water

MDPI

Article

Scenario-Based Diagnosis of Urban Flood-System Responses with FLOOD-ST: A TELEMAC-2D Case Study

Jens Reinert [1,*], David Jonathan Roschinsky [1], Julian Hofmann [2], Adrian Almoradie [3] and Catrina Brüll [1]

[1] Institute of Hydraulic Engineering and Water Resources Management, RWTH Aachen University, 52074 Aachen, Germany; david.roschinsky@rwth-aachen.de (D.J.R.); bruell@iww.rwth-aachen.de (C.B.)
[2] FloodWaive Predictive Intelligence GmbH, 52070 Aachen, Germany; hofmann@floodwaive.de
[3] Emschergenossenschaft und Lippeverband, 45128 Essen, Germany; almoradie.adrian@eglv.de
* Correspondence: reinert@iww.rwth-aachen.de; Tel.: +49-241-80-25747

Abstract

Conventional urban flood assessments primarily characterise inundation under predefined events but provide limited information on how controlled changes in hydrometeorological conditions, antecedent catchment state, and infrastructure operation alter the spatial and temporal response of a flood system. This study operationalises Flood-Oriented Operational Stress Testing (FLOOD-ST) in the Roßbach sub-catchment of the Emscher catchment, Germany, using an established TELEMAC-2D rain-on-grid model. A deliberately selected ensemble of 23 scenarios was evaluated through 45 direct matched contrasts in which one investigated condition changes while the remaining scenario settings are held constant. Flood-system response was assessed using inundation extent together with hydraulic and affectedness-based indicators for streets, buildings, and critical infrastructure, including maximum simultaneous and cumulative affectedness, timing, recurrence, duration, and spatial redistribution. Within the investigated scenario space, changes in hydrometeorological event type and antecedent moisture condition produced comparatively pronounced differences in affectedness magnitude and timing; for example, changing from the stationary cloudburst to the moving convective storm increased cumulative affectedness by 14.90% for streets, 15.52% for critical infrastructure, and 5.63% for buildings. Infrastructure-state perturbations generally produced smaller aggregate differences but could substantially redistribute affectedness in space and time. In the BL05–CS11 matched clogging-state contrast, net inundation increased by 34,308 m^2 while gross spatial turnover reached 73,104 m^2; at building level, a net increase of 25 affected objects coincided with a gross turnover of 61 buildings. The case study demonstrates the diagnostic value of operationalising FLOOD-ST beyond conventional inundation-area comparison by identifying where, when, and for how long affectedness changes across controlled scenarios. These outputs can support the prioritisation of locations and system states for further inspection, operational review, and detailed risk analysis, while the findings remain conditional on the investigated scenario ensemble and model configuration.

Keywords: flood stress testing; scenario-based diagnostics; urban flooding; TELEMAC-2D; hydraulic affectedness; spatial redistribution

Check for updates

Academic Editor: Daniela Gogoase Nistoran

Received: 25 August 2026
Revised: 15 September 2026
Accepted: 23 September 2026
Published: 19 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Flood risk management in urban catchments requires accounting for flood processes that result from the interaction of hydrometeorological event characteristics, antecedent

catchment conditions, and the state of technical infrastructure [1–5]. This is particularly relevant for short-duration heavy-rainfall events, for which spatially heterogeneous precipitation and short catchment response times can lead to rapidly changing hydraulic conditions [4,5]. Reconstructions of the July 2021 flood event in western Germany and neighbouring regions have shown that severe consequences were not determined by precipitation magnitude alone, but developed through the interaction of rainfall-event characteristics, antecedent catchment conditions, exposure, and infrastructure-related factors [1–3]. For densely built urban systems, this interaction creates a need to assess not only the resulting inundation pattern under a reference condition, but also how the hydraulic system response changes under varying hydrometeorological conditions and infrastructure states.

The relevance of these interactions is further increased by the variability and possible intensification of hydrometeorological extremes. Climate projections indicate changes in severe convective environments and precipitation extremes across Europe, while the flood response to increasing rainfall extremes can vary nonlinearly depending on catchment and moisture conditions [6–8]. At the event scale, flood generation may therefore depend not only on the magnitude of individual drivers, but also on their joint occurrence and interaction. Compound-event research has shown that dependencies among meteorological and hydrological drivers can substantially alter hazard and risk estimates compared with assessments that treat these drivers independently [9,10]. In the context of the established FLOOD-ST [11], such relevant hydrometeorological, catchment-related, and infrastructure-related conditions are represented as stressors that can be varied systematically. A diagnostic assessment should therefore allow controlled combinations of these stressors to be compared rather than relying on a single reference condition.

Regulatory flood-hazard assessment in Europe is primarily organised around predefined probability- or design-based scenarios that provide standardised reference conditions for spatial planning, risk communication, and flood-risk management [12]. Such hazard products characterise the expected spatial extent and, depending on the product, hydraulic characteristics of flooding under specified scenarios. However, they do not systematically examine how the same flood system responds when antecedent catchment conditions, conveyance capacity, or operational infrastructure states vary. Consequently, established hazard assessment and diagnostic stress testing address complementary questions: the former provides standardised reference conditions, whereas the latter investigates how the hydraulic system responds when selected hydrometeorological and infrastructure-related conditions deviate from those reference states.

Two-dimensional hydrodynamic modelling and increasingly detailed spatial and rainfall information provide the technical basis for analysing urban flood processes at high spatial and temporal resolution [4,5]. Previous work for the Roßbach catchment in the Emscher catchment area in Germany, established a TELEMAC-2D rain-on-grid model that explicitly represents relevant surface characteristics, hydraulic structures, and operational elements [13]. However, increasing model detail alone does not ensure that differences between scenarios can be interpreted unambiguously. Model evaluation, antecedent catchment conditions, scenario definition, and the choice of response indicators all influence the resulting interpretation [14,15].

Existing compound-flood and multi-driver modelling studies demonstrate that interactions among hydrological processes and event characteristics can substantially modify flood response [16–18]. Similarly, studies of critical infrastructure show that flood consequences depend not only on inundation magnitude but also on the spatial distribution and functional context of exposed assets [19]. These approaches provide important evidence on multi-driver flood processes, but they do not generally focus on a controlled

diagnostic comparison of hydrometeorological conditions, antecedent catchment states, and operational infrastructure states within one fixed urban hydraulic model.

FLOOD-ST was introduced to address this diagnostic perspective by transferring the general principle of stress testing to flood-system analysis [11,20–22]. Rather than replacing established hazard assessment, the framework systematically varies selected stressors while maintaining transparent model and evaluation conditions. The generic framework and the underlying Roßbach hydraulic model have therefore already been developed independently. The remaining application gap is to demonstrate how both can be combined in a traceable hydrodynamic case study and what additional diagnostic information is obtained from controlled scenario contrasts, object- and time-resolved affectedness indicators, and spatial redistribution analysis.

Accordingly, the contribution of the present study is not the development of a new hydraulic model or a new global sensitivity method. Instead, it operationalises FLOOD-ST as a scenario-based diagnostic experiment within an established TELEMAC-2D model. The application links a documented scenario matrix with admissible matched contrasts and consistently derived hydraulic and affectedness-based response indicators. This enables the analysis of changes in magnitude, timing, duration, recurrence, and spatial redistribution while explicitly restricting interpretation to the investigated scenario space.

Accordingly, the study addresses three research questions:

1. How do changes in hydrometeorological event characteristics, antecedent moisture conditions, and infrastructure states alter the magnitude, timing, and duration of flood-system responses within the investigated scenario space?
2. To what extent can comparatively small aggregate changes between matched scenarios conceal spatial and temporal redistribution of affectedness across streets, buildings, and critical infrastructure?
3. What additional diagnostic information can cross-scenario recurrence, duration, and consistently unaffected assets provide beyond conventional inundation-extent comparison?

2. Materials and Methods

2.1. Conceptual Application Workflow and Scope

This study combines two previously developed components: the established TELEMAC-2D rain-on-grid model of the Roßbach sub-catchment in the Emscher catchment, Germany [13], and the generic FLOOD-ST framework [11]. The present study neither develops a new hydraulic model nor extends the generic framework. Instead, it operationalises both components in a controlled case-study application to examine how changes in hydrometeorological event characteristics, antecedent catchment conditions, and infrastructure states are reflected in spatially and temporally resolved flood-system responses.

The analysis follows a scenario-based diagnostic stress-testing design rather than a full-factorial experiment or a global sensitivity analysis. The 23 simulations represent a deliberately bounded scenario ensemble. Comparisons between scenarios are restricted to matched contrasts in which one investigated stressor dimension differs while the remaining scenario dimensions, hydraulic model configuration, numerical settings, and post-processing rules are kept constant.

Within this study, stressors are defined as controlled changes in hydrometeorological conditions, catchment state, or infrastructure state that may alter the magnitude, timing, duration, or spatial distribution of the hydraulic system response. The diagnostic workflow comprises following principal steps:

- Definition of the investigated stressor space and scenario matrix;
- Identification of admissible matched contrasts;

- Consistent extraction of hydraulic and affectedness-based response indicators from all simulations;
- Interpretation of response differences within the investigated scenario combinations and perturbation levels.

The overall application workflow is summarised schematically in Figure 1. The approach provides case-study-based evidence of relative flood-system response within the investigated stress-test definition set. It does not aim to estimate global variance contributions, derive a statistically representative ranking of stressor importance, or determine the probability of occurrence of individual stressor combinations. All interpretations remain conditional on the investigated scenarios, perturbation levels, underlying hydraulic model, and applied post-processing definitions.

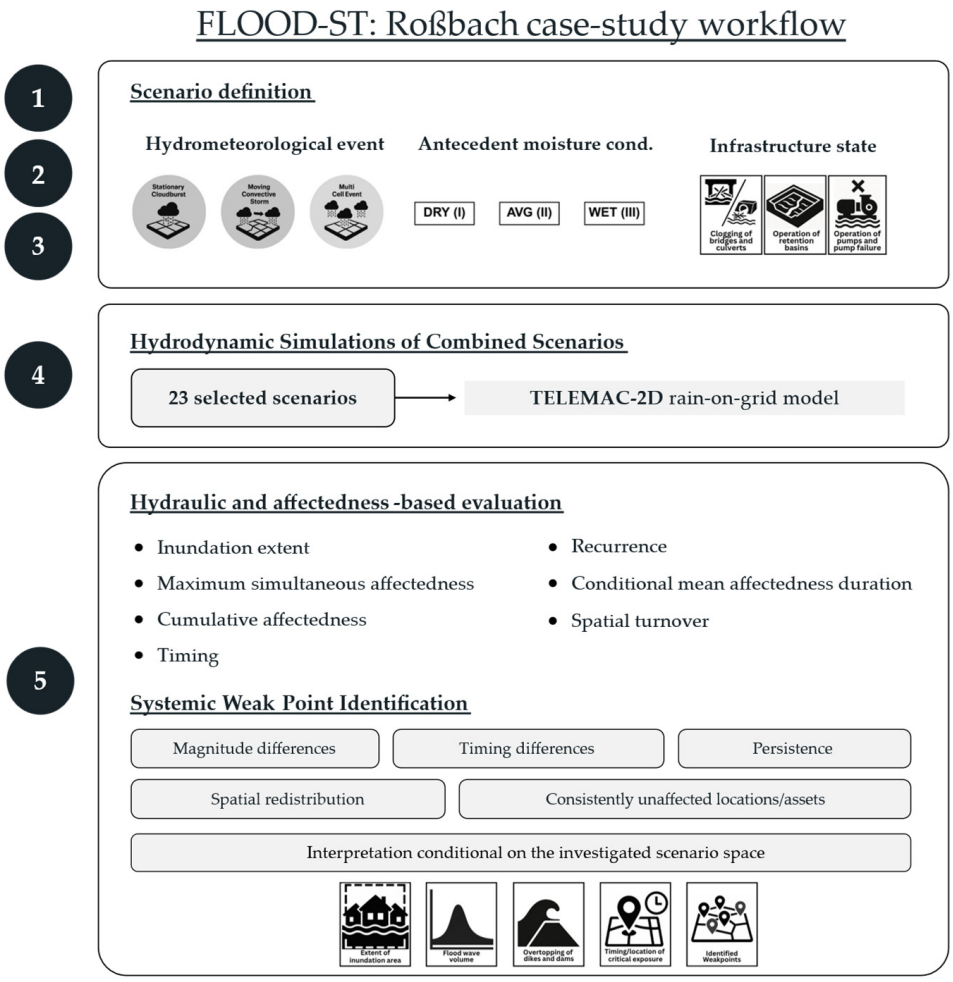


Figure 1. Simplified conceptual workflow of the FLOOD-ST application in the Roßbach case study, illustrating the progression from scenario definition and hydrodynamic simulation to affectedness-based evaluation and diagnostic interpretation. Adapted and extended from the generic FLOOD-ST framework presented in [11].

2.2. Study Area and Analysed Asset Classes

The case study is located in the Roßbach sub-catchment of the Emscher river basin in Germany (Figure 2). The Roßbach catchment covers approximately 30.88 km^2 and is divided into five hydraulic model sub-units. The present diagnostic analysis focuses on the urbanised downstream sub-unit 29000 (Figure 3), which contains the infrastructure elements varied in the stress-test scenarios, including selected culvert and bridge crossings, pumping stations, and the Martener Straße flood-detention basin (dt. Hochwasserrück-

haltebecken; HRB). Geospatial processing, spatial analysis, and map preparation were performed using QGIS (Version 3.28.4 and v3.40.5; QGIS.org, Open Source Geospatial Foundation).

