction principle is that high-resolution rasters are first converted into binary 0/1 rasters and then summarized via zonal mean over HydroBASINS Level 9 polygons. Because the input values are binary, the zonal mean is mathematically equivalent to the proportion of valid pixels that satisfy the relevant criterion within each sub-basin.

2.1. Study Area and Analytical Design

The study area covers Sumatra, Indonesia (Figure 1), a humid tropical region with varied topography, large river systems, lowland and coastal zones, peatlands, forests, plantations, and rapidly growing settlements. These conditions make Sumatra a suitable setting to evaluate how daily extreme rainfall, vegetation, built-up surfaces, and flood inundation are jointly associated within sub-basin units.

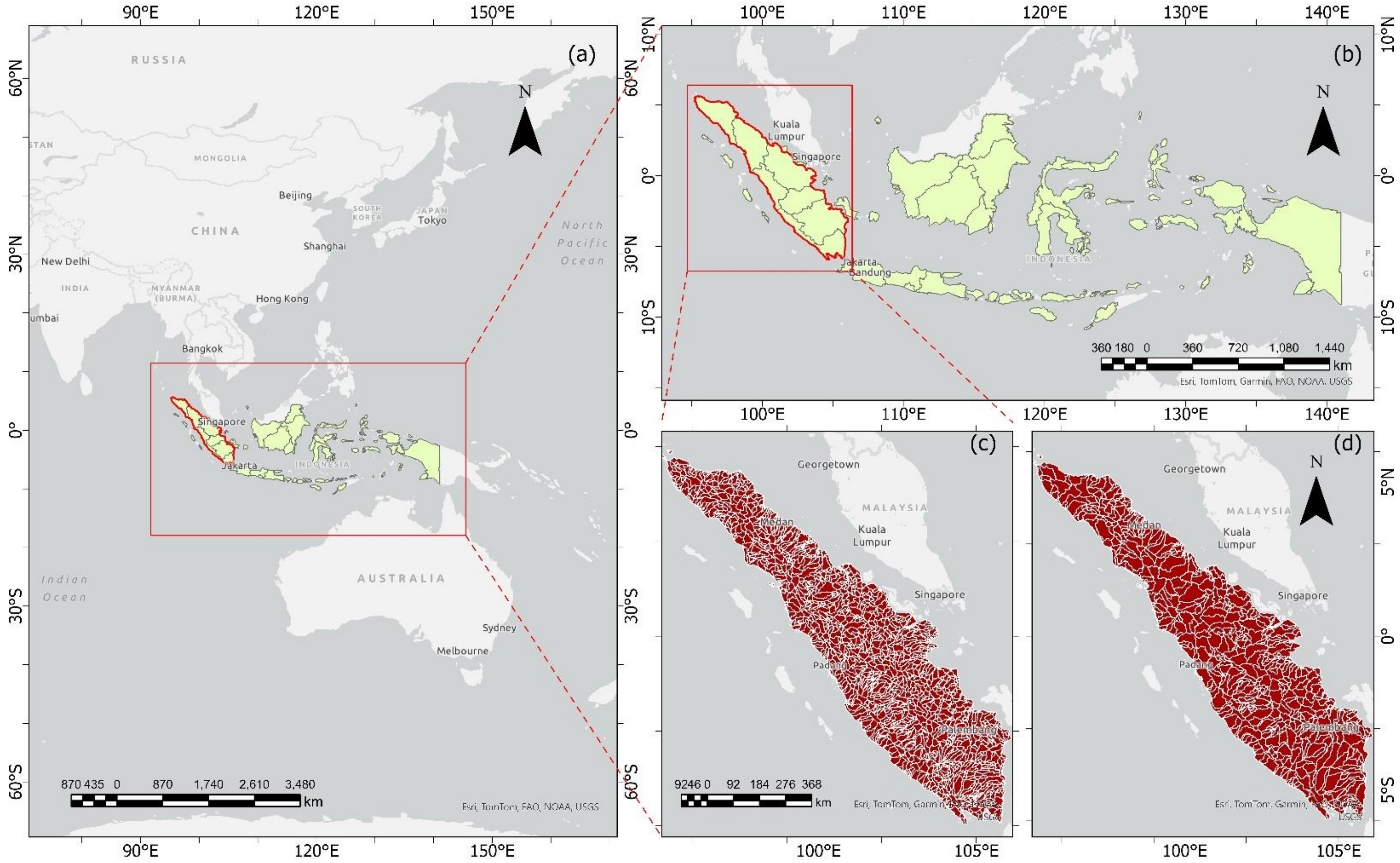

Figure 1. The Study Area. (a) Relative position of the study area between Asia and Australia, (b) Indonesia, (c) HydroBASINS Level 9, (d) HydroBASINS Level 8.

The design is multi-event, with one annual area-mean daily rainfall maximum selected for each year from 2017 to 2025. This rule preserves interannual coverage and prevents a small number of wet seasons from dominating the sample, but it also conditions the analysis on major rainfall days. Consequently, the nine selected dates do not represent the full distribution of flood-generating rainfall conditions and are insufficient for calibrating an operational warning threshold. Each of the 1,601 HydroBASINS Level 9 polygons is evaluated for every selected event, yielding 14,409 polygon-event records before missing-value screening.

We implement a hybrid computational workflow. Google Colab Pro (high-RAM) with Earth Engine API is used to access and process CHIRPS, Sentinel-1, Sentinel-2, DEM, and permanent water layers at the regional scale (Table 1). ArcGIS Pro is used for raster-tile mosaicking and Zonal Statistics as Table, because these steps are more stable for large rasters summarized over HydroBASINS polygons. Python in Google Colab Pro is used for CSV harmonization, completeness checks, descriptive statistics, correlation analysis, Generalized Linear Models, visualization, and tabulation. This division of platforms preserves computational efficiency while keeping each stage of the analysis reproducible. The end-to-end processing sequence is summarized in Figure 2. Whilst the full analytical script and data are provided in the supplementary material.

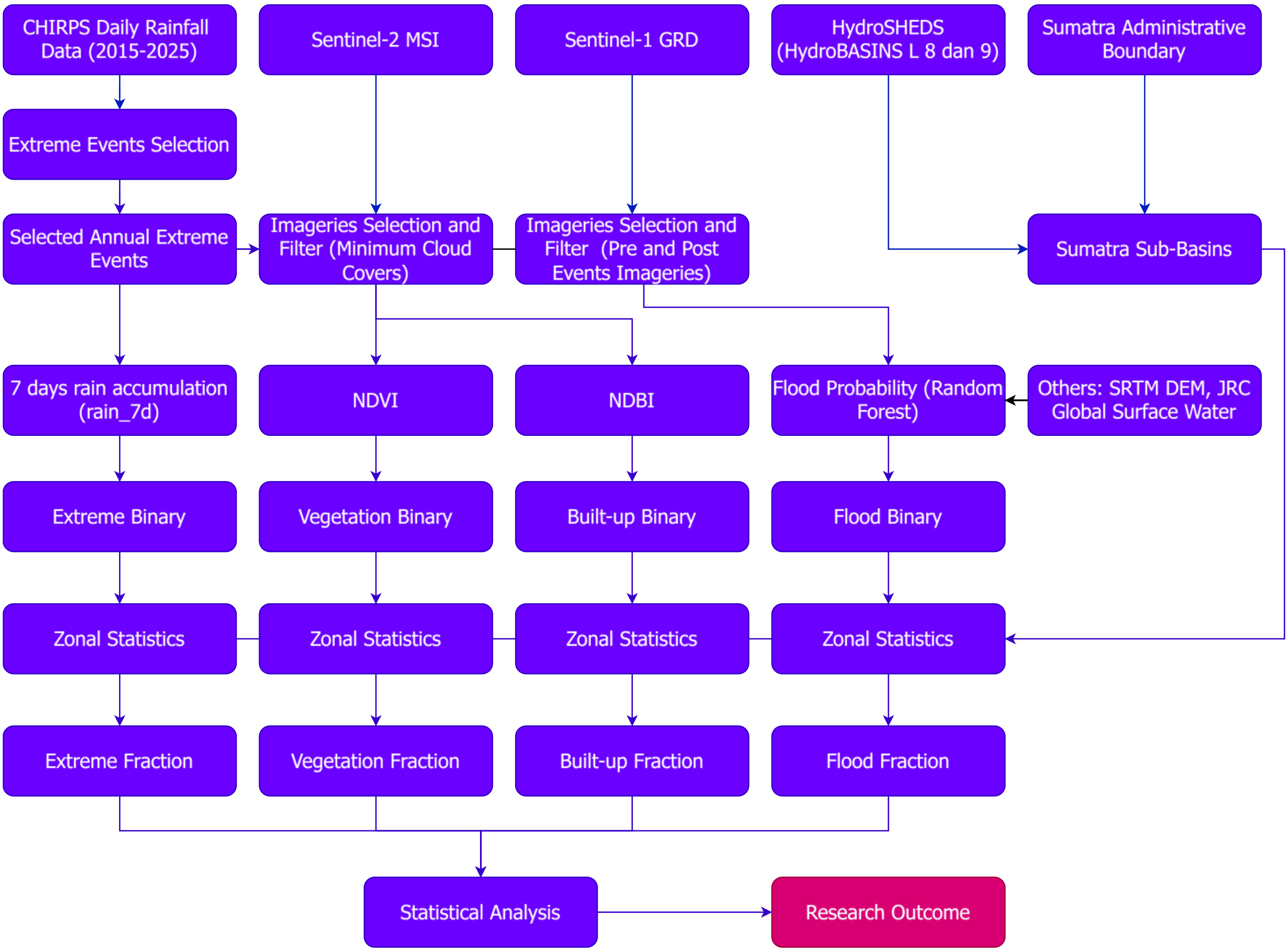

Figure 2. Analytical workflow for the multi-event HydroBASINS-based flood attribution framework. Note. The workflow integrates annual extreme-event selection from CHIRPS, derivation of binary remote-sensing indicators, zonal aggregation to HydroBASINS Level 9 polygons, panel construction, and multivariate modelling.

Table 1. Data sources, resolutions, and analytical functions in the HydroBASINS Level 9 workflow.

Data	Source/collection	Resolution/period	Analytical function
HydroBASINS Level 9	HydroSHEDS/HydroBASINS	Global sub-basin vector	Primary analytical zone; HYBAS_ID identifier; basis of the sub-basin–event panel
CHIRPS Daily	UCSB-CHG/CHIRPS/DAILY	Daily; ≈0.05° grid	Event selection (area-mean daily peak); rain_7d; local P95; extreme_binary; extreme_fraction
Sentinel-2 SR Harmonized	COPERNICUS/S2_SR_HARMONIZED	10–20 m; event year	NDVI; NDBI; vegetation_binary; builtup_binary
Sentinel-1 GRD	COPERNICUS/S1_GRD	10 m; pre- and post-event windows	Backscatter VV/VH; change features; Random-Forest flood probability; flood_binary
SRTM DEM	USGS/SRTMGL1_003	~30 m	Elevation and slope; ancillary features and steep-area masking
JRC Global Surface Water	JRC/GSW1_4/GlobalSurfaceWater	30 m	Permanent water mask to prevent confusion with transient inundation
Climate indices	NOAA CPC ONI / DMI	Monthly	Climate-regime annotation for event interpretation

2.2. Annual Extreme Rainfall Event Selection

Extreme rainfall events are identified from CHIRPS daily rainfall over Sumatra (Figure 3). For each daily timestep we compute several spatial metrics: the area-mean daily rainfall (mean_mm_day), the area-maximum daily value (max_mm_day), the share of area exceeding 50 mm/day, and the share of area exceeding 100 mm/day. mean_mm_day is used as the primary ranking metric because it represents broadly distributed rainfall intensity over the island, whereas max_mm_day and wet-area fractions are retained as supporting characteristics describing local intensity and spatial coverage.

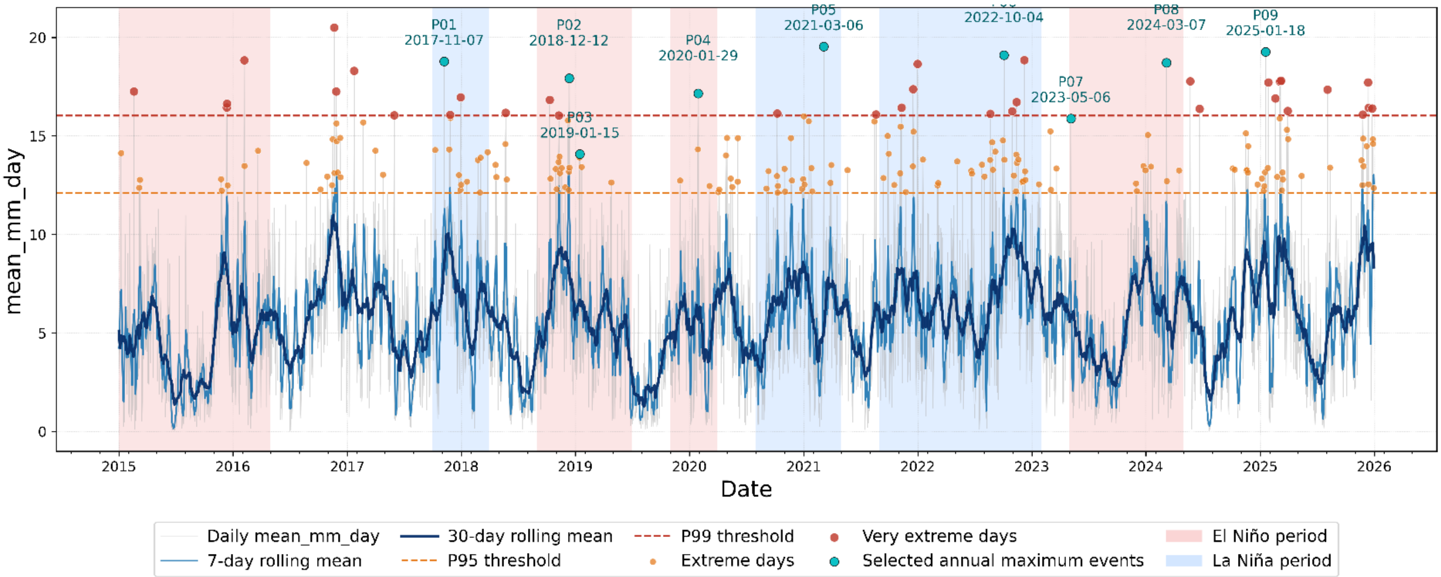

Figure 3. Annual extreme daily rainfall events selected for multi-event analysis (2017–2025). Note. Event dates correspond to the day with the highest area-mean daily rainfall in each calendar year over Sumatra.

For each year from 2017 to 2025, the day with the highest mean_mm_day is selected as the annual extreme rainfall event. This rule prevents temporal clustering of selected events in a single year and yields a sample representative of inter-annual variability. Each event receives a period_no and event_id, for example P01/Y2017_20171107, allowing every derived raster, zonal table, and panel record to be traced consistently. The Extreme/Very extreme classification follows the historical daily-rainfall distribution previously computed during event identification. Two complementary uses of rainfall information are distinguished in this design. Event selection uses the area-mean single-day peak to fix the temporal anchor of each annual event—the day on which island-wide daily rainfall is most intense—whereas the sub-basin extreme_fraction (Section 2.4.3) uses 7-day accumulated rainfall to characterize the multi-day hydrological forcing within each polygon. The two operate at different temporal scales and serve different purposes—temporal anchoring versus hydrologically meaningful forcing characterization—and need not coincide.

2.3. Hydrobasins Level 9 Analytical Units

The spatial unit of analysis is the HydroBASINS Level 9 polygon (Figure 1) clipped to Sumatra. Level 9 is used as a hydrologically aligned compromise between spatial detail and multi-sensor data completeness. Unlike a regular raster cell, a HydroBASINS polygon is defined from river-network topology and catchment structure; however, this study does not benchmark Level 9 against regular grids or administrative units and therefore does not claim that HydroBASINS is statistically superior. Its role is to provide a process-aligned spatial support for the association analysis. The entire workflow is repeated at Level 8 as a scale-sensitivity test rather than as a superiority test.

In the ArcGIS implementation, the HydroBASINS layer is copied to an output geodatabase, then assigned ZONE_INT as a unique sequential integer. All fraction values are subsequently joined back to the event-level feature classes via ZONE_INT. All 1,601 HydroBASINS Level 9 polygons are exported for every event. The full panel therefore contains 14,409 records; the complete analytical panel, which retains only records with all four core parameters populated, contains 11,818 records. The 2017 event has notably low completeness because most vegetation_fraction and builtup_fraction values are missing for that year (due to Sentinel-2 data scarcity). A sensitivity model is consequently estimated with events of n_complete < 500 excluded.

2.4. Construction of Fractional Parameters

Four principal parameters are defined as fractions on the 0 to 1 scale (Table 2). For instance, a value of 0.25 means that 25% of valid pixels within a HydroBASINS Level 9 polygon meet the corresponding criterion, while a value of 1.00 means that all valid pixels do. This definition is more informative than a single binary status, because a sub-basin can simultaneously contain heterogeneous vegetation, built-up surfaces, extreme rainfall, and inundation. The fractional approach also supports multi-resolution integration: CHIRPS, Sentinel-1, and Sentinel-2 are not forced to a single resolution but are aggregated to a consistent hydrological unit.

Table 2. Operational definitions of the four fractional parameters.

Parameter	Source	Operational definition	Scale
vegetation_fraction	Sentinel-2 NDVI	MEAN of binary raster (1 if NDVI ≥ 0.40, 0 otherwise) within each HydroBASINS L9 polygon	0–1
builtup_fraction	Sentinel-2 NDBI	MEAN of binary raster (1 if NDBI ≥ 0.00, 0 otherwise) within each HydroBASINS L9 polygon	0–1
flood_fraction	Sentinel-1 RF flood_probability	MEAN of binary raster (1 if flood_probability ≥ 0.50, 0 otherwise) within each polygon	0–1
extreme_fraction	CHIRPS rain_7d & local P95	MEAN of binary raster (1 if rain_7d > local monthly P95, 0 otherwise) within each polygon	0–1

2.4.1. Vegetation Fraction and Built-Up Fraction from Sentinel-2

Landscape condition is derived from Sentinel-2 Surface Reflectance Harmonized imagery. For each event, an annual median composite is constructed for the event year. Annual composites maximize the number of valid pixels under persistent tropical cloud cover and reduce dependence on a single, potentially clouded scene. Before compositing, imagery is filtered by area and date, restricted to CLOUDY_PIXEL_PERCENTAGE below 80%, and masked using the Scene Classification Layer (SCL) to remove clouds. The resulting median composite is then used to compute NDVI and NDBI.

NDVI is computed as (B8 − B4) / (B8 + B4), where B8 is the near-infrared band and B4 is the red band. vegetation_binary is assigned 1 where NDVI ≥ 0.40 and 0 otherwise; vegetation_fraction is its zonal mean. NDBI is computed as (B11 − B8) / (B11 + B8), where B11 is shortwave infrared and B8 is near-infrared. builtup_binary is assigned 1 where NDBI ≥ 0.00 and 0 otherwise; builtup_fraction is its zonal mean. These thresholds are operational classification choices rather than universal physical boundaries. NDVI = 0.40 was selected conservatively to emphasize actively vegetated pixels rather than sparse or mixed cover, whereas NDBI = 0.00 follows the sign-based interpretation of positive NDBI as built-up-like spectral response. Because threshold values can vary with land-cover type, season, atmosphere, and sensor composite, the resulting fractions should be interpreted conditional on these definitions. A formal multi-threshold sensitivity analysis was not available for the present event set and is identified as a requirement for subsequent model validation.

2.4.2. Flood Fraction from Sentinel-1 with Random Forest Classifier

Flood fraction is derived from a Sentinel-1 (SAR) Random-Forest flood probability (flood_probability) classification. For each event, pre-event imagery is drawn from a 60-day window preceding the event with a 7-day gap to avoid contamination by event-day conditions. Post-event imagery is drawn from the event date through 18 days after the event. Median composites of VV and VH polarizations are constructed for both windows, followed by spatial filtering to mitigate speckle.

Classification features include pre_VV, pre_VH, post_VV, post_VH, pre–post differences and pre/post ratios for both polarizations, SRTM elevation and slope, and the JRC Global Surface Water permanent-water mask. Event-specific pseudo-labels are generated conservatively: a pixel is labeled as flooded when post-event VV < −14 dB, the pre–post VV decrease exceeds 1.5 dB, permanent water is absent, and slope is < 5°. Remaining pixels are assigned to the non-flood class. Stratified sampling then draws up to 1,500 pixels per class per event at the Sentinel-1 processing scale, producing an intentionally balanced pseudo-labeled training set. The Earth Engine smileRandomForest classifier uses 200 trees, minLeafPopulation = 3, bagFraction = 0.7, and random seed = 42. All pseudo-labeled samples are used to fit the event-specific classifier; no independent train/test split is treated as a validation set because such a split would only evaluate agreement with the same pseudo-label rules rather than with independently observed flood extent. The classifier is applied in MULTIPROBABILITY mode, and the probability of class 1 is retained as flood_probability. flood_binary is defined as flood_probability ≥ 0.50, and flood_fraction is the zonal mean of flood_binary. The 0.50 cutoff is the conventional equal-cost decision boundary for a two-class probability output, but it is an operational choice rather than an independently calibrated flood threshold.

2.4.3. Flood-Classification Validation and Uncertainty

No independent, spatially complete ground-truth flood dataset was available for all nine events; consequently, confusion-matrix metrics such as overall accuracy, precision, recall, F1-score, kappa, and IoU cannot be reported without introducing non-independent or fabricated reference data. This is a substantive limitation because flood_fraction is the response variable in the statistical analysis. In the revised interpretation, flood_fraction is therefore treated explicitly as a Sentinel-1-derived flood-detection indicator rather than a fully validated inundation product. The pseudo-label rules, balanced sampling, classifier settings, permanent-water exclusion, slope constraint, and cross-event consistency provide procedural control, but they do not substitute for independent validation. Any operational or predictive application of the framework requires event-specific validation against independent flood observations and sensitivity analysis of the 0.50 probability threshold.

2.4.4. Extreme Fraction from CHIRPS Daily Rainfall (7 Days)

Extreme fraction (extreme_fraction) is constructed from CHIRPS daily rainfall accumulated over a 7-day window. For every CHIRPS pixel, rain_7d is the total rainfall over the seven days ending on the event date. This value is compared against a local P95 threshold, defined as the 95th percentile of the rain_7d distribution for the same calendar month over the climatological period 2001–2024. Extreme binary (extreme_binary) is set to 1 when rain_7d exceeds the local P95 and 0 otherwise. The 7-day accumulation is chosen because, at the HydroBASINS Level 9 sub-basin scale (median area on the order of 100–300 km^2), runoff concentration and flood-front propagation typically span one to several days, so multi-day accumulated rainfall is more hydrologically consistent with the post-event Sentinel-1 observation window than single-day rainfall.

2.5. Raster Export and Zonal Aggregation

To preserve computational stability over Sumatra, each parameter is exported as a separate GeoTIFF. This reduces the risk of timeout and facilitates per-parameter and per-event audit. When Earth Engine subdivides a raster into tiles, all tiles for a given parameter and event are first mosaicked in ArcGIS Pro. The mosaic raster is then used as the value raster in Zonal Statistics as Table.

Zonal Statistics as Table is executed in ArcGIS Pro with the working copy of HydroBASINS Level 9 as the zone layer and ZONE_INT as the zone field. The statistic is MEAN. For binary rasters, MEAN equals the fraction = the count of pixels with value 1 divided by the count of valid pixels within the polygon. NoData is preserved as NoData rather than converted to 0, because NoData can represent the absence of valid observations due to cloud cover, sensor geometry, raster coverage, or tile boundaries. This decision is essential to prevent bias: areas without valid observations must not be treated identically to areas without vegetation, built-up surface, flooding, or extreme rainfall.

The ArcGIS workflow produces one feature class per event and one CSV per event containing all HydroBASINS Level 9 polygons. A Python workflow then merges the nine event-level CSVs into a full panel and a complete-only analytical panel. The complete-only panel retains only records with full information across all core parameters. Completeness is audited per event and per parameter to ensure that statistical models are not biased by partial missingness.

The panel is constructed by combining the nine event CSV files. The base structure of the panel comprises ZONE_INT, period_no, event_id, year, event_date, centroid, sub-basin area, vegetation_fraction, builtup_fraction, flood_fraction, and extreme_fraction. The full panel retains all records, including those with incomplete parameters. The complete-only analytical panel drops any record missing one or more of the four core parameters. This subset is the principal dataset for statistical analysis.

2.6. Statistical Analysis

The statistical analysis is designed to test whether variation in flood_fraction across HydroBASINS Level 9 polygons is associated with vegetation_fraction, builtup_fraction, and extreme_fraction after controlling for common inter-event heterogeneity. The first step is descriptive analysis of distributions, completeness, and event-level variation, followed by Pearson and Spearman correlations. The inferential model is intentionally parsimonious and addresses the association question posed by the study; it does not include the full set of physical controls on flooding. Elevation and slope enter the Sentinel-1 classification workflow, but topographic, hydrological, and substrate variables such as relative elevation, flow accumulation, drainage density, distance to river, TWI, upstream contributing area, soil permeability, and bedrock permeability are not included as covariates in the final GLM. The estimated landscape coefficients must therefore be interpreted as conditional associations, not causal effects or evidence that landscape composition is the primary physical control on flooding.

The main inferential model uses event fixed effects to absorb common between-event variation arising from regional atmospheric conditions, SAR acquisition timing, event intensity, and data coverage differences. Continuous predictors are standardized (z-scored) so that coefficients can be interpreted as the change in the outcome associated with a one standard deviation increase in the predictor. A fractional logit-style Generalized Linear Model is used for the continuous outcome flood_fraction because the outcome is bounded on [0, 1]. To avoid extreme values at the boundaries of the logit link, flood_fraction is clipped to the open interval using an epsilon of 1×10^{-6}. The fractional response specification is presented as Equation 1.

$$logit\left(E\left[flood_{fraction_{ie}}\right]\right) = \beta_0 + \beta_2 zBuilt_{ie} + \beta_3 zExtreme_{ie} + \beta_4 (zExtreme \times zVeg)_{ie} + \beta_5 (zExtreme \times zVeg)_{ie} + Y_e + \varepsilon_{ie} \quad (1)$$

where i indexes the HydroBASINS Level 9 polygon, e indexes the event, and γ_e denotes the event fixed effects. As a complementary specification, logistic regression is fitted on flood_status, a binary indicator equal to 1 when flood_fraction ≥ 0.01 and 0 otherwise. The 0.01 cut-off identifies polygons in which at least 1% of valid pixels are detected as flooded. Standard errors are estimated using cluster-robust covariance based on zone_id where feasible, because each polygon appears across several events. Where cluster-robust estimation fails, HC3 heteroskedasticity-consistent standard errors are used as a fallback. A sensitivity analysis is conducted with events of n_complete < 500 excluded, to verify that the main results are not driven by events with extremely low coverage.

To assess whether the principal findings depend on the chosen analytical unit, the entire workflow was repeated at the coarser HydroBASINS Level 8 scale. Level 8 sub-basins are approximately three to five times larger than Level 9 polygons and therefore internalize more of the upstream–downstream connectivity within each unit. The four binary rasters were re-aggregated to Level 8 polygons using identical Zonal Statistics procedures, and the same fractional logit and logistic specifications were estimated on the resulting panel.