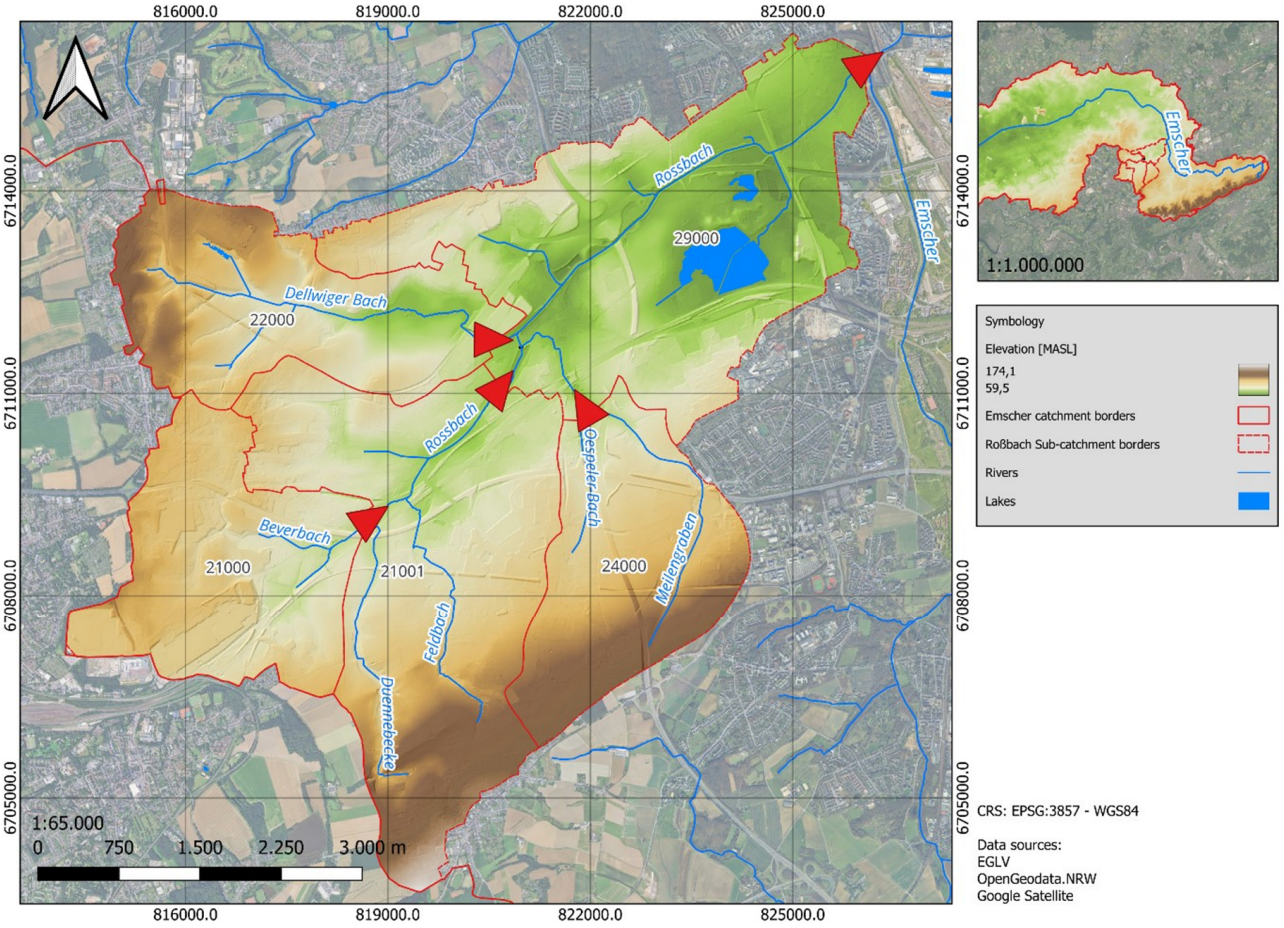


Figure 2. Overview of the Roßbach catchment in the upper Emscher river basin and its five model sub-units. Red triangles indicate the hydraulic coupling points at which scenario-specific discharge hydrographs are transferred between adjacent model sub-units.

Three asset classes were considered for the subsequent affectedness analysis: streets, buildings, and critical infrastructure. Building geometries and the road-network information were derived from the official ATKIS geodata provided through OpenGeodata.NRW. The building dataset comprises 8273 objects within the analysed area. Critical infrastructure was identified manually using Google Maps (version 2025, Google LLC, Mountain View, CA, USA) and subsequently compiled into a dedicated geospatial dataset. The final critical-infrastructure dataset comprises 58 functionally relevant objects, including police and fire-service facilities, administrative buildings, energy-supply infrastructure, schools, kindergartens, and care facilities. These three asset classes form the basis for the area- and object-based affectedness indicators described in Section 2.5. Supplementary Table S7 provides further dataset provenance, processing information, and known metadata limitations.

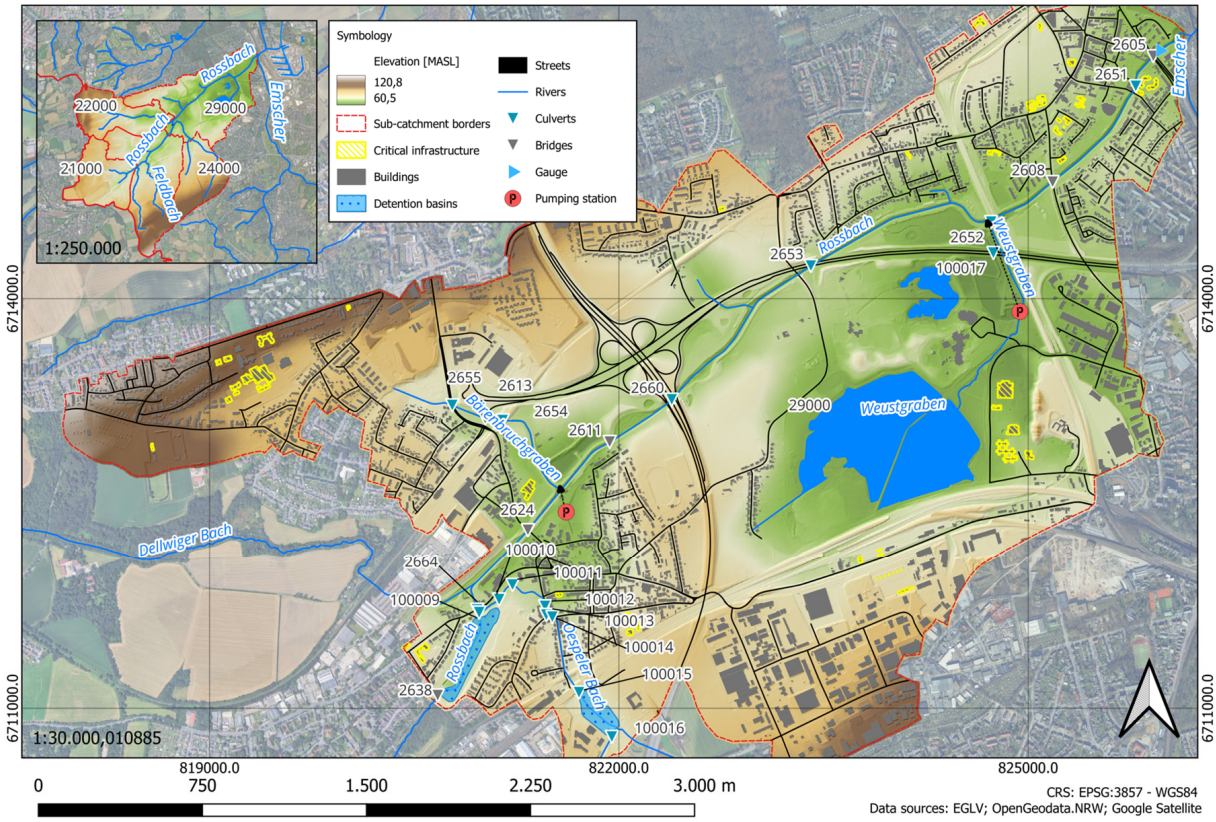


Figure 3. Location and extent of model sub-unit 29000 within the Roßbach sub-catchment.

2.3. Hydraulic Model Basis

The hydrodynamic simulations were conducted with TELEMAC-2D v8.4 using the established rain-on-grid model of the Roßbach catchment [13]. Table 1 provides a detailed summary. Pre-processing and post-processing were performed using Python (v3.12.4 and v3.12.2, respectively; Python Software Foundation, Beaverton, OR, USA). Numerical and spatial analyses were conducted using NumPy v2.2.4, pandas v2.2.3, GeoPandas v1.0.1, Shapely v2.1.1, Rasterio v1.4.3, and GDAL v3.9.3.

The model solves the depth-averaged shallow-water equations on an unstructured triangular mesh. The model topography is based on the 2018 Digital Terrain Model DGM1 provided by Geobasis NRW, Bezirksregierung Köln, with a native raster resolution of 1 m and derived from airborne laser scanning data, while land-use and building information were obtained from OpenGeodata.NRW and the ATKIS database. Land-use classes were used to assign spatially distributed Manning roughness coefficients and Curve Numbers for the SCS-CN runoff representation. Buildings are represented as elevated blocks within the model topography.

The stress-test configuration builds on the previously developed Roßbach TELEMAC-2D model [13]. However, the configuration used here contains stress-test-specific modifications to rainfall inputs, boundary conditions, and hydraulic and operational structures. Quantitative evaluation statistics reported for earlier model-development configurations are therefore not transferred directly to the present stress-test setup. The simulations analysed here were not individually calibrated or recalibrated; instead, the study evaluates relative response differences within a consistently applied model configuration.

Antecedent moisture conditions are represented through the AMC classes of the SCS-CN formulation, with AMC I, II, and III representing dry, average, and wet conditions, respectively. Their implementation within the scenario design is described in Section 2.4.

Table 1. Model discretisation, numerics, and settings used in the Roßbach case study.

Parameter	Specification
Hydrodynamic engine	TELEMAC-2D v8.4 (rain-on-grid)
Mesh type	Unstructured triangular mesh
Runtime computational mesh	37,937 points; 75,155 triangular elements
Manning roughness	Spatially distributed by ATKIS land-use class; n = 0.03–0.13 s m$^{-1/3}$
SCS-CN parameterisation	Spatially distributed by ATKIS land-use class; land-use-specific CN values range from 55 to 100
Time-step control	Variable time step; maximum Courant number 0.9
Linear solver	Conjugate gradient; tolerance 1.0×10^{-4}
Simulation duration per scenario run	12 h
Paired output time steps per scenario run	145 depth-velocity pairs
Rainfall event definition	Synthetic raster sequences representing the stationary cloudburst (SCB), multi-cell event (MCE), and moving convective storm (MCS) archetypes
Catchment-mean precipitation depth	120 mm over valid catchment cells for each trigger after volume correction
Temporal resolution of rainfall forcing	37 rainfall states from 0 to 180 min with 5 min intervals
Spatial resolution of rainfall forcing	100 m × 100 m; 6004 cells total; 3056 valid cells in Roßbach
Upstream boundary treatment	Hydrographs from upstream subcatchments inserted as discharge inflows
Downstream boundary treatment	Fixed downstream boundary condition; unchanged across scenarios

The stress-test scenarios use three synthetic hydrometeorological event types defined within FLOOD-ST: a stationary cloudburst (SCB), a multi-cell event (MCE), and a moving convective storm (MCS). The rainfall distributions are represented on a 100 m × 100 m grid comprising 6004 cells, of which 3056 intersect the Roßbach study area. Each event comprises 37 rainfall states at 5 min spacing and spans three hours. The events were volume-corrected to a common catchment-mean precipitation depth of approximately 120 mm so that the investigated differences primarily represent changes in spatial and temporal rainfall structure rather than total catchment rainfall volume (Figures 4–6).

All stress-test simulations were evaluated over an approximately 12 h observation period using a common temporal output interval. This duration reflects the study's focus on flash-flood-type responses, which are characterised by rapid hydrological response following short-duration, high-intensity rainfall. The synthetic rainfall events themselves extend over approximately three hours, so that the 12 h observation window covers both the direct event response and several subsequent hours of hydraulic recession and redistribution. The same observation period was applied to all scenarios to ensure consistent comparison of peak, cumulative, timing, and duration-based affectedness indicators. The 12 h period is therefore treated as a fixed, deliberately extended analysis window for

short-duration flood processes rather than evidence of complete hydraulic convergence; the methodological limitations consider potential effects of the finite observation window on cumulative affectedness.

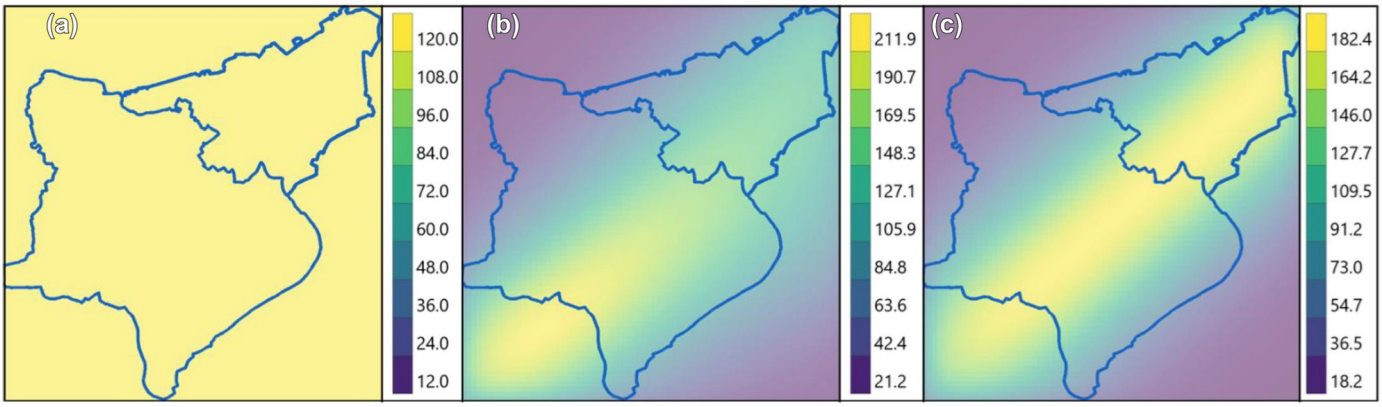


Figure 4. Accumulated precipitation (in mm) for the three synthetic hydrometeorological event types: (**a**) stationary cloudburst (SCB), (**b**) multi-cell event (MCE), and (**c**) moving convective storm (MCS).

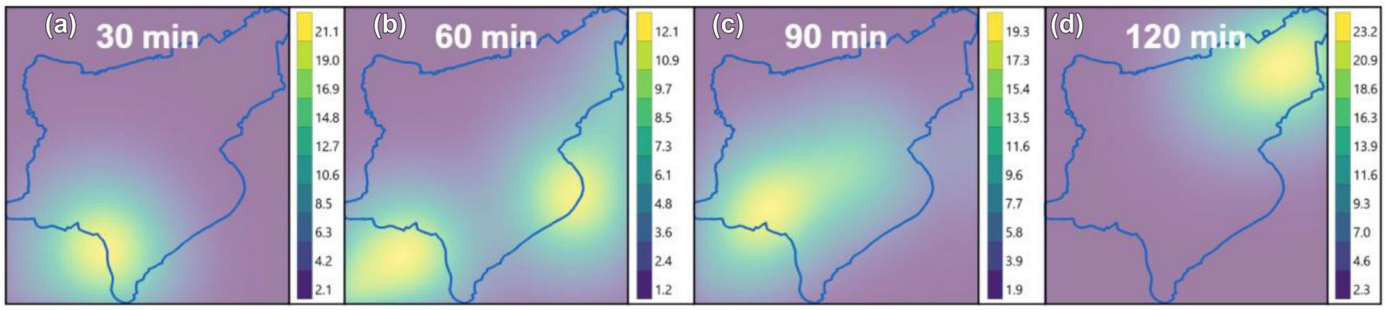


Figure 5. Spatial precipitation distribution (in mm) of the multi-cell event (MCE) at (**a**) 30 min, (**b**) 60 min, (**c**) 90 min, and (**d**) 120 min after the start of the synthetic rainfall event.

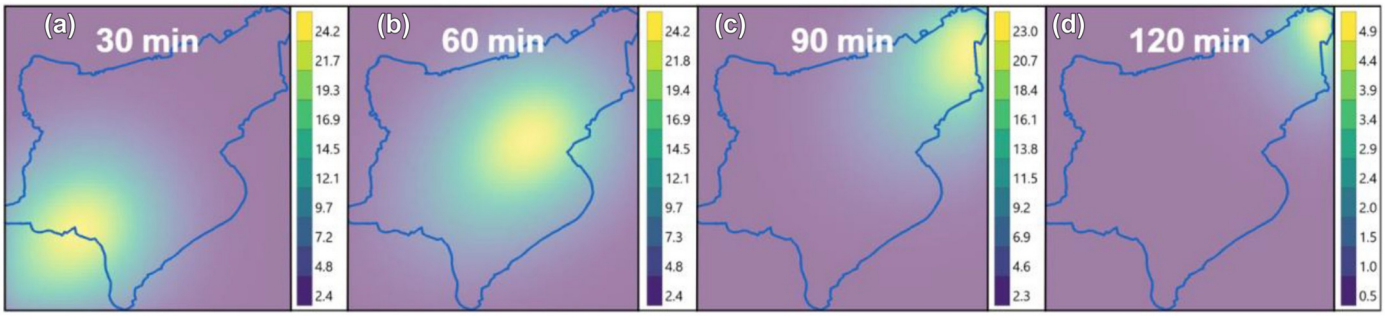


Figure 6. Spatial precipitation distribution (in mm) of the moving convective storm (MCS) at (**a**) 30 min, (**b**) 60 min, (**c**) 90 min, and (**d**) 120 min after the start of the synthetic rainfall event.