3. Results

This section presents the analytical results based on the HydroBASINS Level 9 unit. The analysis uses every HydroBASINS Level 9 polygon for each event; the unit of observation is the sub-basin–event pair, yielding a multi-event panel. The four parameters analyzed are vegetation_fraction, builtup_fraction, flood_fraction, and extreme_fraction, all defined on a [0, 1] scale. The results therefore not only describe the distribution of flooding, but also test how landscape structure and extreme-rainfall coverage relate to flood response within a hydrologically consistent unit.

3.1. Records Structure and Data Completeness

The full data comprises 14,409 records (1,601 HydroBASINS Level 9 polygons × 9 annual extreme events between 2017 and 2025). After screening for complete information across all four core parameters, the analytical panel contains 11,818 observations, equivalent to approximately 82.0% of all polygon–event combinations. Completeness is not uniform across events. The 2017 event has the lowest completeness, with only 140 complete records (8.74% of total polygons), driven by 90.38% missingness in vegetation_fraction and builtup_fraction. Events in 2018 and 2019 each have 1,408 complete records, whereas events from 2020 to 2025 each have 1,477 complete records. This pattern indicates that the principal constraint is not the number of HydroBASINS polygons but the coverage of Sentinel-2 optical parameters in the earliest year of the analysis. Figure 4 visualizes the event-wise pattern of completeness and highlights the pronounced constraint in the 2017 data.

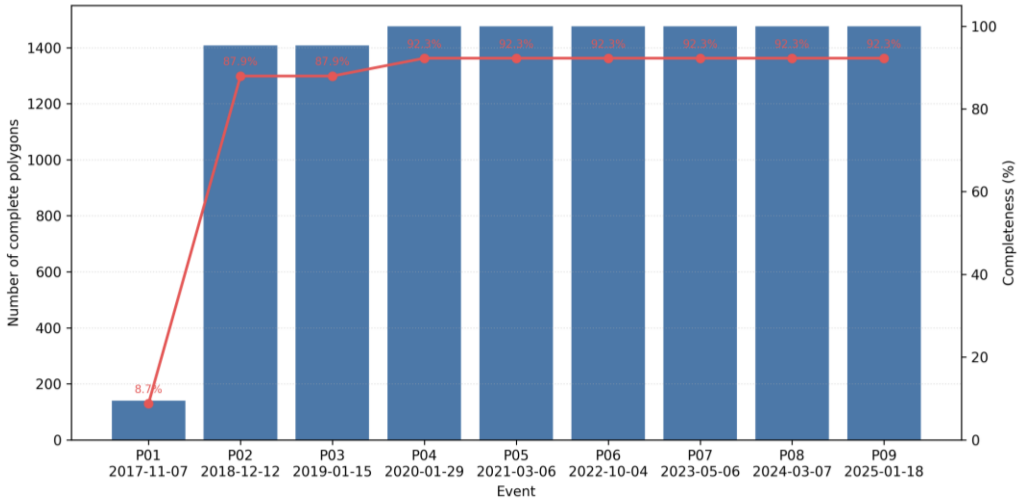

Figure 4. Event-wise completeness of the HydroBASINS Level 9 analytical panel. Bars show the number of complete HydroBASINS Level 9 polygons per event, while the line shows the percentage of complete records.

3.2. Descriptive Statistics of the Principal Parameters

The analytical panel reveals a vegetated landscape dominating the sub-basin units of the study area. The mean vegetation_fraction is 0.924 with a median of 0.962, showing that the great majority of polygons have very high vegetation cover. Built-up surfaces, in contrast, are limited at the sub-basin scale: builtup_fraction has a mean of 0.040 and a median of 0.023, with a maximum of 0.653. The flood_fraction is right-skewed, with a mean of 0.075 and a median of 0.048, yet with a maximum of 1.000; this skewness implies that most sub-basins experience limited inundation while a subset of polygons exhibit very high flood fractions. The extreme_fraction is strongly bimodal, with a mean of 0.391, a median of 0.140, and a third quartile of 1.000, indicating that the spatial coverage of extreme rainfall within sub-basins is highly heterogeneous: in some polygons no CHIRPS pixel exceeds the local P95, whereas in others every valid pixel does.

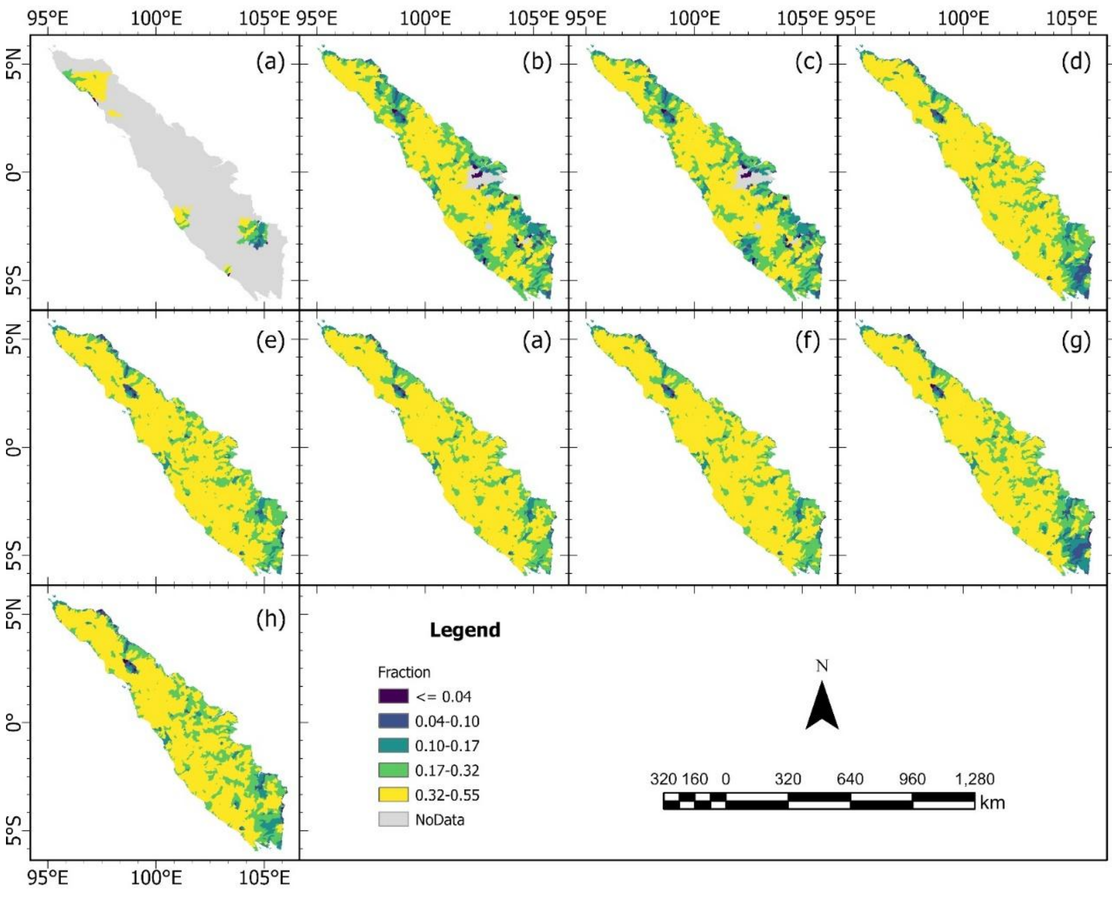

Figure 5. Vegetation Fraction (a to h expresses the years 2017 to 2025, respectively).

The vegetation fraction (Figure 5) is concentrated in the upper portion of the [0,1] range across all nine events, with annual means ranging narrowly from 0.877 (2018, 2019) to 0.946 (2022, 2023) and medians consistently above 0.92. Spatially, the lower tail (vegetation_fraction < 0.80, only 8.8% of records) is distributed across all three zones of Sumatra 384 records in the south (Lampung, South Sumatra), 349 in the north (Aceh, North Sumatra), and 305 in the central belt (Riau, Jambi, West Sumatra) but is disproportionately concentrated along the coastal margin: 39% of low-vegetation polygons are coastal compared with 12% in the panel overall, reflecting the eastern lowland mosaic of oil-palm estates, settlements, and degraded peatland-fringe landscapes. The dynamics across years are subtle: the lower-mean 2018–2019 values likely reflect Sentinel-2 composite limitations rather than real defoliation, while the upward drift from 2020 onward (mean 0.93–0.95) parallels improved valid-pixel coverage.

The built-up fraction (Figure 6) has annual means between 0.031 and 0.076 but maximum reaching 0.32–0.65, reflecting a landscape in which most sub-basins have minimal impervious cover while a small set of urbanized polygons carries the entire upper tail. The 90th percentile clusters at 0.07–0.17 across events, and the distribution sits in the interior of [0,1] (mid-range > 98% in every event) with negligible boundary mass. Spatially, the upper tail is highly localized: only 106 unique sub-basins ever exceed built-up_fraction > 0.20 across the nine events, and these 232 polygon-event records are concentrated in the south (140 records, around the Palembang–Lampung urban corridor) followed by the central east coast (59 records, Pekanbaru–Jambi belt) and the north (33 records, Medan metropolitan area). The pattern matches Sumatra's urbanization geography: the southern Sumatra lowland corridor and the eastern Riau alluvial plain are denser and more urbanized than the western mountainous spine. The temporal dynamics show a lower-mean phase in 2018–2019 (mean ≈ 0.037) constrained by Sentinel-2 composite quality and a higher-mean phase from 2020 onward.

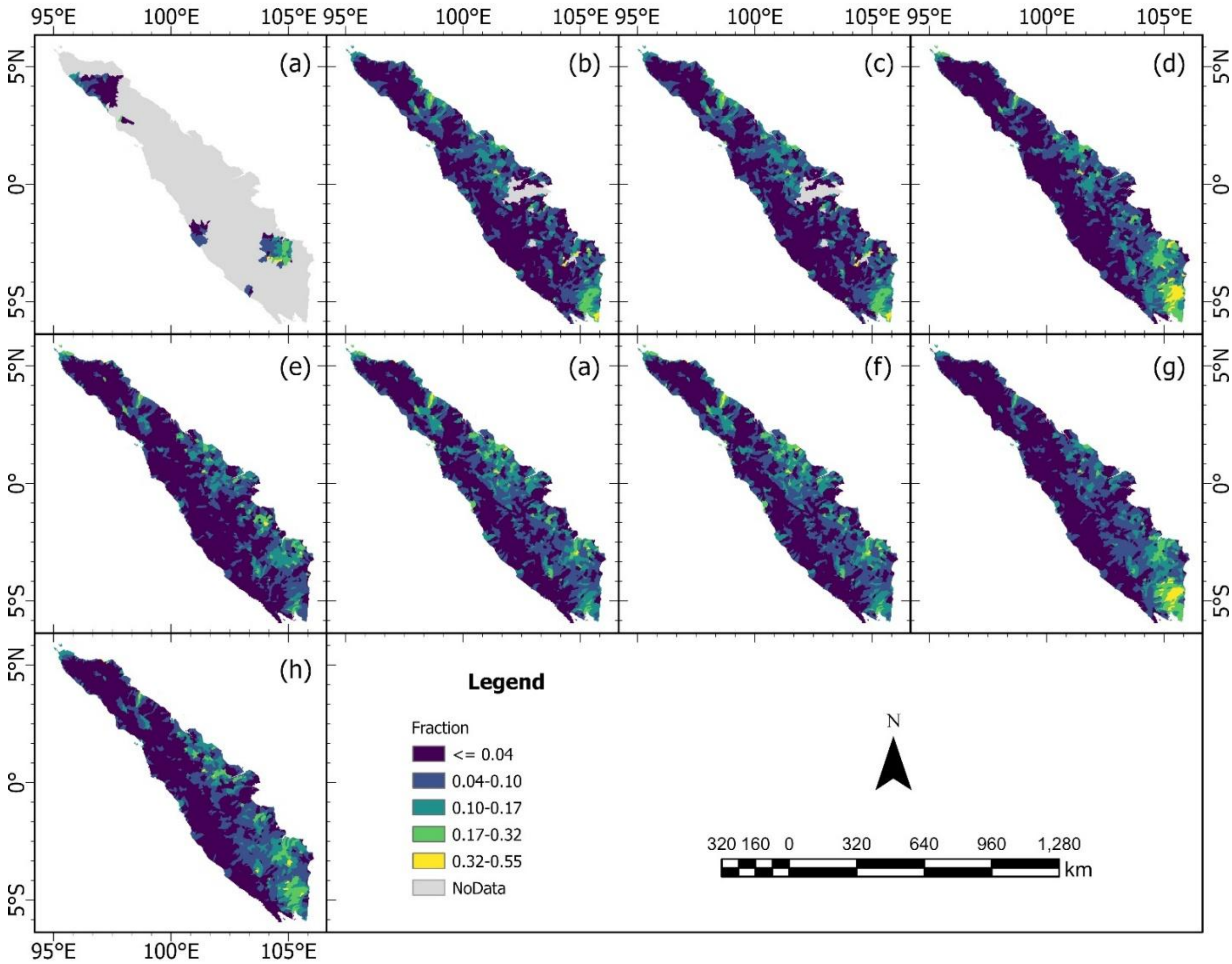

Figure 6. Built-up Fraction (a to h expresses the years 2017 to 2025, respectively).

The extreme fraction (Figure 7) exhibits the largest cross-event variability and the most bimodal distribution of any parameter. Annual means range from 0.014 (2017) and 0.020 (2021) at the dry tail to 0.322 (2024) and 0.342 (2025) at the wet tail. Across the full panel, 71% of records have extreme_fraction = 0, 13% have extreme_fraction = 1, and 16% lie in the interior. Crucially, the spatial signature of "extreme" varies dramatically from event to event: the 2024 event was a southern Sumatra event (mean extreme_fraction 0.587 in the south versus 0.036 in the north), the 2018 and 2022 events were central-Sumatra events (mean 0.428 and 0.443 in the central belt versus 0.06–0.18 elsewhere), while the 2025 event had its widest extreme coverage across the central belt (0.545). The 2017, 2020, and 2021 events were spatially focal, with means below 0.15 in every zone. Importantly, the year with the highest extreme coverage (2025, mean 0.342) is not the year with the highest flood response (2017, mean 0.099), and there is no monotonic correspondence between the spatial location of extreme rainfall and the spatial location of inundation across years.