2.4. Scenario Design and Stressors

The scenario design translates the generic FLOOD-ST framework into a catchment-specific set of controlled hydrodynamic simulations (Table 2). Five conditions were varied: hydrometeorological event type, antecedent moisture condition (AMC), clogging state, pump operation, and detention-basin outflow. Across the defined levels of these five conditions, all combinations would comprise 108 scenarios (Supplementary Table S1). The present study uses a deliberately selected subset of 23 simulations, consisting of five baseline scenarios (BL01–BL05) and 18 compound stress-test scenarios (CS01–CS18). This subset is not meant to represent the full parameter space statistically, but to provide

interpretable comparisons of relevant hydrometeorological and infrastructure states within a computationally feasible case-study design.

Table 2. Overview of the 23-scenario Roßbach matrix, comprising five baseline scenarios (BL) and 18 compound stress-test scenarios (CS).

Stressor	AMC I	AMC II	AMC III	SCB	MCE	MCS	Clogging	No Clogging	Pump Outage	Pump Default	Pump Increased	Basin Default	Basin Throttled
BL_01	x			x				x		x		x	
BL_02		x		x				x		x		x	
BL_03			x	x				x		x		x	
BL_04			x		x			x		x		x	
BL_05			x			x		x		x		x	
CS_01			x	x			x			x		x	
CS_02			x	x			x				x	x	
CS_03			x	x			x			x			x
CS_04			x	x			x				x		x
CS_05			x	x				x			x		x
CS_06			x		x		x			x		x	
CS_07			x		x		x				x	x	
CS_08			x		x		x			x			x
CS_09			x		x		x				x		x
CS_10			x		x			x			x		x
CS_11			x			x	x			x		x	
CS_12			x			x	x				x	x	
CS_13			x			x	x			x		x	
CS_14			x			x	x				x		x
CS_15			x			x		x			x		x
CS_16			x	x				x	x			x	
CS_17			x		x			x	x			x	
CS_18			x			x		x	x			x	

The "x" symbols indicate which components or characteristics are included in each simulation.

Scenario comparisons follow a simple rule: compare two simulations directly only when one condition changes while the remaining scenario settings are identical. Applying the direct-comparison criterion to the simulated subset yields 45 comparisons in which exactly one investigated condition differs while all remaining scenario settings are identical (Supplementary Table S2). These comparisons form the basis for evaluating changes associated with rainfall-event type, AMC, clogging, pump operation, and detention-basin outflow.

The hydrometeorological dimension comprises three synthetic rainfall-event types defined within FLOOD-ST: a stationary cloudburst (SCB), a multi-cell event (MCE), and a moving convective storm (MCS) [11]. All three events were adjusted to a common catchment-mean precipitation depth of approximately 120 mm, while retaining distinct spatial and temporal rainfall patterns. The event fields were generated at a spatial resolution of 100 m $\times$ 100 m and a temporal resolution of 5 min. This controlled design allows examination of differences in the simulated response under rainfall events with comparable catchment-scale volume but different spatial and temporal structures.

The synthetic events represent deliberately severe stress-test conditions rather than reconstructions of historical storms. Their hydrometeorological plausibility was additionally assessed using the Rainstorm Severity Index (RSI) [23]. The SCB corresponds to RSI 10, while the local maxima of the MCE and MCS reach RSI 12, placing all three within the extreme-heavy-rainfall range. These classifications are used only as a plausibility reference for the synthetic rainfall intensities and should not be interpreted as assigning a unique return period to the complete spatial rainfall scenarios.

The 100 m rainfall grid represents a scenario-generation scale rather than an observational rainfall resolution. It was kept identical across all investigated event types so that grid resolution does not vary between direct scenario comparisons. No independent rainfall-grid resolution analysis was conducted. Therefore, potential effects of this representation on absolute hydraulic response are considered a model limitation.

Antecedent moisture condition represents the initial catchment state through the AMC classes of the SCS-CN formulation. AMC I, II, and III represent dry, average, and wet antecedent conditions, respectively. The baseline simulations include all three AMC classes, while most compound stress-test scenarios use AMC III to investigate infrastructure-state changes under wet antecedent conditions.

Clogging was represented as a static reduction in conveyance at selected culverts rather than by explicitly modelling debris transport or time-dependent blockage formation. Prior to the final scenario definition, several alternative culvert configurations (CA1–CA3c) were hydraulically screened by varying opening geometry and loss-related parameters. Configuration CA3a produced the strongest sustained conveyance restriction in these preliminary simulations and was therefore used to inform the severe synthetic clogging configuration (Supplementary Table S4).

In the CA3a pretest at structure 2605, the represented culvert height was reduced to 5% of its reference value while retaining the original width. This configuration reduced peak discharge by approximately 65% and integrated discharge volume by approximately 63%. The final stress-test clogging configuration was subsequently applied to structures 2651 and 100015 using strongly increased hydraulic-loss parameters. At structure 2651, the represented height was reduced from 3.16 m to 0.158 m, corresponding to 5% of its reference height. At structure 100015, the represented height was reduced from 2.0 m to 0.447 m together with the modified loss parameters. The resulting configuration is interpreted as an intentionally severe synthetic stress-test condition rather than an empirically representative blockage fraction or a physically resolved debris blockage.

Pump operation was represented through prescribed discharge states. Both pumping stations operate with a reference capacity of 1 $m^3\ s^{-1}$ in the model. Selected stress-test scenarios represent either increased pumping capacity of 5 $m^3\ s^{-1}$ or pump failure with a capacity of 0 $m^3\ s^{-1}$. The operational information used to define the pumping stations was provided by Emschergenossenschaft/Lippeverband (EGLV).

The Martener Straße detention basin was represented through a prescribed stage–discharge relationship. In addition to the reference relationship, a throttled state was defined by multiplying each discharge ordinate of the reference stage–discharge relationship by 0.7, corresponding to a 30% reduction in prescribed discharge at each stage. This modification represents a simplified operational stress-test condition and is not intended to reproduce adaptive real-time basin control.

2.5. Affectedness Indicators and Post-Processing

Hydraulic affectedness was derived from 145 complete paired water-depth and flow-velocity records available for each simulation run, comprising 144 regularly stored records

and one complete terminal partial-interval record. At each time step t_i [s], specific discharge q [m^2/s] was calculated as

$$q(t_i) = h \times v(t_i),$$

and the Froude number [-] as

$$Fr = \frac{v}{\sqrt{(g\,h)}}$$

where h is water depth [m], v is flow velocity [m/s], and g is gravitational acceleration [m/s^2].

The affectedness assessment used an application-specific conservative threshold set assembled from published flood-hazard criteria for vulnerable persons and flash-flood screening [24,25]. A raster cell was classified as hazard-positive if at least one of the following thresholds was exceeded:

$$I_{cell}(t) = \mathbb{1}[h \geq 0.30\ m \vee v \geq 0.20\ m\ s^{-1} \vee q \geq 0.20\ m^2; s^{-1} \vee Fr \geq 0.12]$$

Here, $\mathbb{1}[\cdot]$ denotes the indicator function, which equals 1 when the condition is fulfilled and 0 otherwise.

The water-depth and specific-discharge thresholds were adopted from published flash-flood screening criteria, while the velocity threshold represents the lower bound associated with increased risk for children and vulnerable persons in hazard classification. The Froude threshold was derived from the selected water-depth and velocity thresholds and was therefore used as an application-specific supplementary screening criterion.

To reduce false detections caused by numerical wet-dry transitions and locally implausible values of specific discharge or Froude number, additional quality constraints were applied.

Velocity and specific-discharge exceedances were considered only where

$$h \geq 0.10\ m.$$

Froude-number exceedances were considered only where

$$h \geq 0.10\ m \text{ and } Fr \leq 5.$$

These thresholds were applied consistently across all scenarios before combining the four indicators.

The Froude number was used as a supplementary hydraulic screening indicator rather than as a stand-alone measure of damage potential or human stability. Because Froude values can become highly sensitive under shallow-flow conditions, exceedances were retained only above the specified minimum water depth and below the upper plausibility limit. Froude-based results were therefore interpreted only in combination with water depth, velocity, and specific discharge.

The spatial affectedness assessment differed between the investigated asset classes. Streets were evaluated on a per-raster-cell basis, and the affected quantity was expressed as the area of affected street in square metres. Buildings and critical infrastructure assets were evaluated as polygon objects using a buffer-ring approach around their boundaries. The critical-infrastructure dataset comprised 58 functionally relevant objects within the study area, including police and fire service facilities, administrative buildings, energy supply infrastructure, schools, and care facilities. A 1 m buffer ring was applied initially and extended to 2 m if fewer than five valid raster cells were available. An object was classified as affected if at least 5% of the valid cells within the buffer ring were hazard-positive.

Because the buffer width and affected-cell threshold represent application-specific classification choices, their influence was examined in a supplementary robustness analysis. The BL05–CS11 comparison was repeated using both the adaptive 1–2 m buffer rule and a fixed 2 m exterior buffer, each combined with affected-cell thresholds of 2.5%, 5%, and 10% (Supplementary Table S6). Across these six configurations, the net difference in cumulatively affected buildings remained positive (+12 to +29 buildings), while gross building turnover ranged from 50 to 68 objects. The absolute affectedness counts therefore depend on the selected classification settings, whereas the central finding of spatial redistribution between the two scenarios remains robust.

Specific discharge and the Froude number are established hydraulic quantities. The asset-class-level affectedness metrics defined below were operationalised for the present case study to translate the FLOOD-ST indicator concept into consistently comparable measures of simultaneous, cumulative, and time-dependent affectedness [11]. They are used as descriptive scenario-comparison metrics and are not intended as standardised damage, risk, or probability indicators.

For each asset class k, the time-resolved affectedness fraction was calculated as

$$I_k(t) = \frac{N_{imp,k}(t)}{N_{ref,k}},$$

where $N_{imp,k}(t)$ denotes the affected reference quantity at time t, and $N_{ref,k}$ denotes the total reference quantity of the respective asset class. For streets, these quantities are expressed as area; for buildings and critical infrastructure, they are expressed as object counts. Absolute values were retained alongside relative fractions throughout the analysis.

The maximum simultaneous affectedness was defined as

$$I_{max,k} = t_{max} \times I_k(t),$$

and $t(I_{max,k})$ denotes the first time at which this maximum was reached.

The cumulative affectedness represents the fraction of the reference quantity affected at least once during the complete simulation period:

$$I_{total,k} = \frac{N_{unique,k}}{N_{ref,k}}$$

where $N_{unique,k}$ is the total street area or number of objects affected at least once during the simulation.

The corresponding timing metric was defined as

$$t(I_{total,k}) = \min \{t: N_{cum,k(t)} = N_{unique,k}\}$$

where $N_{cum,k(t)}$ is the cumulative-affected quantity up to time t. This metric therefore describes the earliest time at which the final simulation-wide cumulative affectedness was reached. The exact definitions of normalisation and timing were specified for this application. They provide a consistent basis for comparing the investigated simulations but have not been independently validated as generally applicable affectedness indices.

Affectedness duration was derived from the time for which an object or street cell was classified as affected within each scenario. For the cross-scenario analysis, recurrence was defined as the number of scenarios in which the respective object or street cell was affected and therefore ranges from 0 to 23. In addition, conditional mean duration describes the mean duration across those scenarios in which affectedness occurred. Recurrence

and duration thus provide complementary information on how consistently and how persistently individual locations respond across the investigated scenario ensemble.

2.6. Matched-Pair Analysis of the Documented Clogging-State Contrast

To examine whether a documented infrastructure-state contrast is associated with temporal and spatial redistribution despite comparatively small aggregate changes, scenarios BL05 and CS11 were selected for a matched-pair analysis. According to the documented scenario matrix, both scenarios represent AMC III, the moving convective storm (MCS) trigger event, default pump operation, and the reference detention-basin outflow condition. They differ by the predefined clogging-state configuration and therefore constitute the documented matched clogging-state contrast analysed in this study.

The comparison was conducted for streets, buildings, and critical infrastructure using both cumulative affectedness and time-resolved affectedness indicators. For the spatial comparison, affected locations and assets were assigned to one of three classes: affected in both scenarios, affected only in BL05, or affected only in CS11. For the independent inundation-depth comparison, the same classification was applied to maximum-water-depth masks derived from all 145 complete hydraulic records using a threshold of $h \geq 0.30$ m.

For each asset class k, the net change in cumulative affectedness was calculated as the difference between CS11 and BL05. In addition, gross spatial turnover was defined as the sum of the reference quantity affected only in BL05 and the reference quantity affected only in CS11. This distinction separates the net change between scenarios from spatial redistribution caused by offsetting local increases and decreases.

Temporal differences were evaluated using the time-resolved affectedness fractions defined in Section 2.5 across all 145 complete records. The pairwise difference was calculated as

$$\Delta I_{k(t)} = I_{k,CS11(t)} - I_{k,BL05(t)},$$

where positive values indicate greater simultaneous affectedness under CS11, representing the predefined clogging state, and negative values indicate greater simultaneous affectedness under BL05, representing the otherwise identical non-clogged baseline state.

2.7. AI Use Statement

Generative AI was used as a supporting tool during the development and refinement of the post-processing workflow. ChatGPT (OpenAI; GPT-5.5 and GPT-5.6) was used to assist with the development, debugging, and refinement of Python-based (Python v3.12.4/v3.12.2) scripts for processing, analysing, and visualising hydrodynamic simulation results. ChatGPT (GPT-5.5 and GPT-5.6) was additionally used for language editing, restructuring, and refinement of manuscript text. All generated code, analytical outputs, and textual suggestions were critically reviewed, tested, and, where necessary, modified by the authors. Generative AI was not used to generate simulation data or as an autonomous source of scientific evidence. All methodological decisions, scientific interpretations, and conclusions were critically assessed, verified, and approved by the authors.

3. Results

3.1. Inundation Extents and Contextual Reference-Map Comparison

The maximum inundation extent at the h = 0.30 m threshold varied substantially across the investigated scenario ensemble (Figure 7). The five baseline simulations ranged from approximately 0.929 km^2 in BL01 to 1.598 km^2 in BL05. For the SCB event, increasing antecedent moisture from AMC I to AMC III increased the inundated area from approximately 0.929 km^2 (BL01) to 1.321 km^2 (BL03). At AMC III, changing the hydrometeorological event

type further increased the inundation extent from SCB (BL03: 1.321 km^2) to MCE (BL04: 1.502 km^2) and MCS (BL05: 1.598 km^2).