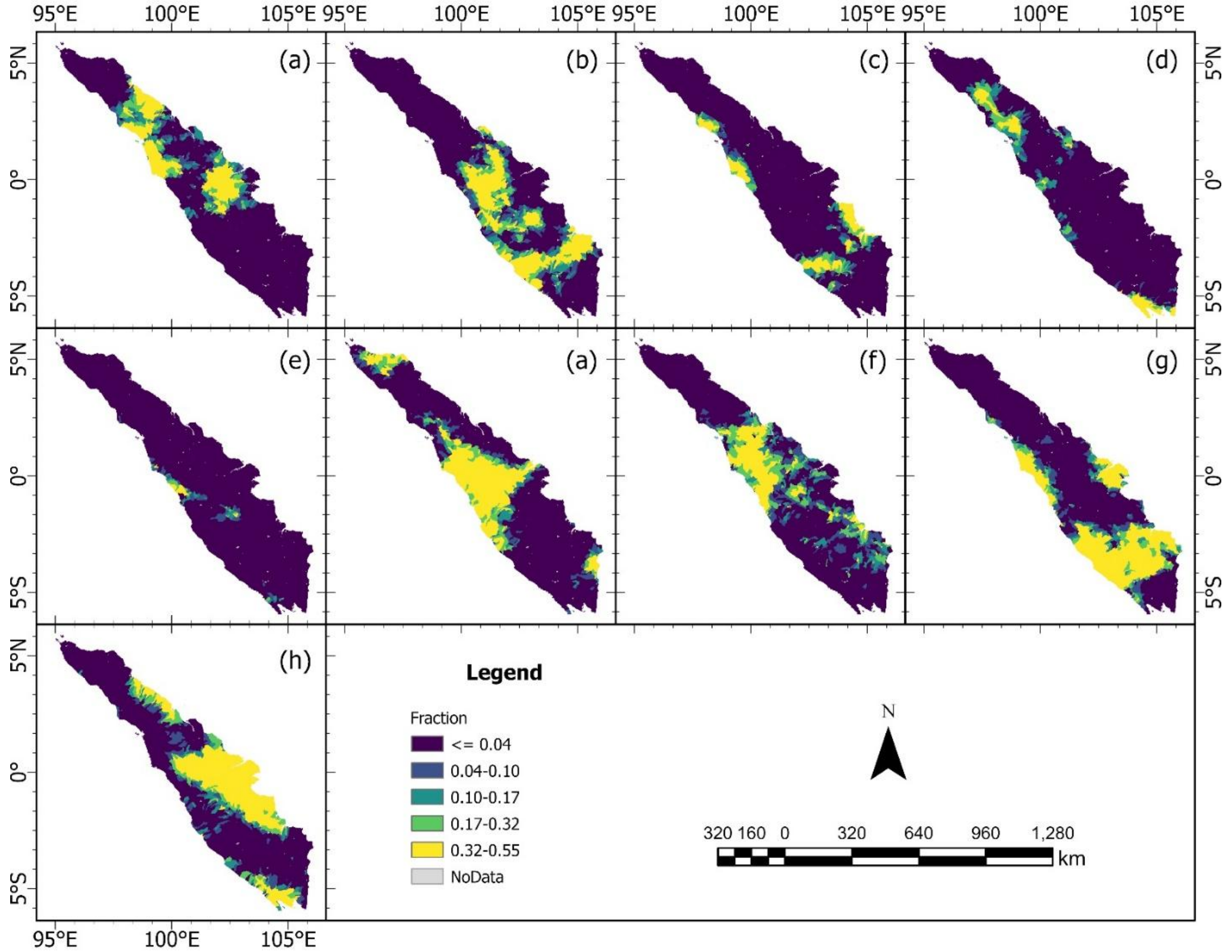

Figure 7. Extreme Rainfall Event Fraction (a to h express the years 2017 to 2025, respectively).

The flood fraction (Figure 8) has annual means in a narrow band from 0.050 (2024) to 0.099 (2017) and medians from 0.033 to 0.073. The 90th percentile in each event lies between 0.11 and 0.22, and maxima reach 0.59–1.00, indicating that while most sub-basins experience flood_fraction below 0.2 there are always a handful of severely inundated polygons — long-tailed behaviour typical of natural hydrological response. Spatially, the southern third of Sumatra (south of 3°S, covering South Sumatra and Lampung) consistently records the highest mean flood_fraction in 7 of 9 events (0.077–0.214 in the south versus 0.046–0.105 elsewhere), reflecting the dominance of the Musi–Banyuasin floodplain and lowland peatlands that respond to upstream contributions rather than local rainfall. Inundation also rises monotonically with upstream contributing area, consistent with downstream accumulation.

Inter-event variation indicates that mean flood intensity does not move in step with extreme-rainfall coverage. The 2023 event has the highest mean flood_fraction (0.099) despite an intermediate extreme_fraction (0.392). The 2022 event has the highest mean extreme_fraction (0.564), yet only an intermediate mean flood_fraction (0.071). This mismatch confirms that flood_fraction is not a simple linear response to extreme_fraction. Hydrologically, sub-basin inundation can be governed by polygon position in the drainage network, upstream runoff contribution, drainage capacity, local topography, antecedent moisture, and Sentinel-1 acquisition timing. Extreme_fraction thus remains useful as an indicator of extreme-rainfall coverage, but is insufficient as the sole predictor of detected flood fraction. The event-level trajectories of the four principal fractions are shown in Figure 9.

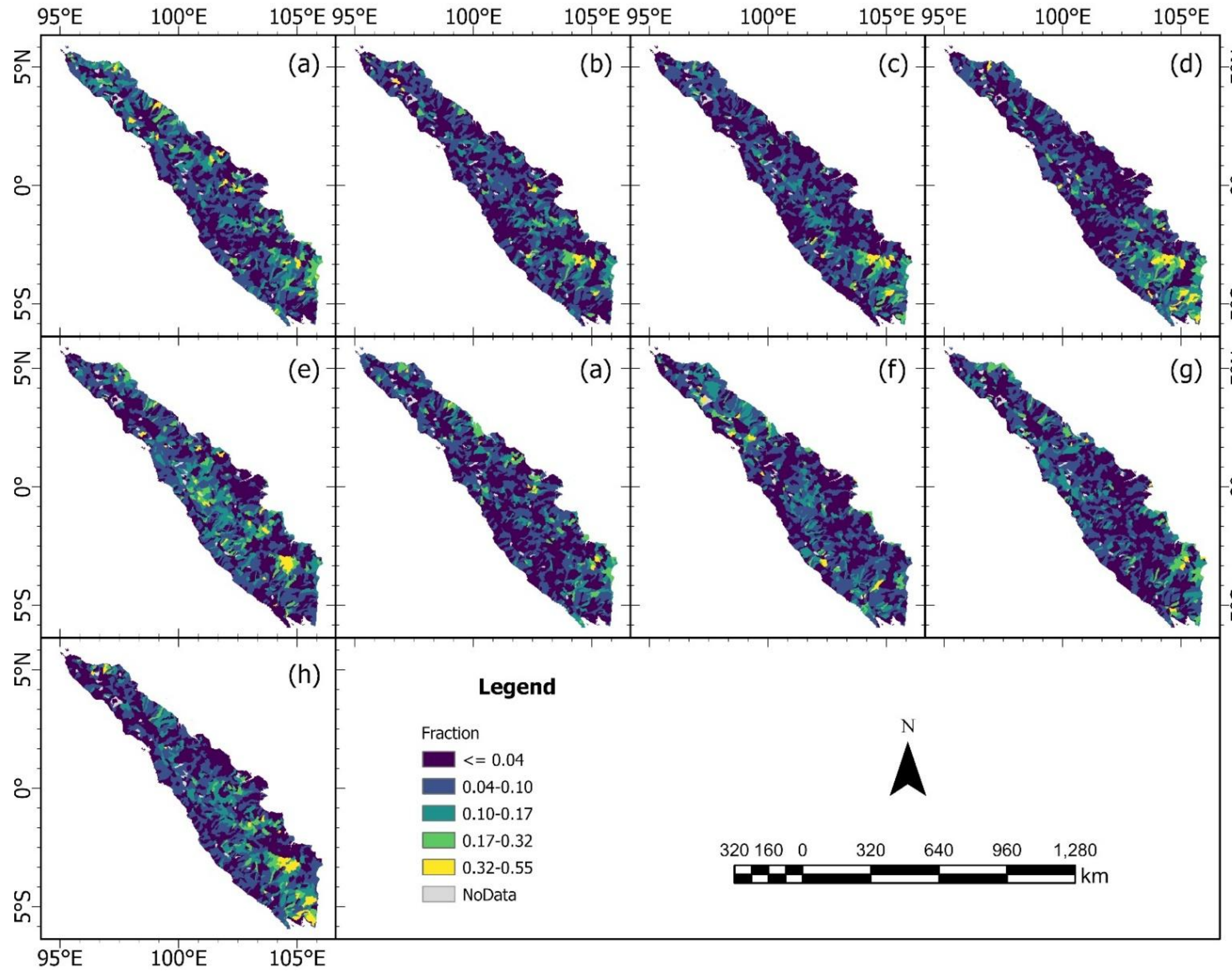

Figure 8. Flood Fraction (a to h expresses the years 2017 to 2025, respectively).

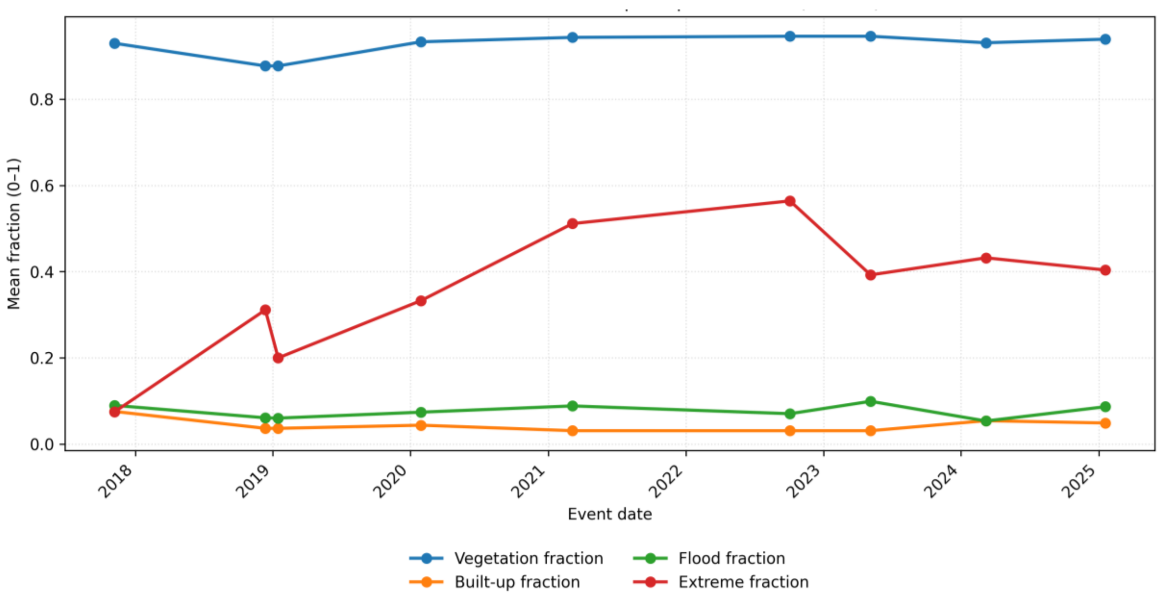

Figure 9. Event-level means of the four principal fractions across the nine annual extreme events. Note. Vegetation fraction remains uniformly high, built-up fraction remains low, flood fraction varies modestly, and extreme fraction shows the strongest inter-event variation.

3.3. Bivariate Correlation Structure

Global correlations are weak in magnitude but the directional associations between landscape variables and flooding are conceptually consistent. Vegetation_fraction correlates negatively with flood_fraction (Pearson r = −0.058; Spearman ρ = −0.088). Built-up fraction correlates positively (Pearson r = 0.081; Spearman ρ = 0.063). Extreme_fraction shows almost no global correlation with flood_fraction (Pearson r = 0.004; Spearman ρ = −0.005). The strongly negative correlation between vegetation_fraction and builtup_fraction (Pearson r = −0.444; Spearman ρ = −0.556) reflects compositional landscape structure: polygons with more vegetation tend to have less built-up surface.

Event-specific correlations reveal sharper heterogeneity than the global aggregate. Vegetation_fraction is generally negatively correlated with flood_fraction, with the strongest negative associations in 2020 (ρ = −0.222), 2022 (ρ = −0.270), and 2024 (ρ = −0.166). Built-up fraction tends to be positively correlated but with variable strength across years; the strongest positive associations occur in 2017 (ρ = 0.105) and 2020 (ρ = 0.217). Extreme fraction shows no stable sign across events. This event-level heterogeneity motivates the use of event fixed effects in the inferential models, because each event carries a distinct combination of rainfall pattern, observational conditions, and spatial context. The global correlation structure is summarized in Figure 10.

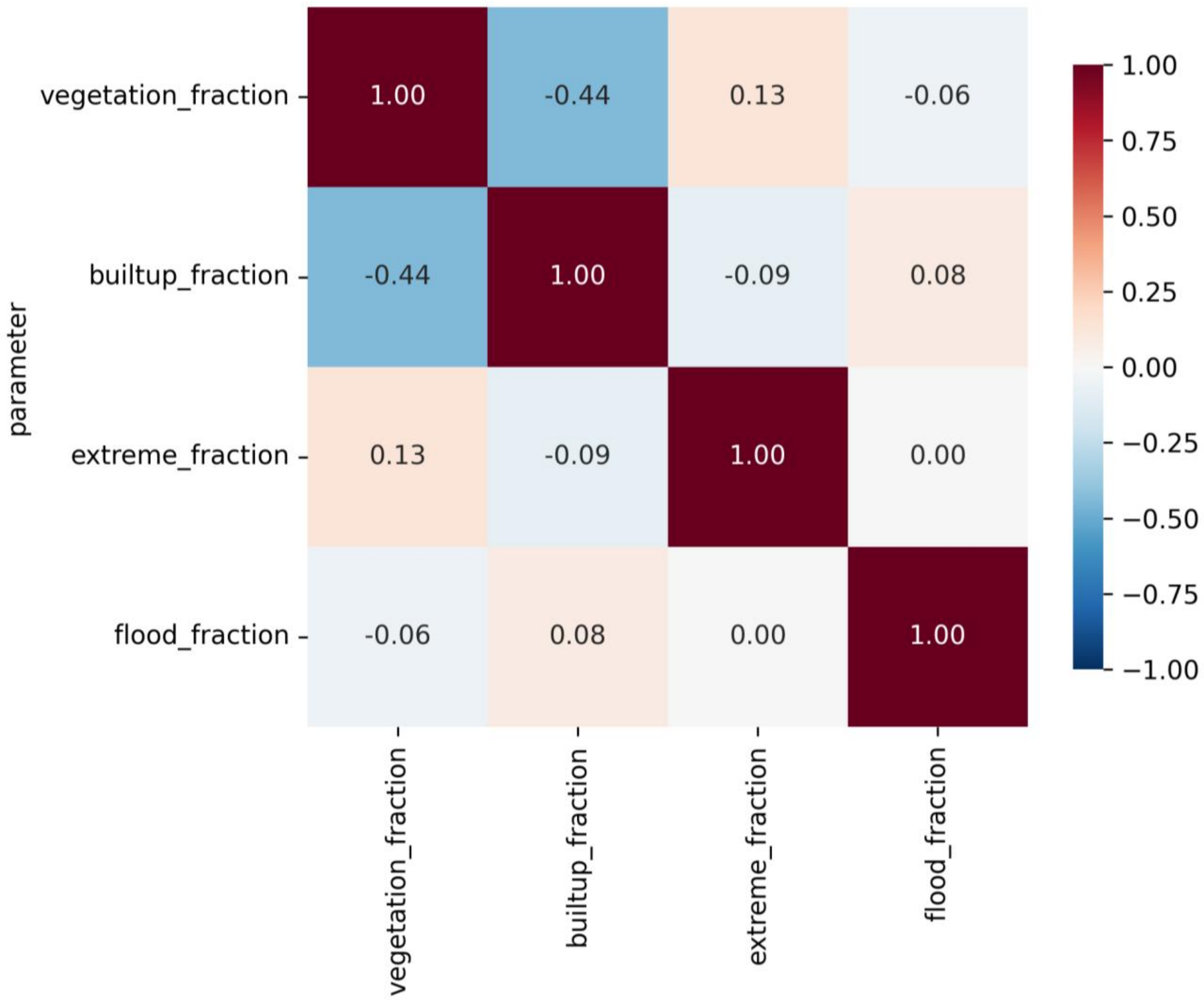

Figure 10. Pearson correlation matrix of the principal parameters in the complete Level 9 panel. Note. The heatmap emphasizes the weak direct correlation between extreme fraction and flood fraction, alongside the modest but interpretable negative association between vegetation and flooding and positive association between built-up area and flooding.

3.4. Multivariate Models of Flood Response

The fractional logit GLM evaluates variation in flood_fraction as a bounded continuous outcome, whereas the logistic GLM evaluates the probability that a polygon exceeds the flood_status threshold (flood_fraction ≥ 0.01). Both specifications include event fixed effects so that polygon-level effects are not confounded with general inter-event differences. All continuous predictors are z-standardized, so coefficients represent the change in the outcome associated with a one standard deviation increase in the predictor.

In the fractional logit specification on the full complete-only panel, vegetation_fraction is negatively associated with flood_fraction ($\beta = -0.073$, 95% CI −0.123 to −0.024, $p = 0.004$), whereas builtup_fraction is positively associated ($\beta = 0.082$, 95% CI 0.040 to 0.123, $p < 0.001$). extreme_fraction is not significant ($\beta = -0.001$, $p = 0.944$); the extreme × vegetation interaction is non-significant, and the extreme × built-up interaction is weak and not robust across scale or model specification. The statistically significant landscape coefficients are modest on the logit link scale: $\exp(-0.073) \approx 0.93$ and $\exp(0.082) \approx 1.09$, corresponding to roughly 7% lower and 9% higher odds of the conditional mean flood fraction per one-standard-deviation change in vegetation and built-up fraction, respectively. These are not percentage-point changes in inundated area and should not be interpreted as large practical effects. The results therefore support modest, directionally consistent landscape associations rather than dominant landscape control.

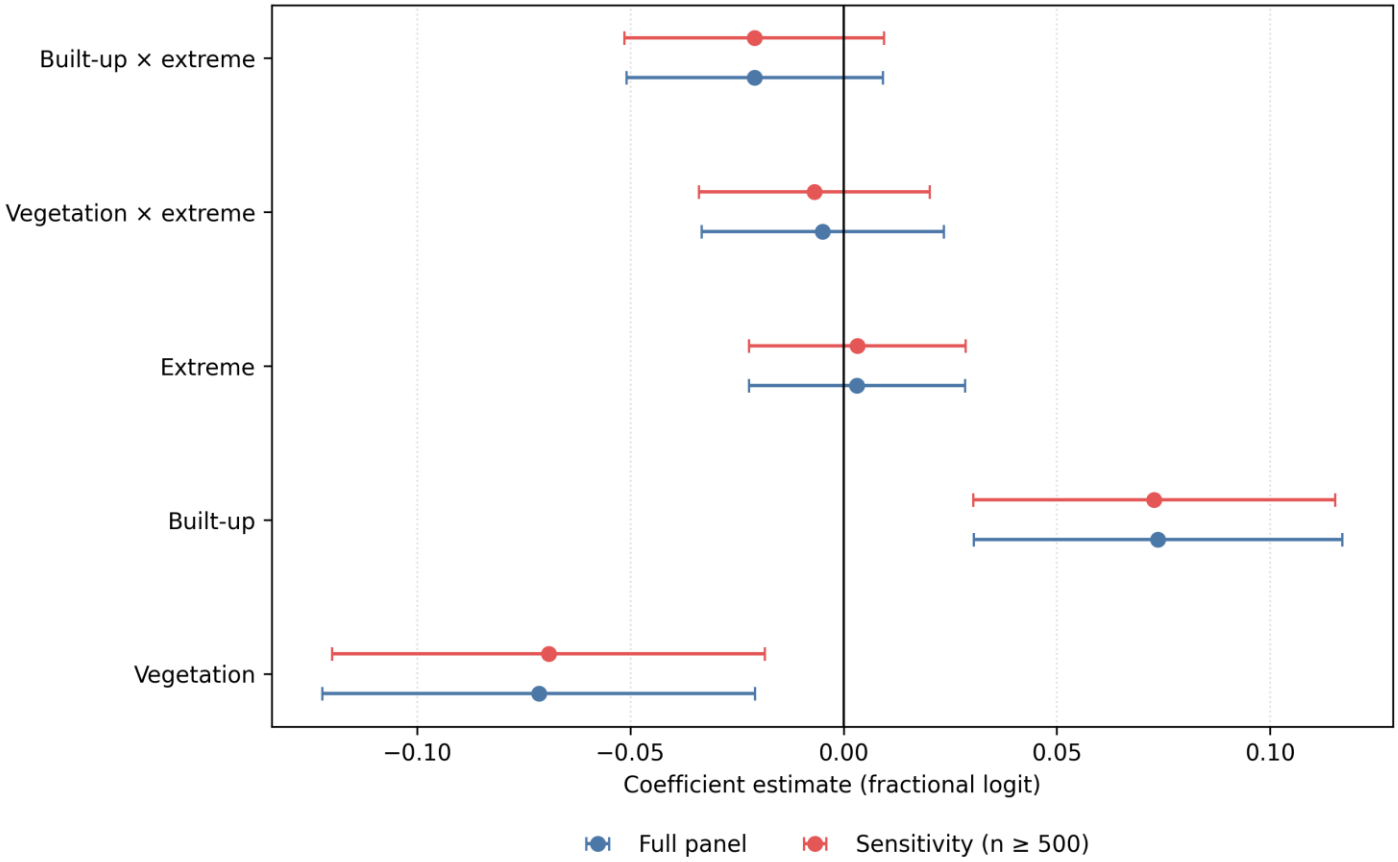

Figure 11. Fractional logit coefficients for the flood-fraction model under the full and sensitivity specifications. Error bars indicate 95% confidence intervals. Built-up fraction is the most stable positive predictor, whereas vegetation fraction is consistently negative. Direct and interaction effects involving extreme fraction are weak.

The logistic model produces a partially aligned but also distinct pattern. Built-up fraction increases the odds of flood occurrence in the full panel (OR = 1.136, $p = 0.019$) and remains significant in the sensitivity model (OR = 1.124, $p = 0.030$). Vegetation_fraction is not significant in the logistic model, indicating that the vegetation signal is stronger for the intensity of flood_fraction than for the binary occurrence of flooding. Extreme_fraction has an odds ratio slightly below one but is not statistically significant (OR = 0.964, $p = 0.150$) in either the full or sensitivity models. Its effect on flood occurrence is indistinguishable from zero once event fixed effects and landscape variables are controlled, consistent with the null direct effect in the fractional model. extreme_fraction, defined as exceedance of the local P95 of rain_7d, does not by itself capture the mechanisms generating inundation at this scale. Flood occurrence is also determined by runoff propagation, sub-basin position, topography, drainage connectivity, and Sentinel-1 observation timing.

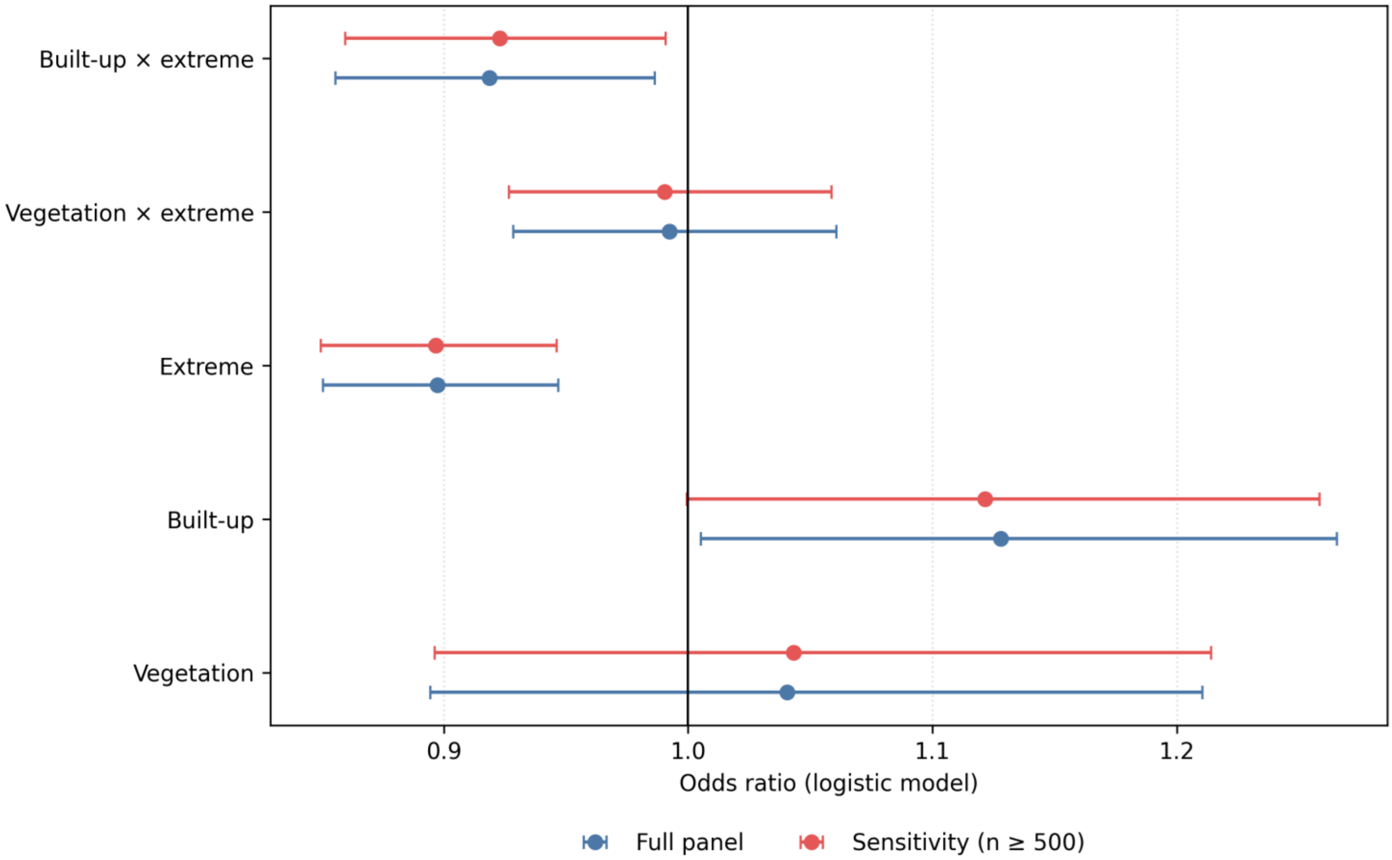

Figure 12. Logistic-model odds ratios for flood occurrence under the full and sensitivity specifications. Odds ratios greater than 1 indicate higher odds of measurable flood occurrence. The built-up effect remains positive, whereas vegetation and extreme fraction do not show consistent direct effects across specifications.