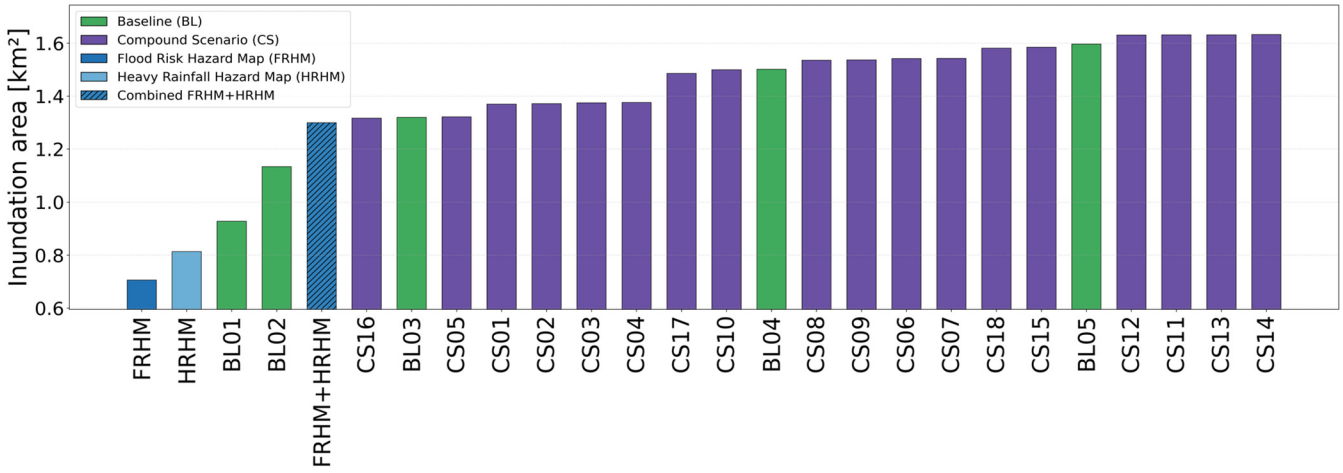


Figure 7. Maximum inundation extent (h $\geq$ 0.30 m) across the 23 FLOOD-ST scenarios in comparison with the official Flood Hazard Map (FRHM) and Heavy-Rainfall Hazard Information Map (HRHM). The reference products are included for contextual comparison only and are not interpreted as calibration or validation data. Spatial-overlap metrics are provided in Supplementary Table S9.

The compound scenarios ranged from approximately 1.317 km^2 (CS16) to 1.633 km^2 (CS14). Within the AMC III scenarios, the inundation extents were primarily grouped according to hydrometeorological event type. SCB-based scenarios occupied the lower range, MCE-based scenarios an intermediate range, and MCS-based scenarios the upper range. Differences among infrastructure-state configurations within each event type were comparatively smaller than the differences among the three event types.

For contextual comparison, the official North Rhine-Westphalian Flood Hazard Map (FRHM) (https://www.flussgebiete.nrw.de/hochwassergefahrenkarten-inhalte-und-symbole) (accessed on 8 September 2026) covers approximately 0.706 km^2 within the analysed area, whereas the Heavy-Rainfall Hazard Information Map (HRHM) (https://www.flussgebiete.nrw.de/hinweiskarte-starkregengefahren) (accessed on 8 September 2026) covers approximately 0.813 km^2. In addition to considering both reference products separately, their spatial footprints were combined into a geometric union. This combined reference footprint comprises every location classified as inundated in either the FRHM or the HRHM and covers approximately 1.300 km^2. The published maps and the FLOOD-ST simulations differ in event definition, probability basis, forcing, boundary conditions, and modelling assumptions. They are therefore used only as contextual spatial references and not as calibration or validation data.

Spatial agreement was additionally assessed using the intersection-over-union (IoU). For each comparison, IoU is defined as the area classified as inundated in both the FLOOD-ST scenario and the respective reference map divided by the total area classified as inundated in either dataset. An IoU of 1 would indicate identical spatial inundation footprints, whereas a value of 0 would indicate no spatial overlap. Across the 23 FLOOD-ST scenarios, IoU ranges from 0.260 to 0.355 for the FRHM and from 0.369 to 0.407 for the HRHM. When the combined FRHM–HRHM reference footprint is used, IoU ranges from 0.461 to 0.506. The comparison therefore indicates partial to substantial spatial correspondence with the published hazard information while also showing scenario-specific inundation outside the reference footprints. These overlap values are interpreted exclusively as contextual spatial benchmarks rather than as measures of model performance or validation (Supplementary Table S9).

3.2. Asset-Based Affectedness

Affectedness was initially assessed at the study-area-wide level using maximum simultaneous affectedness (I_{max}) and final cumulative affectedness (I_{total}). BL05 achieved the highest I_{max} for both streets and buildings, while CS14 recorded the highest I_{max} for critical infrastructure (Figure 8). Except for BL01, all infrastructure assets reached I_{max} within roughly two to three hours into the simulation.

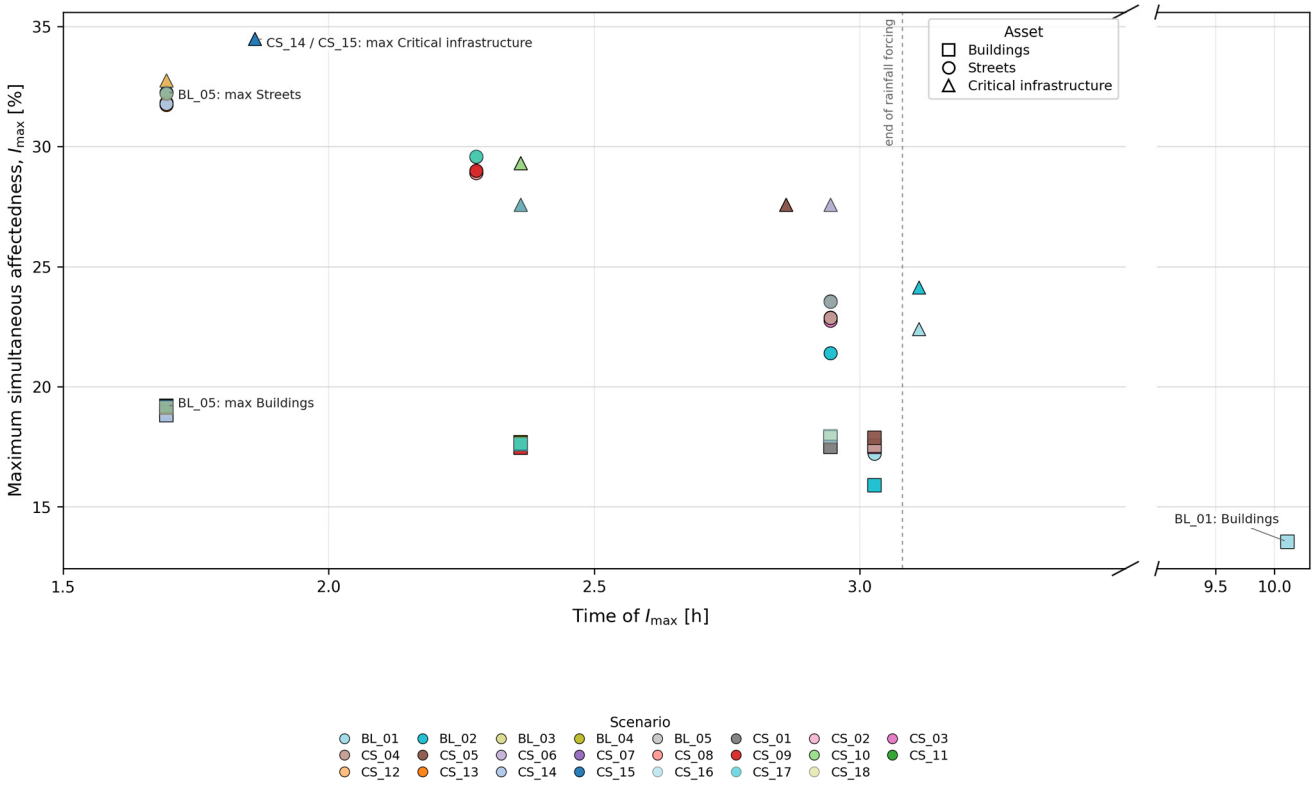


Figure 8. I_{max} per scenario.

For I_{total}, critical infrastructure reached its peak affectedness between two and three hours, considerably earlier than streets, which reached I_{total} at twelve hours, and buildings, which clustered between nine and twelve hours (Figure 9). Across all scenarios, BL05 reached the highest I_{total} for critical infrastructure, while CS13 and CS14 showed the highest values for buildings and streets, respectively.

Although BL05, CS11, and CS12 show similar relative affectedness levels, the corresponding absolute quantities reveal scenario-specific differences. For buildings, BL05 attained the highest I_{max} (1589 objects), followed by CS12 (1571 objects) and CS11 (1560 objects). For critical infrastructure, all three scenarios yielded the same I_{max} and I_{total} values: 19 and 26, respectively. However, in terms of cumulative affectedness, BL05 affected 23 fewer buildings than CS12 and 25 fewer than CS11. This highlights the importance of evaluating absolute values alongside percentages at the study-area scale.

CS13 and CS14 demonstrated temporal affectedness dynamics similar to BL05, CS11, and CS12. In both cases, affectedness rose rapidly around the one-hour mark and reached I_{max} within the first two hours of the simulation. For streets and buildings, I_{max} occurred at approximately 1.69 h in both scenarios, while affectedness on critical infrastructure peaked slightly later in CS14 than in CS13.

Indicator-specific building affectedness varied considerably. Velocity-based exceedances remained comparatively stable at about 11.0% to 11.2% of buildings and were mainly short-lived. Water-depth exceedances affected about 20.0% to 20.3% of buildings

and were characterised by comparatively long exceedance durations, indicating persistent depth affectedness once the threshold was crossed. Froude-number exceedances affected about 15.4% to 15.7% of buildings and, similar to velocity, were predominantly short-lived. Specific-discharge exceedances affected only a small subset of buildings, between 1.4% and 1.8%, and were again mostly short-duration.

When affectedness was defined as the logical or of all four thresholds, the proportion of affected buildings increased to about 28.4–28.8% and remained nearly constant across the six scenarios compared. The choice of indicator significantly influenced the apparent affectedness level: single-indicator evaluations ranged from about 1.4% to 20.3%, whereas the combined multi-indicator definition resulted in about 28.4% to 28.8%. This difference reflects the broader classified affectedness resulting from combining multiple hydraulic screening criteria and should not be interpreted as evidence that any individual indicator systematically underestimates actual flood impact.

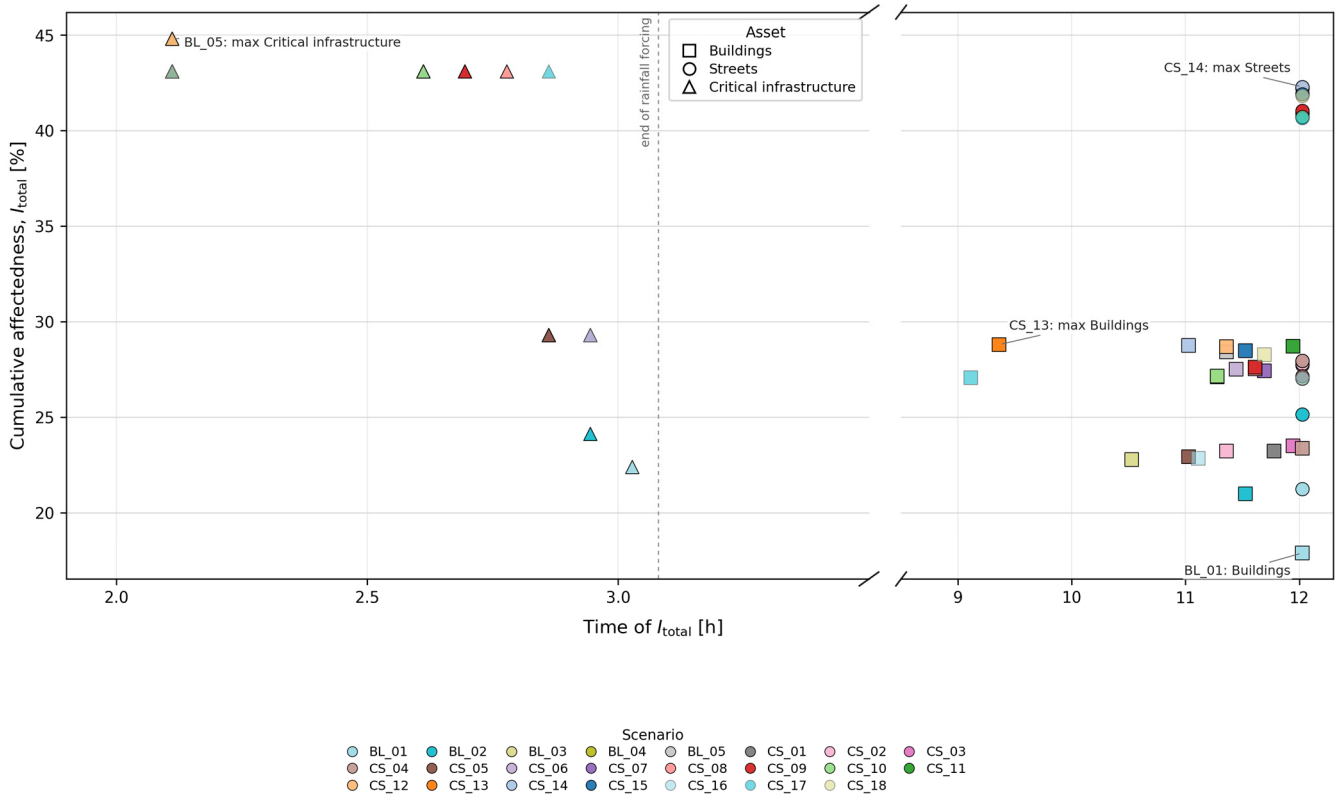


Figure 9. I_{total} per scenario.

3.3. Antecedent Moisture Conditions

Under SCB, increasing AMC from 1 to 3 led to consistently higher affectedness across all asset classes in terms of both cumulative affectedness and peak simultaneous affectedness. For buildings, the increase from AMC1 to AMC3 corresponded to +4.93% (+408 buildings) in cumulative affectedness and +4.39% (+363 buildings) in maximum simultaneous affectedness. The associated peak time shifted markedly from 10.11 h to 2.94 h, while cumulative timing shifted from 12.03 h to 10.53 h.

For streets, the change from AMC1 to AMC3 increased cumulative affectedness by +5.78% (+29,169 m^2) and maximum simultaneous affectedness by +6.34% (+31,981 m^2), with only minor timing shifts. For critical infrastructure, AMC1 to AMC3 resulted in +6.90% (+4 objects) in cumulative affectedness and +5.17% (+3 objects) in maximum simultaneous affectedness, with earlier peak and cumulative timing (Figures 8 and 9).

3.4. Trigger-Event Contrast

At constant AMC3, changing the trigger event from SCB to MCE and MCS resulted in marked differences in affectedness magnitude and timing, particularly for streets and critical infrastructure. Relative to SCB, MCE increased cumulative street affectedness by +13.66% (+68,876 m^2) and cumulative critical-infrastructure affectedness by +13.79% (+8 objects), while maximum simultaneous affectedness increased by +6.02% (+30,335 m^2) for streets and +1.72% (+1 object) for critical infrastructure. In both asset classes, the peak occurred earlier than under SCB.

Relative to SCB, MCS showed even larger increases, with cumulative affectedness rising by +14.90% (+75,217 m^2) for streets and +15.52% (+9 objects) for critical infrastructure, while maximum simultaneous affectedness rose by +8.69% (+43,821 m^2) for streets and +5.17% (+3 objects) for critical infrastructure. Peak timing shifted from 2.94 h to 1.69 h for streets and from 2.86 h to 1.69 h for critical infrastructure.

For buildings, the trigger change primarily affected cumulative affectedness. Compared with SCB, cumulative building affectedness increased by +4.34% under MCE (+359 objects) and by +5.63% under MCS (+466 objects), again with earlier timing (Figures 8 and 9).

3.5. Clogging Contrasts

Across the paired comparisons, clogging produced a consistent trade-off between peak and cumulative affectedness. Under SCB, maximum simultaneous affectedness decreased while cumulative affectedness increased. For buildings, maximum simultaneous affectedness decreased by 0.365% (−31 objects), whereas cumulative affectedness increased by 0.455% (+38 objects). For streets, maximum simultaneous affectedness decreased by −0.725% (−3660 m^2), while cumulative affectedness increased by +0.745% (+3761 m^2). Timing of the peak remained essentially unchanged, but cumulative building affectedness occurred markedly later.

Under MCE, clogging again reduced maximum simultaneous affectedness and slightly increased cumulative affectedness, though to a smaller extent than under SCB. Buildings showed −0.205% (−17 objects) in maximum simultaneous affectedness and +0.435% (+36 objects) in cumulative affectedness, while streets showed −0.625% (−3155 m^2) in maximum simultaneous affectedness and +0.275% (+1389 m^2) in cumulative affectedness. Critical infrastructure showed a small reduction in peak affectedness and nearly unchanged cumulative affectedness.

Under MCS, the documented BL05–CS11 contrast showed the same trade-off, with lower maximum simultaneous affectedness but slightly higher cumulative affectedness. Buildings showed a decrease of 0.351 percentage points (−29 objects) in maximum simultaneous affectedness and an increase of 0.302 percentage points (+25 objects) in cumulative affectedness, while streets showed a decrease of 0.475 percentage points (−2396 m^2) in maximum simultaneous affectedness and an increase of 0.313 percentage points (+1582 m^2) in cumulative affectedness. Critical infrastructure remained unchanged in both peak and cumulative affectedness.

Because these aggregate differences remained small in magnitude (Figures 10 and 11), a direct matched-pair analysis of the documented clogging-state contrast between BL05 and CS11 was conducted to examine whether the observed response reflected spatial turnover and temporally shifting affectedness rather than uniform amplification.

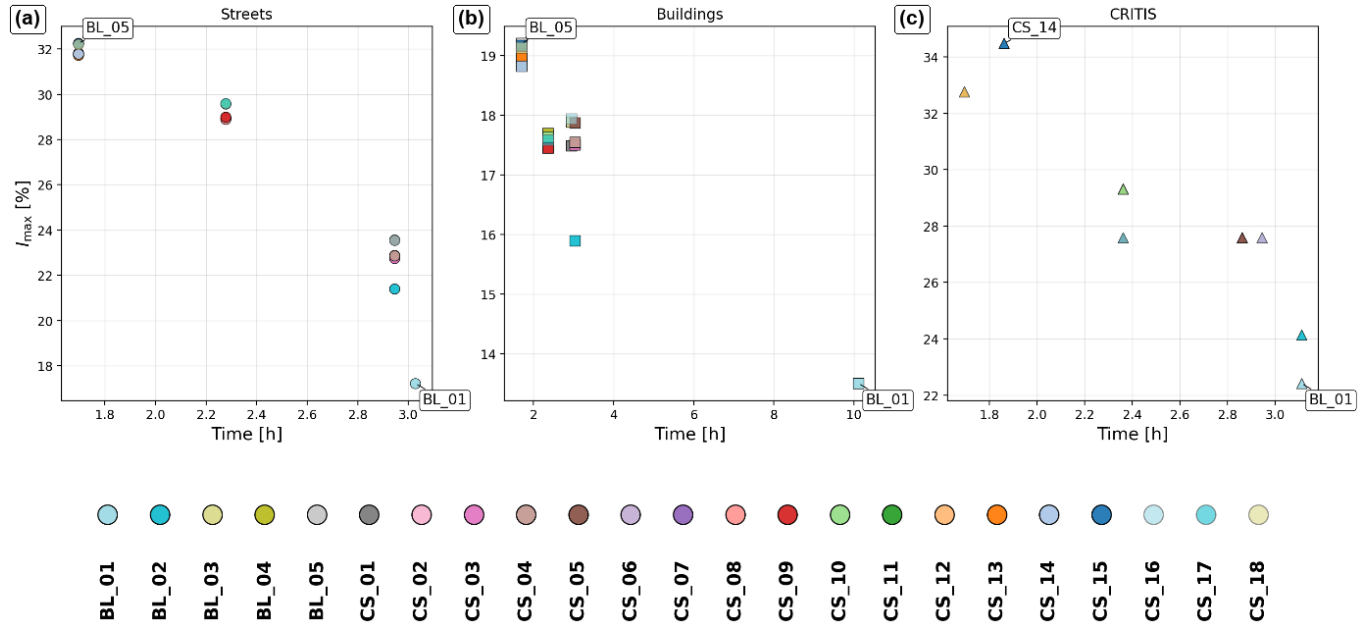


Figure 10. Maximum simultaneous affectedness I_{max} and corresponding timing across the investigated scenarios, shown separately for (**a**) streets, (**b**) buildings, and (**c**) critical infrastructure.

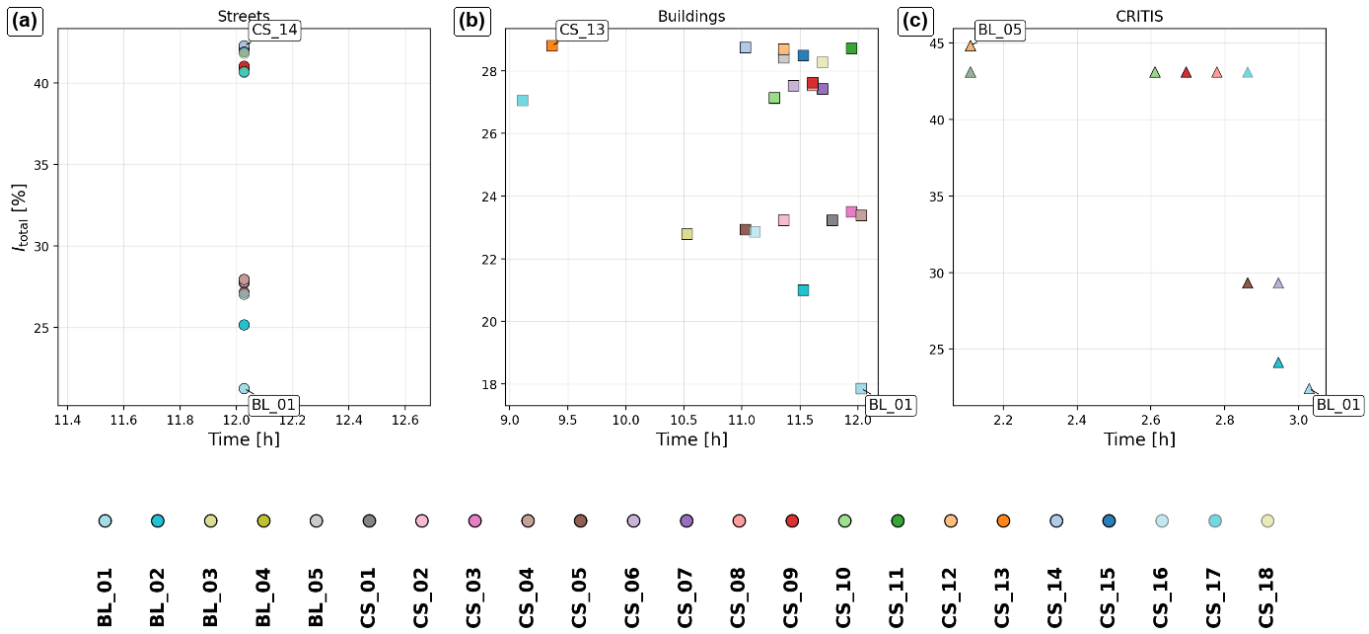


Figure 11. Cumulative affectedness I_{total} and corresponding timing across the investigated scenarios, shown separately for (**a**) streets, (**b**) buildings, and (**c**) critical infrastructure.

3.6. Detention-Basin Throttling Contrasts

Reducing detention basin outflow capacity from ×1 to ×0.7 produced only minor changes in maximum simultaneous affectedness, but slightly increased cumulative affectedness, particularly for buildings and streets. Under SCB, buildings showed +0.025% (+3 buildings) in maximum simultaneous affectedness and +0.215% (+18 buildings) in cumulative affectedness, while streets showed a slight peak decrease and a small cumulative increase. Critical infrastructure remained effectively unchanged.

Under MCE, detention basin throttling again showed very small differences across all asset classes. Buildings showed a minor peak reduction but a slight cumulative increase, with negligible timing shifts. Streets showed similarly small cumulative increases and near-zero peak change, while critical infrastructure again displayed no relevant change.

Under MCS, effects on buildings and streets remained small in magnitude, but the response became more heterogeneous, especially for critical infrastructure. Buildings and streets showed slight peak decreases combined with small cumulative increases. Notably, cumulative building affectedness occurred substantially earlier under discharge throttling, indicating primarily a shift in timing rather than a strong magnitude effect. For critical infrastructure, throttling increased the peak but reduced cumulative affectedness.

3.7. Pump-State Contrasts

Pump-state effects were evaluated using nine direct matched comparisons in which pump operation was varied while the remaining scenario conditions were held constant (Supplementary Table S5). Across these comparisons, changes in maximum simultaneous and cumulative affectedness were generally small, but their direction depended on the hydrometeorological event type and the accompanying clogging and detention-basin states. Pump modification therefore did not produce a uniform monotonic response across the investigated asset classes.

Across SCB configurations, pump effects were generally small, with the clearest responses occurring for streets. Under detention basin $\times 0.7$ and clogging, maximum simultaneous affectedness increased slightly, while cumulative affectedness showed mixed behaviour. Under detention basin $\times 1$ without clogging, pump outage caused only small changes in magnitude, accompanied by moderate timing shifts. Under detention basin $\times 1$ with clogging, streets again showed small increases in peak and cumulative affectedness, whereas buildings remained nearly unchanged in magnitude.

For MCE, pump changes remained secondary, and turning pumps off did not systematically worsen affectedness. Under detention basin $\times 0.7$ with clogging, streets experienced small increases in peak and cumulative affectedness, while buildings changed only slightly and critical infrastructure showed a small peak increase. Under detention basin $\times 1$ without clogging, buildings even decreased in both peak and cumulative affectedness, although cumulative affectedness occurred substantially earlier. Under detention basin $\times 1$ with clogging, streets again increased slightly, whereas buildings showed a small reduction in cumulative affectedness.

Under MCS, pump effects remained small overall but included pronounced timing differences for buildings, while streets responded more consistently in magnitude. Under detention basin $\times 0.7$ with clogging, buildings reduced both peak and cumulative affectedness but with a strong delay in cumulative timing, whereas streets increased in both peak and cumulative affectedness. Under detention basin $\times 1$ without clogging, both buildings and streets decreased. Under detention basin $\times 1$ with clogging, buildings increased in peak but slightly reduced cumulative affectedness, while streets increased in peak and reduced cumulative affectedness, indicating redistribution rather than uniform amplification (Figures 10 and 11).

The directionally mixed responses indicate that changing pump operation primarily modifies local routing and redistribution within the coupled surface-flow system rather than producing a uniform increase or decrease in affectedness. The nine matched pump-state comparisons and their changes in I_{max}, cumulative affectedness, and timing are summarised in Supplementary Table S5.

3.8. Direct Spatial and Temporal Comparison of the Documented Clogging-State Contrast

BL05 and CS11 were analysed as the documented matched clogging-state contrast to examine the diagnostic response associated with the predefined infrastructure perturbation under otherwise comparable conditions. The two scenarios share AMC III, the MCS trigger, default pump operation, and the reference detention-basin outflow condition, but differ in the documented clogging-state configuration.

The independent maximum-water-depth comparison ($h \geq 0.30$ m) showed an increase in inundated area from 1,597,796 m^2 in BL05 to 1,632,104 m^2 in CS11, corresponding to a net increase of 34,308 m^2. At the same time, the spatial overlay revealed 1,578,398 m^2 affected in both scenarios, 19,398 m^2 affected only in BL05, and 53,706 m^2 affected only in CS11. The resulting gross spatial turnover therefore amounted to 73,104 m^2, demonstrating that the net increase alone represents only part of the spatial change between the two scenarios (Figure 12).

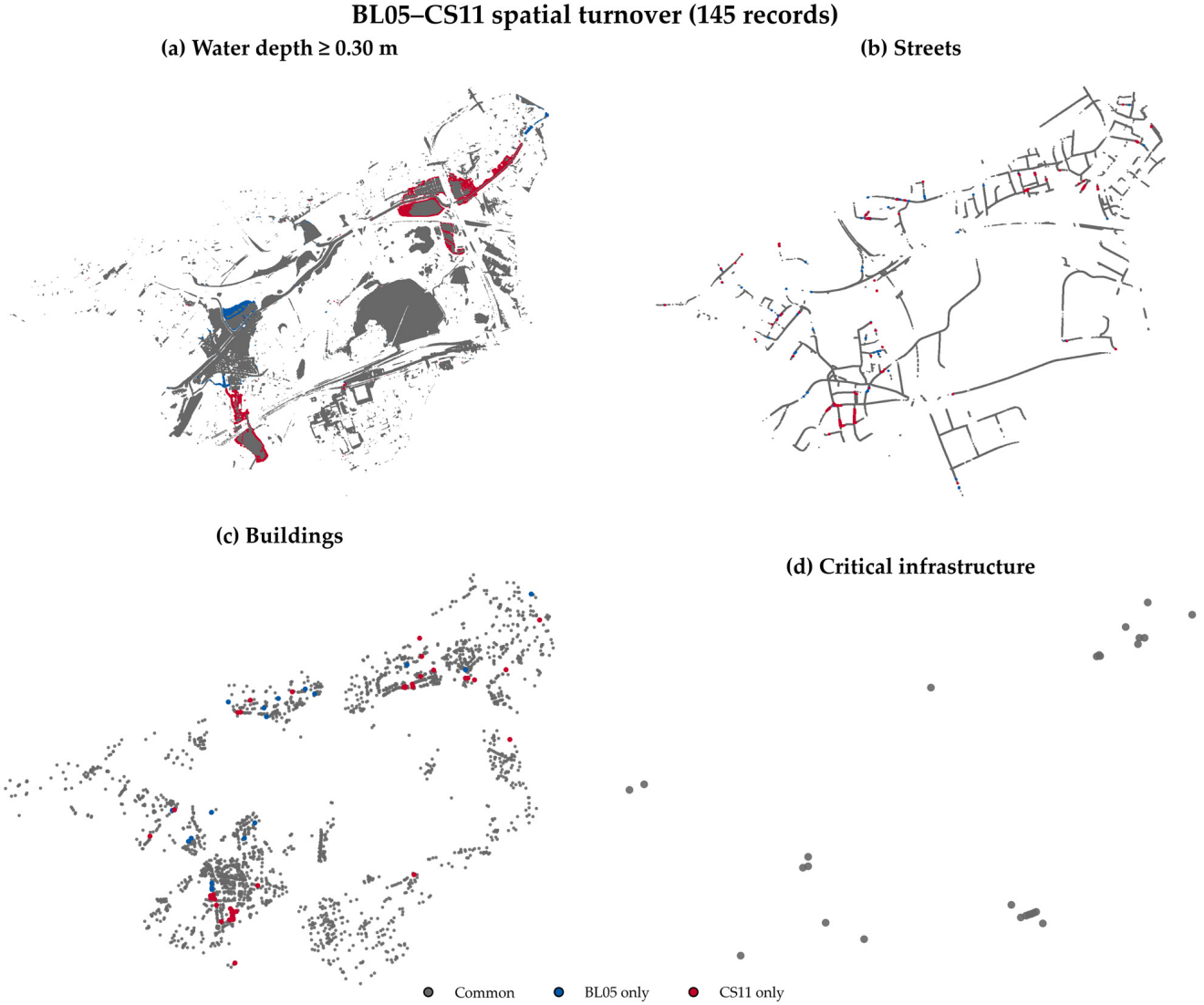


Figure 12. Spatial turnover between BL05 and CS11 for the documented clogging-state contrast. Panel (**a**) shows the maximum-water-depth footprint at $h \geq 0.30$ m, whereas panels (**b**–**d**) show cumulative affectedness for streets, buildings, and critical infrastructure, respectively. Grey indicates locations affected in both scenarios, blue indicates locations affected only in BL05, and red indicates locations affected only in CS11. The comparison demonstrates that relatively small net differences can coincide with spatially distinct local gains and losses in affectedness.