3.5. Sensitivity and Consistency of Findings

The sensitivity analysis excludes events with very few complete records (n_complete < 500). Under this criterion, only the 2017 event is removed and events from 2018 to 2025 are retained. The fractional model preserves the principal directions: vegetation_fraction remains negatively associated with flood_fraction ($\beta = -0.069$, $p = 0.008$), built-up fraction remains positively associated ($\beta = 0.073$, $p < 0.001$), and extreme_fraction and its interactions do not provide a stable direct signal. In the logistic specification, built-up fraction remains positive, whereas vegetation and extreme-rainfall effects are not consistently significant. The core interpretation therefore does not depend on the low-completeness 2017 event, but the sensitivity analysis addresses event completeness only; it does not resolve omitted-variable bias, flood-map validation, or threshold sensitivity.

Comparison of the full and sensitivity models suggests that built-up fraction is the most stable landscape predictor, especially for explaining flood-fraction intensity and flood-status occurrence. Vegetation_fraction is consistently negatively associated with flood_fraction in the fractional response model but not significant in the logistic model, indicating that vegetation relates more to the magnitude of inundation than to its binary presence. Extreme_fraction remains an important descriptor of extreme-rainfall spatial coverage, but cannot stand alone as a direct predictor of flood_fraction. Together, these results reinforce the argument that tropical flood response is mediated by landscape composition and hydrological structure, rather than being a direct function of local extreme-rainfall exceedance.

3.6. Cross-Scale Robustness: Hydrobasins Level 8

To assess whether the principal findings depend on the chosen hydrological aggregation scale, the entire workflow was repeated using HydroBASINS Level 8 sub-basins, which are roughly three to five times larger than Level 9 units and therefore internalize more of the upstream–downstream connectivity within each polygon. The Level 8 complete panel comprises 4,854 records across 601 sub-basins and the same nine annual events, and was analysed with identical specifications: z-standardized predictors, event fixed effects, cluster-robust standard errors by sub-basin, and the same fractional logit and logistic models. The motivation is twofold: to test the Modifiable Areal Unit Problem (MAUP), under which spatial relationships can change with the size of the analytical unit, and to evaluate whether enlarging the unit — so that rainfall and flood response co-occur more often within the same polygon — strengthens the otherwise weak direct rainfall signal.

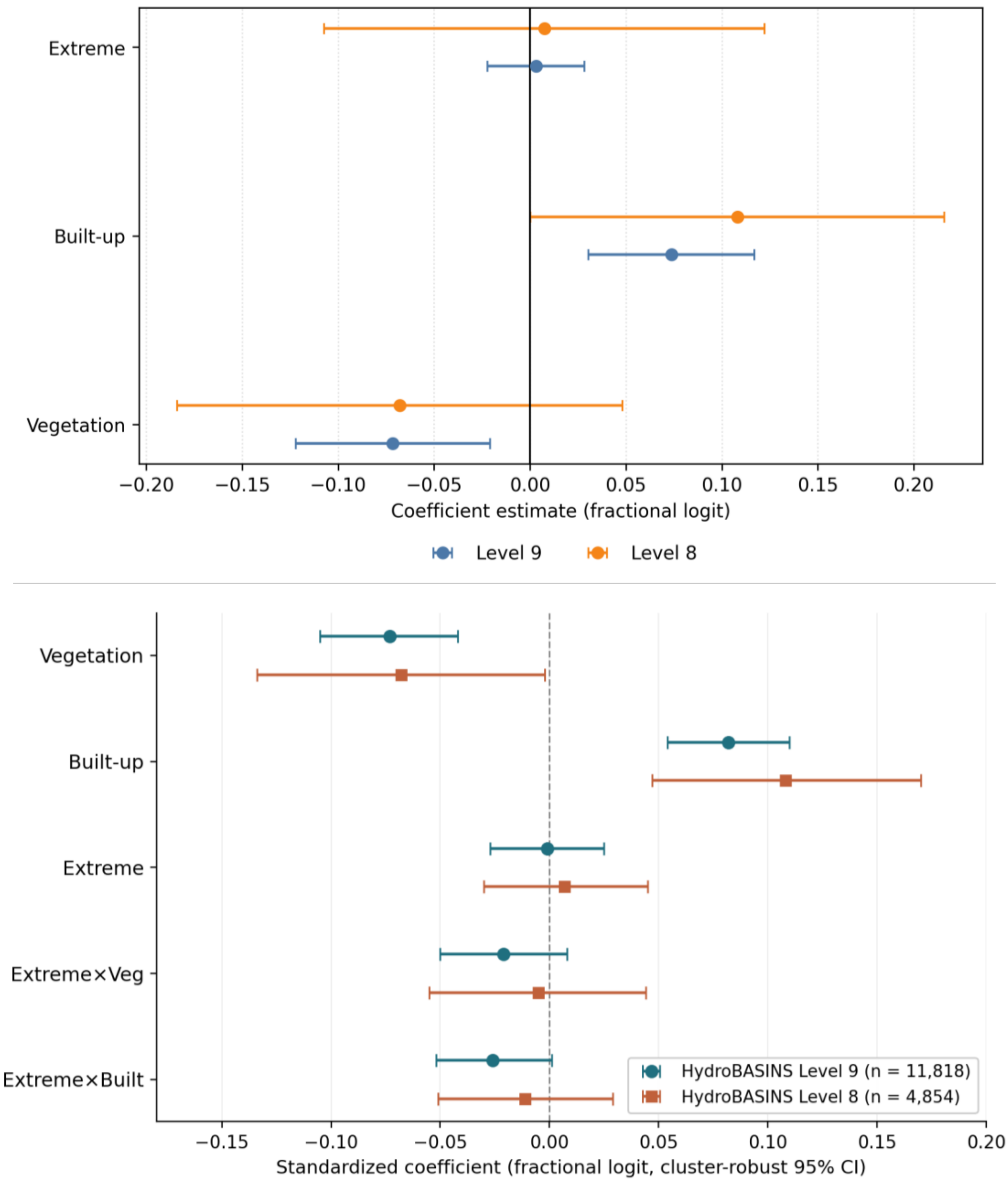

Figure 13. Cross-scale comparison of core fractional-logit coefficients between HydroBASINS Levels 9 and 8. Note. Standardized coefficients with 95% confidence intervals are shown for vegetation fraction, built-up fraction, and extreme fraction. The direction of the vegetation and built-up effects is stable across scales, supporting the robustness of the main Level 9 findings.

The cross-scale comparison is summarized in Figure 13. The two landscape effects are stable in sign across both scales. In the fractional response model, vegetation fraction remains negative at both Level 9 ($\beta \approx -0.073$) and Level 8 ($\beta \approx -0.068$), while built-up fraction remains positive and becomes slightly stronger at the coarser scale.

The direct effect of extreme fraction also remains statistically indistinguishable from zero at Level 8 (fractional logit $\beta = 0.007$, $p = 0.697$; logistic OR = 0.964, $p = 0.475$), and the interaction terms are non-significant. The weak Level 9 extreme × built-up signal does not persist at the coarser scale. This cross-scale result indicates that the observed rainfall–flood association is not strengthened simply by enlarging the hydrological unit. It should not, however, be interpreted as proof that spatial-unit mismatch is irrelevant or that landscape variables dominate omitted hydrological controls; Level 8 and Level 9 remain two levels within the same HydroBASINS hierarchy and do not constitute a benchmark against grid-based or administrative-unit analyses.

The logistic specification at Level 8 reveals one scale-dependent nuance. Vegetation fraction, which is negative in the fractional response, turns positive in the binary occurrence model at Level 8 (OR = 1.164), although it does not reach significance under cluster-robust standard errors ($p = 0.170$). This sign divergence between the fractional and binary outcomes reinforces the dual hydrological role of vegetation discussed in Section 4.3: highly vegetated Level 8 polygons — which frequently encompass riparian zones, swamps, and natural floodplains — are more likely to register some inundation above the flood_status threshold, even as vegetation reduces the overall proportion of the sub-basin that floods. The cross-scale results thus do not contradict the primary analysis; they refine it, showing that the magnitude-reducing role of vegetation and the occurrence-enabling role of floodplain vegetation operate simultaneously.

4. Discussion

Across the nine annual extreme-rainfall events, the analysis identifies weak bivariate correlations, modest but directionally stable landscape coefficients, and no robust direct association between local extreme-rainfall coverage and SAR-derived flood fraction after event fixed effects are introduced. These findings should be interpreted within the restricted covariate set and the event-conditioned design. They do not establish that landscape composition is the primary physical control on flooding. Rather, among the variables evaluated here, vegetation and built-up fractions show more stable associations with flood_fraction than extreme_fraction does. The result is consistent with a hydrological system in which rainfall forcing is translated into inundation through routing, storage, topographic position, surface permeability, drainage structure, antecedent wetness, and observation timing—several of which are not explicitly represented in the final regression model.

By adopting HydroBASINS Level 9 as the analytical unit, the analysis shifts the inferential basis from meteorological grids to hydrological units more aligned with routing, accumulation, and flow connectivity. HydroBASINS provides a globally consistent hydrographic framework for studying river systems and flow networks across scales (Lehner & Grill, 2013); in this study, that framework allows each observation to be interpreted as a sub-basin–event pair, so that flood_fraction represents not just a value at a raster cell, but the inundation response within a hydrologically coherent catchment.

4.1. The Roles of Landscapes in Flood Responses

The analytical panel is dominated by sub-basins with high vegetation_fraction (mean 0.924) and low builtup_fraction (mean 0.040), while flood_fraction is right-skewed (mean 0.075; median 0.048). The maximum flood_fraction reaching 1.000 in some sub-basins implies that, although most hydrological units experience limited inundation, a subset of sub-basins can respond to extreme rainfall with extensive flooding. This pattern is consistent with remote-sensing studies showing that flood extent does not spread uniformly but is controlled by topography, flow connectivity, land cover, drainage capacity, and sensor acquisition timing (Riazi *et al.,* 2023; San Jose *et al.,* 2026; Shampa *et al.,* 2025).

This result aligns with the broader literature that conceptualizes flooding as a multi-factor process irreducible to a single rainfall threshold. Studies of compound flooding and hydrometeorological extremes show that rainfall, soil moisture, river discharge, tide, drainage configuration, and land use can interact non-linearly (Atmaja & Lee, 2026; Mantovani *et al.,* 2025; Pizzorni *et al.,* 2024). In tropical regions such as Sumatra, this complexity is amplified by mesoscale convection, MJO, ENSO, IOD, and local orographic variation that jointly shape the intensity and location of extreme rainfall (Chrysanti & Son, 2025; Saufina *et al.,* 2025; Syamsudin *et al.,* 2026). The contribution here is not to argue that extreme rainfall is unimportant, but to demonstrate that extreme-rainfall indicators must be situated within a more comprehensive hydro-landscape framework.

The weak rainfall–flood relationship requires particular attention because extreme rainfall is the event-selection criterion. Conditioning the analysis on one major area-mean rainfall event per year restricts the range of rainfall conditions represented in the sample. Once events are already selected from the upper tail of rainfall intensity, differences in local P95 exceedance coverage may explain relatively little additional variation in flood extent. This range-restriction mechanism is hydrologically plausible: above a sufficiently high rainfall regime, spatial differences in inundation can increasingly reflect runoff routing, storage, topographic convergence, antecedent wetness, drainage capacity, and surface permeability rather than the local exceedance fraction alone. Multi-factor flood studies likewise show that rainfall, land use, topography, flow accumulation, and drainage interact non-linearly (Riazi *et al.,* 2023; Singha *et al.,* 2024; Mantovani *et al.,* 2025). The near-zero correlation with extreme_fraction therefore does not imply that rainfall is unimportant; it indicates that this particular rainfall metric has limited discriminatory power within a sample already conditioned on major events.

The event-fixed-effect results reinforce that interpretation without resolving the underlying mechanism. extreme_fraction is not significant in the fractional model or the full-panel logistic model, but the model omits several physical controls that can confound or mediate the rainfall–flood relationship. The null coefficient therefore cannot be used to infer a rainfall threshold, nor can it support an early-warning rule. It is more appropriately read as evidence that local P95 exceedance coverage, by itself, is insufficient to discriminate the spatial magnitude of SAR-detected flooding among these nine selected events.

Observation uncertainty further weakens any one-to-one rainfall–flood expectation. Sentinel-1 flood detection depends on acquisition timing relative to the flood peak and recession, radar geometry, roughness, urban backscatter, and flooded vegetation (Chen *et al.,* 2024; Tavus *et al.,* 2022; San Jose *et al.,* 2026). Because the present flood maps are pseudo-label based and lack independent validation across all events, part of the residual variation may reflect observation error. This uncertainty is now treated as a limitation of the response variable rather than as an explanation that can be separated empirically from hydrological processes.

The persistence of this decoupling across two hydrological aggregation scales (Section 3.6) is itself informative. If the weak rainfall–flood association were merely a consequence of forcing and response being measured in mismatched units, then enlarging the unit from Level 9 to Level 8 thereby internalizing more upstream–downstream connectivity within each polygon should have strengthened it. It did not. The direct extreme-fraction effect remained null at both scales, which redirects the explanation away from unit geometry and toward process: routing time scales longer than the event window, antecedent moisture, baseline water bodies, and the timing of SAR acquisition relative to the flood peak.

Landscape variables show more stable directions than extreme_fraction, but the effect sizes remain small. The negative vegetation association is compatible with interception, infiltration, roughness, and storage, while the positive built-up association is compatible with imperviousness and faster runoff generation (Lin *et al.,* 2020; Pal *et al.,* 2022; Hoang & Liou, 2024). However, these mechanisms cannot be isolated from geomorphic position or substrate permeability in the current model. High vegetation fraction may also mark riparian corridors, swamps, or floodplains, and built-up areas may be concentrated in lowlands and river corridors. The results therefore support landscape-sensitive flood response, but not a causal ranking of landscape over topography, hydrology, or rainfall.

The practical significance of the landscape associations is correspondingly modest. A one-standard-deviation increase in vegetation fraction changes the odds of the conditional mean flood fraction by approximately 0.93, while the corresponding value for built-up fraction is approximately 1.09. These effect sizes are meaningful as consistent directional signals across a large panel, but they are not large enough to justify deterministic statements such as 'vegetation prevents flooding' or 'built-up cover controls flooding.' For planning, their value lies in identifying landscape composition as one component of a broader susceptibility context that should be combined with terrain, drainage, and upstream-contributing-area information.

The logistic model nevertheless reveals additional nuance. Built-up fraction remains a stable risk predictor (full-panel OR = 1.136, $p = 0.019$; sensitivity OR = 1.124, $p = 0.030$). Vegetation_fraction, in contrast, is not significant in the logistic model. The difference between fractional and logistic specifications suggests that the mechanism increasing the magnitude of flood_fraction is not identical to the mechanism determining whether a sub-basin crosses the flood_status threshold of 0.01. Vegetation can reduce the magnitude of inundation while still allowing small-area inundation to occur in vegetated sub-basins located within floodplains, swamps, or riparian corridors.

Vegetation should therefore be interpreted with caution. In tropical landscapes, high vegetation_fraction does not necessarily denote upland forest with strong protective hydrological function. High NDVI may equally represent riparian vegetation, swamps, plantations, or lowland areas geomorphologically close to flow pathways or retention zones. The bio-hydromorphological literature shows that riparian and wetland vegetation has a dual role: it attenuates runoff but also marks areas that interact frequently with floodwaters (Maketa *et al.,* 2026; Zhou *et al.,* 2026). The vegetation result here should be interpreted as a context-dependent spatial–hydrological association rather than as a general assertion that vegetation is universally protective.

In contrast, built-up fraction carries a more direct policy interpretation. Although mean builtup_fraction is only 0.040, its effects on flood_fraction and flood_status remain visible. This indicates that even relatively small built-up shares can have hydrological consequences when located in strategic positions such as lowlands, sub-basin outlets, river corridors, or areas with limited drainage. The result is consistent with urban-flood research emphasizing that growth in impervious surface, floodplain development, and inadequate drainage infrastructure elevate flood risk, particularly when infrastructure capacity does not keep pace with land-use change (Bagheri & Liu, 2025; Dharmarathne *et al.,* 2024; Wang *et al.,* 2025).

4.2. Methodological Contribution: Hydrologically Aligned Multi-Sensor Association Analysis

The methodological contribution is deliberately repositioned. CHIRPS, Sentinel-1, Sentinel-2, NDVI, NDBI, and Random Forest are established components of flood research and are not presented here as individually novel. The contribution lies in expressing rainfall, vegetation, built-up cover, and SAR-derived flooding as comparable sub-basin fractions; organizing them as repeated polygon-event observations in HydroBASINS; and testing their associations with event fixed effects and cross-scale sensitivity. This extends a mapping workflow into an inferential framework, while remaining an association study rather than a causal attribution model.

The fractional approach is a central element of the contribution. vegetation_fraction, builtup_fraction, flood_fraction, and extreme_fraction are all defined as proportions of valid pixels meeting specific criteria within each HydroBASINS Level 9 polygon. Parameters are therefore not collapsed into a single category per sub-basin. This preserves internal sub-basin heterogeneity and is well suited to multi-resolution data integration. It also avoids the limitation of binary approaches that would treat a sub-basin with 2% inundation identically to one with 80% inundation as both “flooded”. In risk-management contexts, such intensity differences are substantial because they imply different impacts, intervention priorities, and validation requirements.

Multi-sensor integration also brings uncertainty. Sentinel-1 can detect inundation under cloud, but urban surfaces, flooded vegetation, permanent water, and soil-moisture changes can affect backscatter. Evaluations of SAR datasets indicate that performance may decline when models are applied to regions or events outside the training conditions (San Jose *et al.,* 2026). The Random-Forest approach using conservative pseudo-labels yields consistent flood_probability across events, but should be regarded as an observational indicator rather than a substitute for field validation. Similarly, Sentinel-2 remains sensitive to cloud, as evident from the 2017 event with only 140 complete records. The decision to keep NoData as NoData, and to use a complete-only analytical panel, is a conservative choice to avoid misinterpreting absent observations as absent vegetation, built-up surface, or inundation.

4.3. Implications for Tropical Flood Understanding and Risk Governance

These findings are important for Sumatra because the island combines intense hydrometeorological pressure with rapid landscape transformation. Studies of Sumatra document complex dynamics of forest cover, plantations, fires, and land use, particularly in the lowlands and on economically valuable landscapes (Aso *et al.,* 2024; Candraningrum, 2026; Kartika *et al.,* 2022; Thoha *et al.,* 2024). These changes can alter retention capacity, runoff pathways, and the exposure of settlements to inundation. The finding that built-up fraction is positively associated with flood response provides an empirical foundation for arguing that Sumatran flood risk cannot be managed through rainfall prediction alone; it must be coupled with spatial planning, land-use governance, and protection of landscape hydrological function.

The first practical implication is methodological rather than operational: rainfall indicators should be interpreted together with hydrological position and landscape context. The present analysis does not estimate an alert threshold and should not be used directly as an early-warning model. Because only nine annual major-rainfall dates are analyzed and flood_probability is not independently validated across all events, no defensible answer can be given to the question 'above what rainfall threshold should the population be warned?' Establishing such a threshold would require a substantially larger event inventory that includes both flood and non-flood cases, independent flood labels, multiple rainfall durations and lags, terrain and drainage covariates, and out-of-sample calibration of sensitivity, specificity, and ROC-AUC.

The second implication concerns control of built-up expansion. Increases in built-up fraction do not need to be very large to affect flood odds when they occur in sub-basins with limited drainage or in flow-accumulation pathways. The urban-flood-adaptation literature stresses that growth in impervious surfaces, floodplain development, and undersized drainage infrastructure tend to increase flood risk (Bagheri & Liu, 2025; Dharmarathne *et al.,* 2024; Song *et al.,* 2024). In Sumatra, controlling settlement expansion within river corridors, swamps, lowlands, and sub-basin outlets is a more relevant strategy than expanding drainage capacity after development takes place.

The third implication concerns the interpretation of vegetation. The results do not support a simple generalization that all vegetation reduces flood risk. Upland vegetation, protection forest, riparian vegetation, wetlands, and plantations all have distinct hydrological functions. Vegetation-based policies should therefore differentiate between vegetation as an infiltration and retention provider, vegetation as an indicator of floodplain landscapes, and vegetation as a riparian-ecosystem component to be conserved. This is consistent with nature-based solutions emphasizing location-specific function and ecological connectivity, rather than aggregate vegetation cover alone.

4.4. Limitations and Future Works

Several limitations materially constrain inference. First, the event inventory contains only nine annual area-mean rainfall maxima. This design is useful for repeated-event comparison but restricts rainfall variability and does not support frequency analysis or operational warning calibration. Second, flood_fraction is derived from a Sentinel-1 Random-Forest model trained on conservative pseudo-labels. Independent flood reference data were not available consistently across all nine events, so confusion-matrix metrics, F1-score, kappa, and IoU cannot be reported and the response variable should be regarded as a SAR-derived flood indicator. Third, NDVI $\geq$ 0.40, NDBI $\geq$ 0.00, and flood_probability $\geq$ 0.50 are operational thresholds; although their rationale is now stated, a formal threshold-sensitivity analysis has not yet been performed. Fourth, the final regression omits explicit topographic, drainage, upstream-contributing-area, soil-permeability, and bedrock-permeability covariates. Elevation and slope are used in the SAR classification workflow, but this does not control for geomorphic confounding in the statistical model. Fifth, the current results emphasize coefficient estimates and cross-event/cross-scale robustness; a full predictive-diagnostic suite, including pseudo-R^2, AIC/BIC comparisons and ROC-AUC for the binary model, requires re-estimation from the analytical panel and should be completed before any predictive or warning application is claimed.

These limitations define the next analytical steps. The highest priority is independent validation of flood extent for a representative subset of events using authoritative flood maps, high-resolution imagery, or field/reference observations, followed by sensitivity testing of the flood-probability threshold. The inferential model should then be expanded with elevation, slope, relative elevation, flow accumulation, distance to river, drainage density, TWI, upstream contributing area, and available soil or lithological permeability indicators to distinguish landscape composition from geomorphic position. Rainfall forcing should be tested across rx1day, rx3day, rx5day, rx7day, antecedent rainfall, and lagged accumulations. Finally, the event inventory should be expanded beyond annual maxima to include a larger range of flood and non-flood dates, enabling out-of-sample diagnostics, ROC-AUC calibration, and explicit evaluation of whether the framework has operational warning value.

5. Conclusion

This study evaluates how vegetation fraction, built-up fraction, and local extreme-rainfall coverage are associated with a Sentinel-1-derived flood fraction across HydroBASINS sub-basins in Sumatra. Within the variables evaluated, landscape composition shows more stable associations with flood_fraction than local P95 exceedance coverage, but the analysis does not establish landscape as the primary physical control on flooding. The full panel comprises 14,409 polygon-event records, with 11,818 complete observations used for the principal analysis. Flood response is highly heterogeneous and generally low in areal fraction, with a small subset of sub-basins showing much larger inundation indicators.

The first main finding is that extreme_fraction shows little bivariate association with flood_fraction and no robust direct effect in the event-fixed-effect models. Because the study is conditioned on nine major annual rainfall events, this result is consistent with limited discriminatory power of local P95 exceedance once rainfall is already extreme at the event-selection scale. It does not imply that rainfall is unimportant and cannot be converted into an operational warning threshold from the present sample.

The weak direct extreme-fraction association persists when the analysis is repeated at HydroBASINS Level 8, while the directions of the vegetation and built-up coefficients are broadly stable. This cross-scale consistency strengthens confidence that the reported associations are not unique to one HydroBASINS level, but it does not prove superiority over grid or administrative units and does not eliminate omitted-variable bias from topography, drainage, or substrate permeability.

The second main finding is that vegetation_fraction is negatively associated with flood_fraction in the fractional model, whereas builtup_fraction is positively associated and more consistent across fractional and logistic specifications. These effects are statistically detectable but modest. They should be interpreted as conditional associations among the included variables, not as evidence that vegetation or urbanization alone determines flooding. The role of vegetation is particularly context dependent because high NDVI can represent uplands, plantations, riparian zones, wetlands, or naturally flood-prone lowlands.

Methodologically, the principal contribution is the combination of fractional multi-sensor variables with hydrologically aligned sub-basin units and repeated-event statistical analysis. The framework preserves within-sub-basin heterogeneity and offers a reproducible structure for comparing rainfall, landscape composition, and SAR-derived flood response. Its contribution is therefore integrative and inferential rather than the introduction of a new sensor, index, or classifier.

For flood-risk management, the results support combining rainfall monitoring with landscape, terrain, drainage, and upstream-connectivity information rather than relying on rainfall exceedance alone. The present study should not be interpreted as an early-warning model and does not provide a warning threshold. Before operational use, the framework requires independent flood validation, threshold-sensitivity analysis, a larger event inventory, inclusion of topographic and hydrological controls, and predictive diagnostics using out-of-sample data.

Acknowledgements

MY acknowledges the Research Center for Nuclear Material and Radioactive Waste Technology (PRTBNLR), Research Organization for Nuclear Energy (ORTN), National Research and Innovation Agency (BRIN), Indonesia to permit to conduct postdoctoral at the Interdisciplinary Research Center for Industrial Nuclear Energy (IRC-INE), King Fahd University of Petroleum and Minerals (KFUPM) as part of the institution's capacity building.

Author Contributions

Conceptualization: Priyono, K. D., Susilawati, S. A., Sari, D. N., Sunariya, M. I., Priyana, Y., Pragata, A. A., Rohman, A., Yusuf, M., Ibrahim, M. H., Sattara, F., & Nawaz, M.; **methodology**: Priyono, K. D., Susilawati, S. A.; **investigation**: Priyono, K. D., **writing—original draft preparation**: Priyono, K. D., Yusuf, M.,; **writing—review and editing**: Priyono, K. D., Yusuf, M., Ibrahim, M. H.; **visualization**: Priyono, K. D., Susilawati, S. A., Sari, D. N., Sunariya, M. I., Priyana, Y. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest.

Generative AI Declaration

During the preparation of this work, the authors use ChatGPT to develop the analytical code.

Data availability

The data collection and analysis code is available at https://ums.id/sumatra_code.

Funding

This research funded by Universitas Muhammadiyah Surakarta trough International Research Partnership (RKI) Scheme No. 187.12/DRPPS/A3.9/IV/2026.

References

Allayorov, D., Rakhimov, K., Umarov, U., Atakulov, D., Allayorova, D., & Shaymardonov, S. (2026). Climate Change and Flooding: Bibliometric Analysis to Identify Future Research. *Forum Geografi,* 40, 1, 147-160.