A similar pattern emerged for buildings. Cumulative building affectedness increased from 2351 objects in BL05 to 2376 objects in CS11, corresponding to a net increase of 25 buildings. However, 2333 buildings were affected in both scenarios, while 18 were affected only in BL05 and 43 only in CS11, resulting in a gross turnover of 61 buildings. In contrast, critical infrastructure showed no cumulative turnover, with 26 objects affected in both scenarios and no objects unique to either BL05 or CS11 (Figure 12).

The street-level cumulative comparison also confirmed a small but measurable redistribution. Cumulative affected street area increased from 211,680 m^2 in BL05 to 213,262 m^2 in CS11, corresponding to a net increase of 1582 m^2. Spatially, 210,154 m^2 were affected in both scenarios, 1526 m^2 only in BL05, and 3108 m^2 only in CS11, giving a gross turnover of 4634 m^2 (Figure 12).

The temporal comparison (Figure 13) further showed that the clogging-state contrast did not produce a uniform increase in simultaneous affectedness. At 6100 s (1.69 h), corresponding to the first common peak, CS11 remained 0.475 percentage points below BL05 for streets and 0.351 percentage points below BL05 for buildings, while critical infrastructure showed no relevant difference. Later in the event, the sign of the pairwise difference changed. The largest positive difference for streets reached +3447 m^2 at 14,200 s (3.94 h), and the largest positive difference for buildings reached +53 objects at 19,900 s (5.53 h). Critical infrastructure differed only briefly by one object during the early event phase.

These results show that the documented clogging-state contrast was not primarily expressed through large changes in cumulative totals. Instead, it altered when affectedness occurred, which buildings and infrastructure assets were affected, and where local increases and decreases compensated each other. The matched-pair analysis therefore provides direct case-study evidence that small net changes may conceal diagnostically relevant spatial turnover and temporal redistribution.

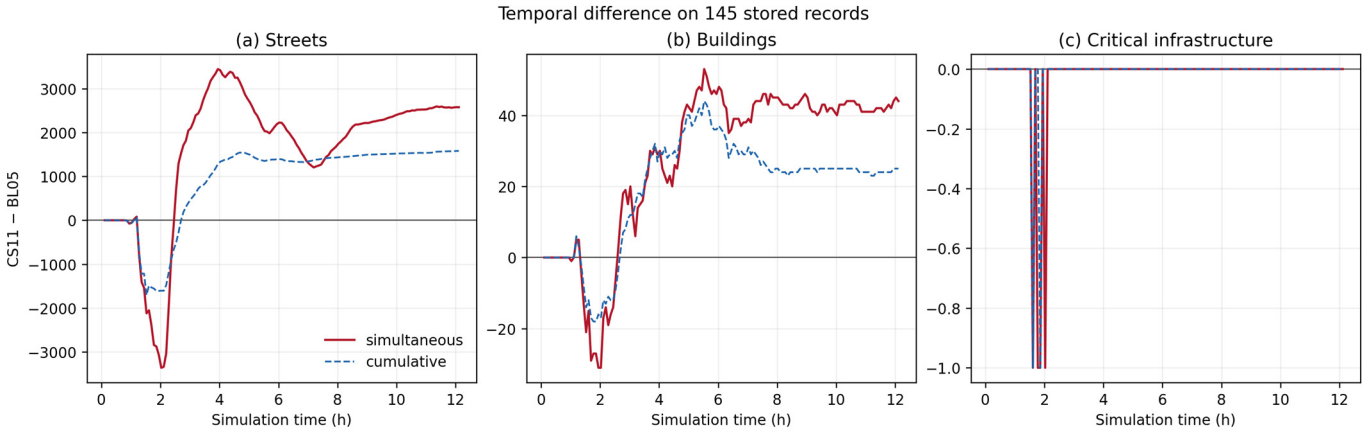


Figure 13. Time-resolved pairwise difference between CS11 and BL05 for the documented clogging-state contrast. Panel (**a**) shows the difference in simultaneously affected street area, panel (**b**) the difference in affected buildings, and panel (**c**) the difference in affected critical infrastructure. Positive values indicate greater simultaneous affectedness under the predefined clogging configuration (CS11), whereas negative values indicate greater simultaneous affectedness under the reference infrastructure state (BL05).

3.9. Hazard Hotspots

Beyond the direct BL05–CS11 matched-pair comparison, broader spatial tendencies across the complete scenario ensemble were further assessed through recurring hotspot patterns.

Hazard hotspots were derived from cross-scenario recurrence and affectedness duration across the 23-scenario ensemble. Recurrence describes the number of scenarios in which a street cell or asset was classified as affected and therefore ranges from 0 to 23, whereas conditional mean duration describes the mean affectedness duration across those scenarios in which affectedness occurred. For streets, an exemplary hotspot area was identified near the Martener Straße detention basin (Figure 14). The recurrence-based representation therefore highlights locations repeatedly affected across different scenario configurations, while the duration-based representation distinguishes locations where affectedness persisted for longer periods.

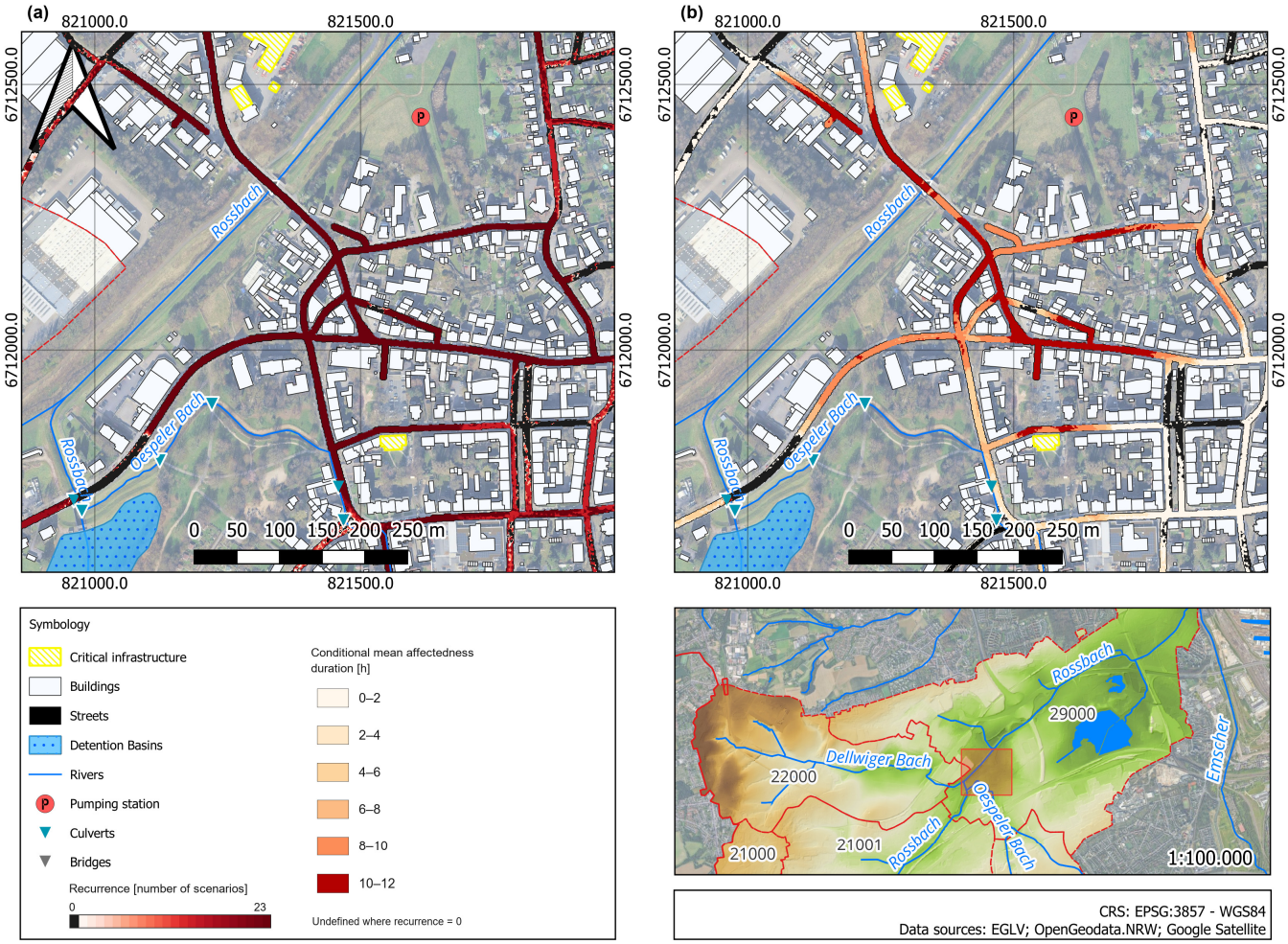


Figure 14. Exemplary street-affectedness hotspot across the 23-scenario ensemble. Panel (**a**) shows the recurrence of affected street cells across the investigated scenarios (0–23 scenarios), whereas panel (**b**) shows the conditional mean affectedness duration across those scenarios in which the respective cell was affected. The inset indicates the location of the displayed area within the Roßbach catchment.

For buildings, hotspot visualisation in an industrial area further to the east shows the same distinction between recurrence and affectedness duration (Figure 15). The duration-based representation highlights buildings or building sections for which affectedness persisted for comparatively longer periods within the scenarios in which they were affected.

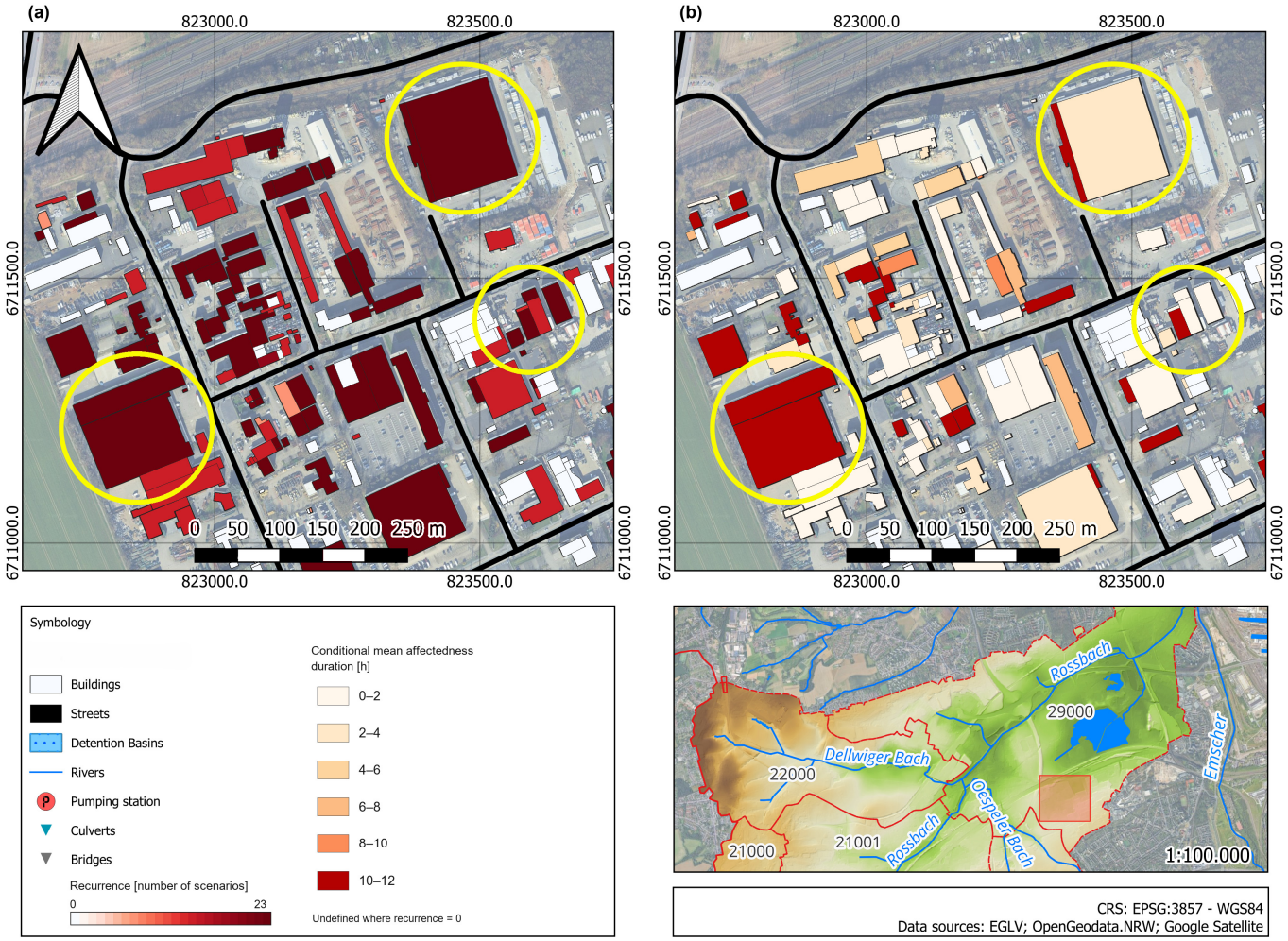


Figure 15. Exemplary building-affectedness hotspot across the 23-scenario ensemble. Panel (**a**) shows the recurrence of affected buildings across the investigated scenarios (0–23 scenarios), whereas panel (**b**) shows the conditional mean affectedness duration across those scenarios in which the respective building was affected. The inset indicates the location of the displayed industrial area within the Roßbach catchment.

3.10. Consistently Unaffected Locations and Assets

Locations and assets with zero recurrence were classified as consistently unaffected within the analysed 23-scenario ensemble and under the applied affectedness thresholds (Figure 16). This applied to 5429 of the 8273 buildings and 30 of the 58 critical-infrastructure objects. For streets, the same criterion was applied at raster-cell level. These locations are not interpreted as inherently safe areas, because their classification is conditional on the investigated scenario ensemble, hydraulic model, observation period, and affectedness thresholds. Instead, they identify locations for which no threshold exceedance was detected in any of the analysed scenarios.

Street cells outside the aggregated depth-based inundation footprint may nevertheless be classified as affected because the multi-indicator assessment additionally considers exceedances of flow velocity, specific discharge, and Froude number. Consequently, consistently unaffected locations are defined using the complete affectedness classification rather than water depth alone.