Amer, R. (2025). Machine learning-driven rapid flood mapping for Tropical Storm Imelda using Sentinel-1 SAR imagery. *Remote Sensing,* 17(11), 1869. https://doi.org/10.3390/rs17111869

Aso, O. P., Muhsandi, K., & Gustiawan, I. (2024). The land use history and origin of land status oil palm plantations in South Sumatera. *IOP Conference Series: Earth and Environmental Science,* 1407, 012012. doi: 10.1088/1755-1315/1407/1/012012

Atmaja, T., & Lee, D.-J. (2026). Integrative statistical analysis of compound flood dynamics under climate variability in China's coastal zone. *Journal of Hydrology,* 671, 135284. doi: 10.1016/j.jhydrol.2026.135284

Bagheri, A., & Liu, G.-J. (2025). Climate change and urban flooding: Assessing remote sensing data and flood modeling techniques: A comprehensive review. *Environmental Reviews*. doi: 10.1139/er-2024-0065

Bazai, N. A., Alam, M., Cui, P., Hao, W., Khan, A. P., Waseem, M., Shunyu, Y., Ramzan, M., Wanhong, L., & Ahmed, T. (2025). Dynamics and impacts of monsoon-induced geological hazards: A 2022 flood study along the Swat River in Pakistan. *Natural Hazards and Earth System Sciences,* 25, 1071–1090. doi: 10.5194/nhess-25-1071-2025

Broich, M., Hansen, M., Stolle, F., Potapov, P., Margono, B. A., & Adusei, B. (2011). Remotely sensed forest cover loss shows high spatial and temporal variation across Sumatera and Kalimantan, Indonesia 2000–2008. *Environmental Research Letters,* 6(1), 014010.

Candraningrum, D. (2026). Floods, elephants, and tropical forest change in Sumatra. *Forests, Trees and Livelihoods*. doi: 10.1080/14728028.2026.2627644

Chen, F., Zhang, M., Zhao, H., Guan, W., & Yang, A. (2024). Pakistan's 2022 floods: Spatial distribution, causes and future trends from Sentinel-1 SAR observations. *Remote Sensing of Environment,* 315, 114055. doi: 10.1016/j.rse.2024.114055

Chen, X., Li, H., Yu, H., Hou, E., Song, S., Shi, H., & Chai, Y. (2025). Counterfactual analysis of extreme events in urban flooding scenarios. *Journal of Hydrology: Regional Studies,* 58, 102166. doi: 10.1016/j.ejrh.2024.102166

Chrysanti, A., & Son, S. (2025). Unraveling atmosphere and surface boundary interactions behind extreme tropical rainfall: A case study in Indonesia using fully coupled atmosphere–hydrology simulations. *Journal of Advances in Modeling Earth Systems.* doi: 10.1029/2024MS004730

Dharmarathne, G., Waduge, A. O., Bogahawaththa, M., Rathnayake, U., & Meddage, D. P. P. (2024). Adapting cities to the surge: A comprehensive review of climate-induced urban flooding. *Results in Engineering.* doi: 10.1016/j.rineng.2024.102123

Fu, L.-L., Pavelsky, T., Crétaux, J.-F., Morrow, R., Farrar, J. T., Vaze, P., Sengenes, P., Vinogradova-Shiffer, N., Sylvestre-Baron, A., Picot, N., & Dibarboure, G. (2024). The Surface Water and Ocean Topography Mission: A breakthrough in radar remote sensing of the ocean and land surface water. *Geophysical Research Letters.* doi: 10.1029/2023GL107652

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data,* 2, 150066. doi: 10.1038/sdata.2015.66

Gulshad, K., Yaseen, A., & Szydlowski, M. (2024). From data to decision: Interpretable machine learning for predicting flood susceptibility in Gdansk. *Remote Sensing,* 16(20), 3902. doi: 10.3390/rs16203902

Hoang, D.-V., & Liou, Y.-A. (2024). Assessing the influence of human activities on flash flood susceptibility in mountainous regions of Vietnam. *Ecological Indicators,* 158, 111417. doi: 10.1016/j.ecolind.2023.111417

Huning, L. S., Margulis, S. A., Guan, B., Waliser, D. E., & Aghakouchak, A. (2024). *Compound flooding from atmospheric rivers and intensified hydrological cycles.* Nature Reviews Earth & Environment.

Ibrahim, A., Wageeh, A., Hamouda, M. A., Ahmed, A., & Gad, A. (2026). Impact of LULC classification methods on runoff simulation in an arid mountainous watershed using remote sensing and machine learning. *Earth,* 7(1), 26. doi: 10.3390/earth7010026

Jumadi, J., Sari, D. N., Umrotun, U., Musiyam, M., Nurmantyo, C., Muhammad, S. F., & Ibrahim, M. H. (2024). Remote Sensing And Gis-Driven Model For Flood Susceptibility Assessment In The Upper Solo River Watershed. *Geographia Technica,* 19(2)

Kajari, B., Tobak, Z., Turi, N., Bozan, C., & Van Leeuwen, B. (2024). Prediction of inland excess water inundations using machine learning algorithms. *Water*, 16(9), 1267. doi: 10.3390/w16091267

Kartika, T., Kushardono, D., Setiawan, Y., Ibrahim, A., Kustiyo, Sutanto, A., Noviar, H., Khomarudin, M. R., Carolita, I., Tosiani, A., Ubay, M. S. N., Taufik, M., & Al-Khudri, A. (2022). Identification method of vegetation cover changes derived from mosaic Landsat-8 data: Case study Sumatera, Kalimantan, and Java Island. *IOP Conference Series: Earth and Environmental Science*, 1109, 012060. doi: 10.1088/1755-1315/1109/1/012060

Kusratmoko, E., Albiyanti, R. H. M., & Marko, K. (2017). Modelling rainfall-runoff process using HEC-HMS: A case study in Ciliwung-Cisadane River Basin. *IOP Conference Series: Earth and Environmental Science,* 79, 012023.

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrological Processes,* 27(15), 2171–2186. doi: 10.1002/hyp.9740

Lin, W., Sun, Y., Nijhuis, S., & Wang, Z. (2020). Scenario-based flood risk assessment for urbanizing deltas using future land-use simulation (FLUS): Guangzhou Metropolitan Area as a case study. *Science of the Total Environment,* 739, 139899. Doi: 10.1016/j.scitotenv.2020.139899

Liu, Y., Samat, A., Du, P., Chen, J., Abuduwaili, J., Luo, K., Zhu, E., & Shokparova, D. (2026). High-resolution spatio-temporal analysis and driver attribution of floods in Kazakhstan using SHAP and remote sensing integration. *Climate Risk Management,* 52, 100783. doi: 10.1016/j.crm.2025.100783

Llasat, M. C., Marcos-Matamoros, R., Pascual, R., Rigo, T., Insua-Costa, D., & Crespo-Otero, A. (2025). Western Mediterranean flash floods through the lens of Alcanar (NE Iberian Peninsula): Meteorological drivers and trends. *Atmospheric Research*. doi: 10.1016/j.atmosres.2025.108266

Maketa, M. J., Esler, K. J., Roffe, S. J., & Rebelo, A. J. (2026). Investigating the impact of severe floods on biogeomorphic change in wetlands using remote sensing. *Wetlands,* 46, 24. doi: 10.1007/s13157-026-02043-w

Mantovani, J. R., Alcantara, E., Baiao, C. F., Pampuch, L., Curtarelli, M. P., Ribeiro, J. V. M., Guimaraes, Y. C., Londe, L., Massi, K., Marengo, J., Nobre, C. A., Assireu, A. T., Bortolozo, C. A., Simoes, S. J., Tomasella, J., Leal de Moraes, O. L., & Park, E. (2025). Unprecedented flooding in Porto Alegre Metropolitan Region (Southern Brazil) in May 2024: Causes, risks, and impacts. *Journal of South American Earth Sciences,* 160, 105533. doi: 10.1016/j.jsames.2025.105533

Musiyam, M., Jumadi, J., Fikriyah, V.N., Masruroh, H., Septiyani, E.D., Amin, C., Hadibasyir, H.Z., Sattar, F., and Nawaz, M.. (2025). Flood Risk Analysis Using Spatial Synthetic Population in the Upper Bengawan Solo Watershed, Indonesia. *Forum Geografi,* 39, 3, 383-397.

Nasri, N., Nursaputra, M., Arif, A. R., Dwiyani, R., Chania, L., Agastya, I. B. O., & Arif, D. A. A. (2025). Ecological Disaster in Northern Sumatra: How Extreme Rainfall and Land-Cover Disturbance Triggered Widespread Flash Flooding. *Jurnal Penelitian Kehutanan Wallacea*, *14*(2), 111-121.

Pal, S. C., Chowdhuri, I., Das, B., Chakrabortty, R., Roy, P., Saha, A., & Shit, M. (2022). Threats of climate change and land use patterns enhance the susceptibility of future floods in India. *Journal of Environmental Management,* 305, 114317. doi: 10.1016/j.jenvman.2021.114317

Petiangma, M., Kountouris, Y., & Macpherson, R. (2025). Multi-day rainfall accumulation and the spatial structure of flood damage in tropical settings. *Hydrology and Earth System Sciences.* doi: 10.5194/hess-2024-345

Pizzorni, M., Innocenti, A., & Tollin, N. (2024). Droughts and floods in a changing climate and implications for multi-hazard urban planning: A review. *City and Environment Interactions.* doi: 10.1016/j.cacint.2024.100169

Rahman, M. T., Ibukun, J. A., Olubaju, A. E., Sodipo, E. O., & Oyetunji, S. O. (2026). Integrating SAR remote sensing and machine learning for comprehensive flood risk assessment in Kogi State, Nigeria. *Journal of African Earth Sciences*, 241, 106168. doi: 10.1016/j.jafrearsci.2026.106168

Riazi, M., Khosravi, K., Shahedi, K., Ahmad, S., Jun, C., Bateni, S. M., & Kazakis, N. (2023). Enhancing flood susceptibility modeling using multi-temporal SAR images, CHIRPS data, and hybrid machine learning algorithms. *Science of the Total Environment,* 871, 162066. doi: 10.1016/j.scitotenv.2023.162066

San Jose, I. K., Wiguna, S., Kametaka, R., Adriano, B., Mas, E., & Koshimura, S. (2026). Evaluation of open-source SAR-based flood datasets for flood extent mapping in emergency settings. *Progress in Disaster Science*, 29, 100507. doi: 10.1016/j.pdisas.2025.100507

Saufina, E., Trismidianto, Satiadi, D., Harjupa, W., Risyanto, Purwaningsih, A., Fathrio, I., Praja, A. S., Juaeni, I., Witono, A., Rahmatia, F., Pratama, R., Putri, M. S., & Wulandari, P. (2025). The role of Madden–Julian Oscillation, Westerly Wind Bursts, and Kelvin Waves in triggering extreme rainfall through Mesoscale Convective Systems: A case study of West Sumatra, March 7–8, 2024. *Atmospheric Research,* 318, 107993. doi: 10.1016/j.atmosres.2025.107993

Shampa, Nasir, N. N., Winey, M. M., Dey, S., Zahid, S. M. T., Tasnim, Z., Islam, A. K. M. S., & Hussain, M. A. (2025). Integration of remote sensing and machine learning approaches for operational flood monitoring. Water, 17(15), 2189. doi: 10.3390/w17152189

Singha, C., Rana, V. K., Pham, Q. B., Nguyen, D. C., & Lupikasza, E. (2024). Integrating machine learning and geospatial data analysis for comprehensive flood hazard assessment. *Environmental Science and Pollution Research.* doi: 10.1007/s11356-024-34286-7

Song, J., Shao, Z., Zhan, Z., & Chen, L. (2024). State-of-the-art techniques for real-time monitoring of urban flooding: A review. *Water,* 16(17), 2476. doi: 10.3390/w16172476

Sun, Q., Fang, J., Dang, X., Xu, K., Fang, Y., Li, X., & Liu, M. (2022). Multi-scenario urban flood risk assessment by integrating future land use change models and hydrodynamic models. *Natural Hazards and Earth System Sciences,* 22, 3815–3835. doi: 10.5194/nhess-22-3815-2022

Syamsudin, F., Lestari, S., Awaludin, A., Siswanto, Yulihastin, E., Marpaung, F., Sulistyowati, R., Sinatra, T., & Purbantoro, B. (2026). Contributions of local and large-scale climate drivers on the extreme rainfall events in the Megacity of Jakarta. *Atmospheric Research*. Doi: 10.1016/j.atmosres.2025.108537

Tavus, B., Can, R., & Kocaman, S. (2022). A CNN-based flood mapping approach using Sentinel-1 data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-3-2022, 549–556. doi: 10.5194/isprs-annals-V-3-2022-549-2022

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., & Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature,* 596, 80–86. doi: 10.1038/s41586-021-03695-w

Thoha, A. S., Nasution, Y. R. P., Bahar, Y., & Saharjo, B. H. (2024). Land use/land cover change in protected forest of Tangkahan, North Sumatra, Indonesia. *IOP Conference Series: Earth and Environmental Science,* 1352, 012044.

Ulloa, N. I., Yun, S.-H., Chiang, S.-H., Furuta, R., & Datta, K. (2022). Sentinel-1 spatiotemporal simulation using convolutional LSTM for flood mapping. *Remote Sensing,* 14(2), 246.

Wang, J., Zhao, M., Xu, M., Li, Y., & Gou, A. (2025). Effect of land cover types evolution in megacities on flood regulation capacity: The case of Zhengzhou since 1990. *Natural Hazards*. doi: 10.1007/s11069-024-06915-4

Yao, J., Wang, M., Xu, N., Liu, T., Mo, F., Li, X., Cao, Y., Chen, K., Sun, J., & Lu, H. (2025). 3-D flood mapping from SWOT observations during extreme rainfall: A case study of Gangnan Reservoir. *International Journal of Digital Earth.* doi: 10.1080/17538947.2025.2544916

Zhang, Y., Wu, L., Ren, H., Liu, Y., Zheng, Y., Liu, Y., & Dong, J. (2024). *Compound flooding: A bibliometric review and an integrated risk assessment framework.* Environmental Modelling & Software.

Zhou, Y., Toda, Y., & Zhu, R. (2026). Spatiotemporal dynamics of Sentinel-2 NDVI as indicators of bio-hydromorphological interactions: Implications for river management. *Earth Surface Dynamics,* 14, 233–246. doi: 10.5194/esurf-14-233-2026

Zittis, G., Bruggeman, A., & Lelieveld, J. (2022). Revisiting future extreme precipitation trends in the Mediterranean. *Weather and Climate Extremes*. doi: 10.1016/j.wace.2021.100380