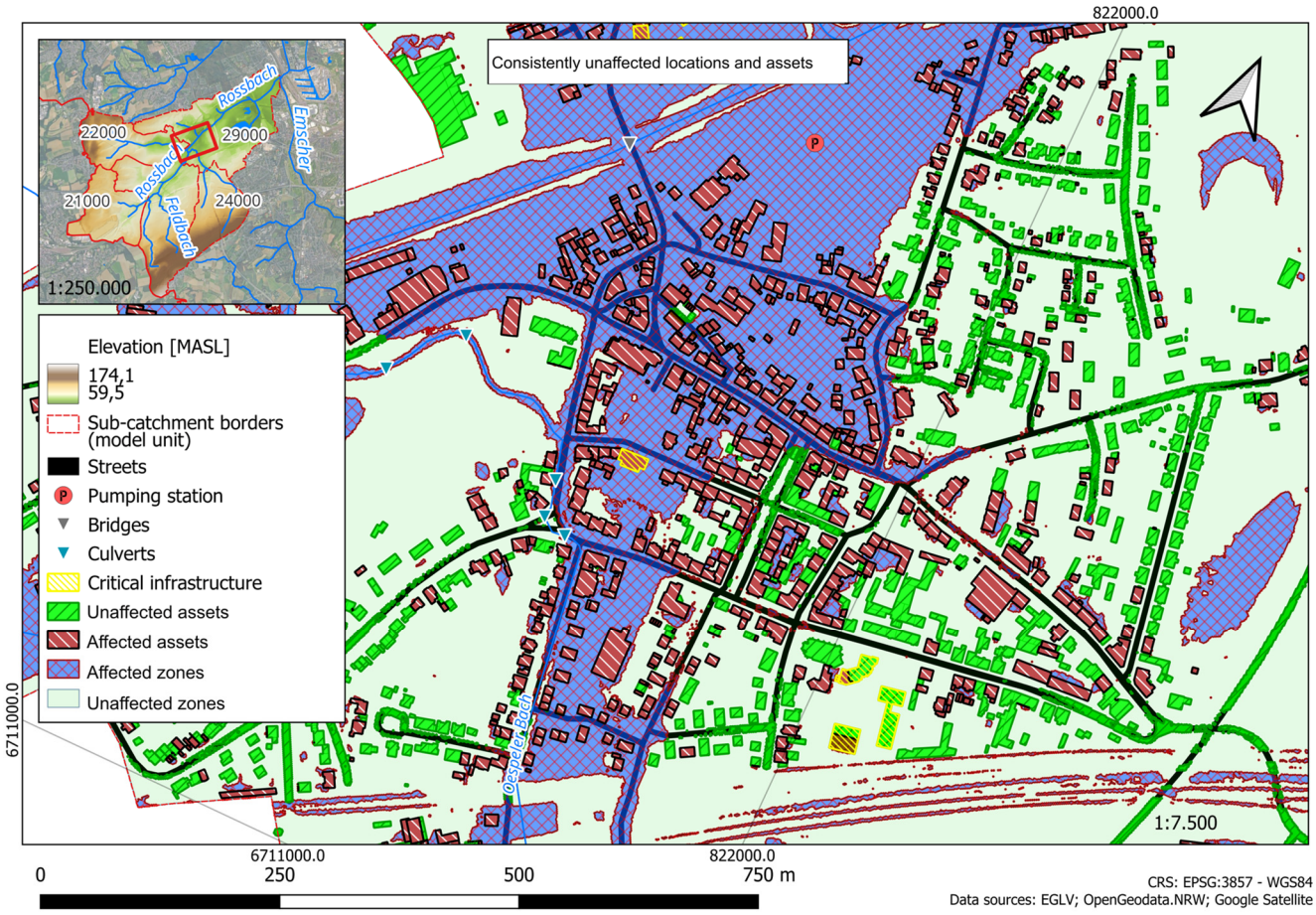


Figure 16. Consistently unaffected locations and assets within the analysed 23-scenario ensemble. The displayed locations were not classified as affected in any investigated scenario under the applied multi-indicator thresholds.

4. Discussion

4.1. Diagnostic Value of Controlled Flood Stress Testing Beyond Hazard-Only Comparison

Conventional flood-hazard and flood-risk maps, together with heavy-rainfall hazard maps, remain essential planning and reference products [12]. Flood-risk maps extend hazard information by incorporating exposed assets and potential adverse consequences. However, these products are generally based on probabilistically derived flood events and do not explicitly diagnose how flood-system response changes when event characteristics, antecedent catchment conditions, or infrastructure performance are deliberately perturbed. FLOOD-ST adds this diagnostic layer by linking hydrometeorological events, transparent scenario definition, admissible comparisons, and consistent area- and object-based affectedness indicators into one traceable workflow. Accordingly, the FRHM and HRHM comparisons in this study are interpreted as contextual spatial benchmarks rather than as independent validation of the stress-test simulations.

The results illustrate this value in a rain-on-grid setting. Trigger event type and antecedent moisture condition produced the clearest differences in inundation extent, magnitude, and timing. Infrastructure-state perturbations such as clogging, pump operation, and detention-basin throttling produced more subtle but diagnostically relevant changes, especially in timing and redistribution. This distinction is important because diagnostically relevant system responses are not necessarily expressed as large changes in inundated area. Some investigated conditions primarily altered affectedness magnitude, whereas others were expressed more strongly through changes in where, when, or for how long

affectedness occurred. A hazard-only reading would capture part of this response, but it would not reveal which investigated condition was associated with the observed shift.

Asset-class-specific affectedness further illustrates this distinction. Streets and critical infrastructure showed particularly pronounced responses to changes in rainfall-event type, whereas building affectedness changed more moderately in relative terms. For example, relative to SCB, MCS increased cumulative affectedness by 14.90% for streets and 15.52% for critical infrastructure, compared with 5.63% for buildings. These differences show that the same hydrometeorological perturbation is translated differently across the transport network, the built environment, and functionally relevant infrastructure.

The matched-pair analysis of a baseline scenario (BL05) and a compound scenario (CS11) provides direct case-study support for this diagnostic value. The clogging-state contrast showed that net changes in cumulative affectedness remained comparatively small, yet these small net differences masked larger local changes in which areas and assets were affected. For the independent depth-based comparison, net inundation increased by 34,308 m^2 while the gross spatial turnover amounted to 73,104 m^2. At building level, a net increase of 25 affected objects coincided with 61 buildings changing their affectedness status between the two scenarios. This distinction between net change and gross turnover is central to the diagnostic interpretation because it reveals spatial response differences that would remain hidden in a total-area-only comparison.

The temporal dimension reinforces the same conclusion. During the early event phase, CS11 produced slightly lower simultaneous affectedness than BL05 for both streets and buildings, whereas later time steps showed the opposite behaviour. The clogging-state contrast therefore did not simply amplify the hydraulic response but redistributed affectedness in time and space. This is precisely the type of system response that FLOOD-ST is intended to diagnose.

Therefore, the scenario design should be understood as a diagnostic design rather than a full-factorial experiment. Its purpose is to support interpretable contrasts between comparable runs and not to simulate all theoretically possible combinations. This is crucial for infrastructure-related stress testing, where effects may be redistributive rather than monotonic. If scenarios are not matched by design, differences in affectedness can easily be mistaken for differences arising from setup, hydrometeorological input, or post-processing. FLOOD-ST reduces this ambiguity by defining in advance which stressors are varied and which model components are kept constant.

The asset-class-specific indicators are central to this contribution. Buildings, streets, and critical infrastructure do not provide interchangeable views of the same flood response. Streets are particularly informative for pathway redistribution and access-related exposure, buildings indicate asset-level affectedness, and critical infrastructure highlights functionally relevant elements even when object counts are small. This differentiation helps to explain why inundation extent alone is not enough to diagnose flood-system response. Timing shifts, duration, first affectedness, and cumulative affectedness provide additional information on how affectedness develops and which parts of the system show the clearest response differences.

Cross-scenario hotspot and recurrence analyses further enhance the diagnostic value of the approach. Recurrence identifies locations that are repeatedly affected across the investigated scenario ensemble, while conditional mean affectedness duration provides complementary information on persistence. Conversely, locations and assets with zero recurrence remained consistently unaffected within the analysed 23-scenario ensemble and under the applied thresholds. These locations should not be interpreted as inherently safe areas, because their classification remains conditional on the investigated scenarios, model configuration, observation period, and affectedness definitions. Both recurring hotspots

and consistently unaffected locations can therefore be used as screening information for more detailed assessment, for example with respect to accessibility, local inspection, or emergency planning.

Overall, FLOOD-ST complements established hazard assessment by shifting the analytical focus from inundated areas to the response of the flood system. This makes it particularly useful in settings where system performance depends on the interactions between event structure, catchment state, and infrastructure behaviour.

4.2. Methodological Applicability, Representation Scale, and Limits of Interpretation

The methodological relevance of the case study lies in demonstrating the diagnostic utility of the FLOOD-ST workflow rather than in establishing a universally transferable model configuration. The framework does not prescribe an identical hydraulic setup for every catchment. Its general elements comprise explicit stressor definition, controlled scenario design, transparent scenario labelling, consistent indicator extraction, and bounded interpretation, whereas their practical implementation remains catchment-specific.

General framework-level claims remain grounded in the underlying methodological work [11]. The present study supports a more specific application-level claim: within the documented Roßbach scenario space, controlled matched contrasts reveal differences in magnitude, timing, duration, and spatial redistribution that inundation extent alone does not capture. The analysis therefore supports condition-specific interpretation within the simulated scenario set, but it does not constitute a formal global sensitivity analysis, variance decomposition, or statistically complete ranking of stressor importance.

This limitation follows directly from the scenario design. The five investigated dimensions define 108 theoretical combinations, of which 23 were simulated. The selected ensemble was designed to provide interpretable baseline and compound scenarios rather than statistical coverage of the complete parameter space. Within this subset, 45 direct comparisons can be constructed in which one investigated condition changes while the remaining scenario settings are identical. These matched contrasts improve interpretability of pairwise differences, but they do not establish the variance contribution or global importance of individual stressors across the complete theoretical scenario space.

Representation scale also affects interpretation. The investigated asset classes differ strongly in spatial granularity and numerical size. Street cells are much more numerous and spatially detailed than buildings or critical-infrastructure objects; consequently, a pronounced street-level response should not be interpreted as directly equivalent to a proportional consequence for buildings or critical infrastructure. At the same time, street-level differences are not merely a counting artefact. In an urban channel–street system, changes in street affectedness can indicate shifts in flow pathways, accessibility, and hydraulic connectivity. Buildings translate these hydraulic differences into asset-level affectedness, while critical infrastructure provides a smaller but functionally relevant diagnostic layer.

Critical infrastructure requires similar caution. Because only 58 critical-infrastructure objects were included, the corresponding results should not be interpreted statistically in the same way as results for streets or buildings. Their relevance lies primarily in identifying functionally important individual cases and scenario-specific changes. This is consistent with the broader understanding that flood impacts on infrastructure depend not only on direct exposure, but also on interdependencies, function, and network context [19].

The same bounded interpretation applies to the hydrodynamic model configuration. The stress-test configuration builds on the previously developed Roßbach TELEMAC-2D model, but the present setup includes stress-test-specific rainfall inputs, boundary conditions, and hydraulic and operational modifications. Quantitative performance statistics from earlier model-development configurations are therefore not transferred directly to the

present setup, and no independent calibration or validation dataset was available for the complete stress-test configuration. Similarly, no mesh-convergence analysis was available for the runtime computational mesh. The FRHM and HRHM comparisons consequently provide contextual spatial benchmarks rather than independent model validation. These limitations primarily constrain interpretation of absolute hydraulic quantities; relative differences between consistently configured matched simulations constitute the stronger evidence provided by this study.

Additional uncertainty arises from the representation of the synthetic rainfall events. SCB, MCE, and MCS were intentionally normalised to approximately 120 mm catchment-mean precipitation to isolate differences associated with their spatial and temporal structure. Their RSI classification confirms that the imposed intensities fall within an extreme-heavy-rainfall range, but this does not imply that the complete synthetic fields represent equally probable or historically observed events. Moreover, the 100 m rainfall grid is a scenario-generation scale rather than an observational resolution, and no independent rainfall-grid resolution sensitivity analysis was conducted. Absolute inundation and affectedness values may therefore depend on the selected rainfall representation, even though we held its resolution constant across all direct scenario comparisons.

The severe clogging configuration requires a similarly bounded interpretation. It represents an intentionally adverse synthetic infrastructure state derived from preliminary hydraulic screening and not an empirically calibrated blockage probability or universally representative clogging fraction. The preliminary tests demonstrate that different structure configurations produce different hydraulic responses, but they do not constitute a systematic probabilistic sensitivity analysis across blockage severities. Consequently, the clogging scenarios are used to diagnose system response under an explicitly severe infrastructure condition rather than to predict the expected real-world frequency or magnitude of blockage.

The finite 12 h observation period also limits the interpretation of cumulative affectedness. The simulations were intentionally extended well beyond the three-hour rainfall period to capture delayed short-duration flood responses, but the 12 h endpoint is not interpreted as evidence of complete hydraulic convergence. In particular, cumulative street affectedness continued to increase slightly towards the end of the observation period, and several building series reached their final observed cumulative value only late in the simulation. Cumulative affectedness therefore describes affectedness within the fixed 12 h analysis window. The endpoint diagnostics provided in Supplementary Table S8 quantify the remaining late-period changes and should be considered when comparing cumulative metrics.

Object-level affectedness additionally depends on application-specific classification choices. The supplementary robustness analysis shows that varying the affected-cell threshold between 2.5%, 5%, and 10% and comparing the adaptive 1–2 m buffer with a fixed 2 m buffer changes absolute building counts. However, across all six tested configurations, the BL05–CS11 net difference remained positive (+12 to +29 buildings) and gross turnover remained substantially larger (50–68 buildings). The exact magnitude of object-level affectedness is therefore parameter-dependent, whereas the qualitative interpretation of spatial redistribution in this matched contrast is robust to the tested classification settings (Supplementary Table S6).

The Froude-based component requires particular caution. Its contribution depends on the selected minimum-depth criterion and upper plausibility limit and should not be interpreted as an independent consequence metric. The combined classification should therefore be understood as a multi-indicator screening approach in which exceeding any

one threshold is sufficient to indicate classified affectedness. Froude-specific results are interpreted only as relative hydraulic patterns within the fixed analysis configuration.

4.3. Practitioner Relevance and Future Extensions

The practical relevance of FLOOD-ST lies in its ability to translate hydrodynamic simulations into structured diagnostic evidence. Many operational questions are not limited to whether an area is inundated under a reference condition. Practitioners may also need to understand which investigated conditions are associated with larger response differences, where affectedness emerges first, which pathways repeatedly become affected, and how infrastructure states alter the spatial or temporal distribution of affectedness. The presented results show that FLOOD-ST can support this reasoning by combining controlled scenario design with area- and object-based affectedness indicators and transparent comparisons.

For water-management authorities and infrastructure operators, the Roßbach case study illustrates how different dimensions of system response can be distinguished. Within the investigated contrasts, hydrometeorological event type and antecedent moisture condition produced comparatively large changes in magnitude and timing, whereas infrastructure-state perturbations were more often expressed through smaller aggregate changes, redistribution, or timing shifts. Such information can support the prioritisation of further model refinement, local inspection, maintenance, or operational review, particularly where total inundation extent alone provides limited diagnostic information.

For emergency planning and municipal actors, the study highlights the relevance of timing and pathway redistribution. Earlier affectedness can reduce the time available for local response, even when the total number of additionally affected buildings remains small. Street-level redistribution may also be operationally relevant because streets can function as access routes, evacuation corridors, and pathways for response organisations. FLOOD-ST can therefore indicate where more detailed operational assessment may be warranted, but it does not directly prescribe emergency actions.

The BL05–CS11 matched-pair analysis is also relevant from a practitioner perspective because it illustrates how small aggregate changes can coincide with operationally relevant spatial and temporal redistribution. Even where cumulative totals changed only slightly, the timing of affectedness and the spatial identity of affected streets or buildings shifted measurably. Such redistributive effects may therefore warrant closer examination in emergency planning, maintenance prioritisation, and local inspection, even when changes in total inundation extent remain small.

In the Roßbach application, the cross-scenario hotspot analysis translated the 23-scenario ensemble into spatially explicit screening information. The identified street hotspot near the Martener Straße detention basin and the building hotspot in the eastern industrial area illustrate how recurrence identifies locations repeatedly affected across different scenario configurations, while conditional mean affectedness duration provides complementary information on persistence. Conversely, locations and assets with zero recurrence remained consistently unaffected within the analysed 23-scenario ensemble and under the applied thresholds. These outputs can support the prioritisation of locations for field inspection, maintenance, access-route assessment, or further local validation. Their interpretation, however, remains conditional on the investigated scenario ensemble, model configuration, observation period, and affectedness definitions.

The asset-class-specific indicators further demonstrate how the FLOOD-ST concept can be operationalised for consequence-oriented screening in the Roßbach case study. Streets provide information on pathway and accessibility changes, buildings translate hydraulic changes into asset-level affectedness, and critical infrastructure adds a functionally relevant perspective. Together with the controlled variation in hydrometeorological and

infrastructure conditions, these outputs provide a practical example of stress testing as conceptually framed in the FLOOD-ST Technical Note [11] and within the DPSIR framework [26,27], without constituting a complete DPSIR analysis, damage assessment, or operational risk assessment.

Future extensions should build on this diagnostic function without weakening the discipline of controlled comparison. FLOOD-ST is well suited for future measure testing when interventions or model adaptations are compared under otherwise consistent conditions. Coupling with damage models, accessibility or transport models, sewer-system models, or additional hydrological components could further increase practical relevance. Such extensions should not simply add complexity, but preserve the core principles of controlled contrasts, transparent scenario labelling, consistent indicators, and bounded interpretation.

The overall practitioner value of FLOOD-ST is therefore not that it reduces flood-system complexity to a single indicator. Its value lies in making different response dimensions explicit and comparable. By showing how investigated hydrometeorological conditions and infrastructure states are associated with changes in affectedness, timing, persistence, and hydraulic pathways, the workflow supports a more operationally meaningful interpretation of flood-system response. It complements hazard mapping and expert judgement by providing structured evidence on where response differences occur, how they manifest, and where more detailed analysis is warranted.

5. Conclusions

This study demonstrates the case-study-based diagnostic application of Flood-Oriented Operational Stress Testing (FLOOD-ST) in the Roßbach catchment using an established TELEMAC-2D rain-on-grid model. A deliberately selected 23-scenario ensemble, controlled matched comparisons, and consistent hydraulic and affectedness-based indicators were combined to examine how changes in hydrometeorological event type, antecedent moisture condition, and infrastructure state are associated with differences in flood-system response.

Within the investigated scenario space, changes in hydrometeorological event type and antecedent moisture condition produced more pronounced differences in affectedness magnitude and timing, whereas infrastructure-state perturbations more often produced smaller aggregate changes and context-dependent spatial or temporal redistribution. The BL05–CS11 matched comparison illustrates this diagnostic value particularly clearly: relatively small net changes coincided with substantially larger spatial turnover and shifts in the timing and identity of affected streets and buildings. Cross-scenario recurrence and affectedness-duration analyses further identified locations repeatedly affected across the ensemble as well as locations that remained consistently unaffected under the investigated scenarios and applied thresholds. These results show that inundation extent alone does not capture all diagnostically relevant dimensions of flood-system response.

The scientific contribution of the study lies in operationalising the existing FLOOD-ST framework within a detailed hydrodynamic case study and demonstrating how controlled scenario contrasts can be translated into spatially and temporally explicit evidence for further flood-management assessment. The workflow can complement conventional hazard mapping by supporting the prioritisation of locations and system states for closer inspection, maintenance, operational review, or more detailed risk analysis. The findings remain conditional on the investigated scenario ensemble, model configuration, synthetic rainfall representation, observation period, and affectedness definitions; they do not constitute a formal global sensitivity analysis, independent validation of the hydraulic model, or evidence of universal transferability. Future applications should therefore test the ap-

proach across additional catchments and model configurations while preserving controlled comparisons, transparent scenario definition, and bounded interpretation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/w18202604/s1. Table S1: Theoretical 108-combination FLOOD-ST design space across event type, AMC, clogging, pump and detention-basin factors, indicating the 23 simulated combinations; Table S2: Complete response table for all 45 admissible one-condition matched contrasts, evaluated from all 145 complete stored records; Table S3: Spatial and temporal metrics of the SCB, MCE and MCS rainfall archetypes; configured timestamps from 0 to 180 min are distinguished from the 185-min frame-hold integration convention; Table S4: Parameter and approximate hydraulic pretest evidence for the clogging configurations; discharge metrics originate from reproducible plot digitisation and are not field validation; Table S5: Full pump-state and detention-basin $\times 0.7$ matched contrasts by asset class and response metric; Table S6: BL05–CS11 building and critical-infrastructure classification robustness for two buffer rules and three affected-cell thresholds using all 145 records; Table S7: Asset-dataset provenance, processing and known metadata limitations; Table S8: Temporal endpoint diagnostics at 10, 11 and 12 h and at the final stored record for all scenarios and asset classes; Table S9: All-scenario spatial overlap of FLOOD-ST inundation with FRHM and HRHM contextual references.

Author Contributions: Conceptualization, J.R.; methodology, J.R. and D.J.R.; investigation, J.R. and D.J.R.; formal analysis, J.R. and D.J.R.; data curation, D.J.R.; visualisation, D.J.R. and J.R.; writing—original draft preparation, J.R.; writing—review and editing, J.R., D.J.R., J.H., A.A., and C.B.; supervision, J.H. and C.B.; project administration, C.B.; funding acquisition, C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Interreg North-West Europe Programme through the Flash-FloodBreaker project (project ID NWE0200167), co-funded by the European Regional Development Fund (ERDF).

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. The complete hydrodynamic model setup and underlying geospatial input datasets cannot be made publicly available as a single repository because parts of the input data were provided by third parties and project partners and are subject to data-use and licensing restrictions. Derived simulation results and analysis data generated within this study can be made available by the authors upon reasonable request, subject to the applicable data-use conditions.

Acknowledgments: The authors gratefully acknowledge the support of all partners involved in the FlashFloodBreaker project. The authors particularly acknowledge David Jonathan Roschinsky for the contribution of results generated within his master's thesis, which provided an essential basis for the presented case-study analysis and visualisation of simulation results. During the preparation of this manuscript, the authors used ChatGPT (OpenAI; GPT-5.5 and GPT-5.6) for assistance with Python-based data processing and visualisation scripts as well as language editing and manuscript refinement. The authors reviewed and edited all outputs and take full responsibility for the content of this publication.

Conflicts of Interest: Author Julian Hofmann was employed by the company FloodWaive Predictive Intelligence GmbH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
AMC	Antecedent Moisture Condition

BL	Baseline scenario
CS	Compound stress-test scenario
DEM	Digital Elevation Model
DPSIR	Driver–Pressure–State–Impact–Response framework
DWD	German Weather Service (Deutscher Wetterdienst)
EGLV	Emschergenossenschaft und Lippeverband
ERDF	European Regional Development Fund
FEM	Finite Element Method
FLOOD-ST	Flood-Oriented Operational Stress Testing
FRHM	Flood Risk Hazard Map
GIS	Geographic Information System
HRB	Hochwasserrückhaltebecken (flood detention basin)
HRHM	Heavy-Rainfall Hazard Map
MCE	Multi-Cell Event
MCS	Moving Convective Storm
NRW	North Rhine-Westphalia
QGIS	Quantum Geographic Information System
RWTH	RWTH Aachen University
SCB	Stationary Cloud Burst
SCS-CN	Soil Conservation Service Curve Number method
TELEMAC-2D	Two-dimensional hydrodynamic modelling system within the TELEMAC-MASCARET system

References

1. Mohr, S.; Ehret, U.; Kunz, M.; Ludwig, P.; Caldas-Alvarez, A.; Daniell, J.E.; Ehmele, F.; Feldmann, H.; Franca, M.J.; Gattke, C.; et al. A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe—Part 1: Event description and analysis. *Nat. Hazards Earth Syst. Sci.* **2023**, *23*, 525–551. [CrossRef]
2. Ludwig, P.; Ehmele, F.; Franca, M.J.; Mohr, S.; Caldas-Alvarez, A.; Daniell, J.E.; Ehret, U.; Feldmann, H.; Hundhausen, M.; Knippertz, P.; et al. A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe—Part 2: Historical context and relation to climate change. *Nat. Hazards Earth Syst. Sci.* **2023**, *23*, 1287–1311. [CrossRef]
3. Lehmkuhl, F.; Schüttrumpf, H.; Schwarzbauer, J.; Brüll, C.; Dietze, M.; Letmathe, P.; Völker, C.; Hollert, H. Assessment of the 2021 summer flood in Central Europe. *Environ. Sci. Eur.* **2022**, *34*, 107. [CrossRef]
4. Guo, K.; Guan, M.; Yu, D. Urban surface water flood modelling—A comprehensive review of current models and future challenges. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 2843–2860. [CrossRef]
5. Thorndahl, S.; Einfalt, T.; Willems, P.; Nielsen, J.E.; ten Veldhuis, M.-C.; Arnbjerg-Nielsen, K.; Rasmussen, M.R.; Molnar, P. Weather radar rainfall data in urban hydrology. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 1359–1380. [CrossRef]
6. Púčik, T.; Groenemeijer, P.; Rädler, A.T.; Tijssen, L.; Nikulin, G.; Prein, A.F.; van Meijgaard, E.; Fealy, R.; Jacob, D.; Teichmann, C. Future Changes in European Severe Convection Environments in a Regional Climate Model Ensemble. *J. Clim.* **2017**, *30*, 6771–6794. [CrossRef]
7. Brunner, M.I.; Swain, D.L.; Wood, R.R.; Willkofer, F.; Done, J.M.; Gilleland, E.; Ludwig, R. An extremeness threshold determines the regional response of floods to changes in rainfall extremes. *Commun. Earth Environ.* **2021**, *2*, 173. [CrossRef]
8. Tabari, H. Climate change impact on flood and extreme precipitation increases with water availability. *Sci. Rep.* **2020**, *10*, 13768. [CrossRef] [PubMed]
9. Zscheischler, J.; Westra, S.; van den Hurk, B.J.J.M.; Seneviratne, S.I.; Ward, P.J.; Pitman, A.; AghaKouchak, A.; Bresch, D.N.; Leonard, M.; Wahl, T.; et al. Future climate risk from compound events. *Nat. Clim. Change* **2018**, *8*, 469–477. [CrossRef]
10. Bevacqua, E.; Suarez-Gutierrez, L.; Jézéquel, A.; Lehner, F.; Vrac, M.; Yiou, P.; Zscheischler, J. Advancing research on compound weather and climate events via large ensemble model simulations. *Nat. Commun.* **2023**, *14*, 2145. [CrossRef] [PubMed]
11. Reinert, J.; Klopries, E.-M.; Hassenjürgen, P.; Menz, A.; Hofmann, J.; Brüll, C.; Klein, A.; Schüttrumpf, H. Technical note: Quantifying System Limits: Flood Oriented Operational Stress Testing under Intensified Precipitation Regimes in Central Europe. *EGUsphere* **2026**, *2026*, 1–21. [CrossRef]
12. European Parliament; Council of the European Union. Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the Assessment and Management of Flood Risks. *Off. J. Eur. Union* **2007**, *L 288*, 27–34.
13. Reinert, J.; Hofmann, J.; Almoradie, A.; Brüll, C. Assessing the Challenges of Urban Flood Modelling: A Sensitivity Analysis Using a TELEMAC-2D Rain-on-Grid Framework in the Emscher Catchment. *Water* **2026**, *18*, 1224. [CrossRef]

14. Mizukami, N.; Rakovec, O.; Newman, A.J.; Clark, M.P.; Wood, A.W.; Gupta, H.V.; Kumar, R. On the choice of calibration metrics for "high-flow" estimation using hydrologic models. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 2601–2614. [CrossRef]
15. Ran, Q.; Wang, J.; Chen, X.; Liu, L.; Li, J.; Ye, S. The relative importance of antecedent soil moisture and precipitation in flood generation in the middle and lower Yangtze River basin. *Hydrol. Earth Syst. Sci.* **2022**, *26*, 4919–4931. [CrossRef]
16. Peña, F.; Nardi, F.; Melesse, A.; Obeysekera, J.; Castelli, F.; Price, R.M.; Crowl, T.; Gonzalez-Ramirez, N. Compound flood modeling framework for surface–subsurface water interactions. *Nat. Hazards Earth Syst. Sci.* **2022**, *22*, 775–793. [CrossRef]
17. Eilander, D.; Couasnon, A.; Leijnse, T.; Ikeuchi, H.; Yamazaki, D.; Muis, S.; Dullaart, J.; Haag, A.; Winsemius, H.C.; Ward, P.J. A globally applicable framework for compound flood hazard modeling. *Nat. Hazards Earth Syst. Sci.* **2023**, *23*, 823–846. [CrossRef]
18. Pasquier, U.; He, Y.; Hooton, S.; Goulden, M.; Hiscock, K.M. An integrated 1D–2D hydraulic modelling approach to assess the sensitivity of a coastal region to compound flooding hazard under climate change. *Nat. Hazards* **2019**, *98*, 915–937. [CrossRef]
19. Khanam, M.; Sofia, G.; Koukoula, M.; Lazin, R.; Nikolopoulos, E.I.; Shen, X.; Anagnostou, E.N. Impact of compound flood event on coastal critical infrastructures considering current and future climate. *Nat. Hazards Earth Syst. Sci.* **2021**, *21*, 587–605. [CrossRef]
20. Giardini, D.; Mignan, D.; Pinto, A.; Tsionis, G. (Eds.) *Harmonized Approach to Stress Tests for Critical Infrastructures Against Natural Hazards*; Publications Office: Luxembourg, 2016.
21. Brown, C.; Ghile, Y.; Laverty, M.; Li, K. Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector. *Water Resour. Res.* **2012**, *48*, W09537. [CrossRef]
22. Whateley, S.; Steinschneider, S.; Brown, C. A climate change range-based method for estimating robustness for water resources supply. *Water Resour. Res.* **2014**, *50*, 8944–8961. [CrossRef]
23. Schmitt, T.G.; Krüger, M.; Pfister, A.; Becker, M.; Mudersbach, C.; Fuchs, L.; Hoppe, H.; Lakes, I. Einheitliches Konzept zur Bewertung von Starkregenereignissen mittels Starkregenindex. *KA Korresp. Abwasser Abfall* **2018**, *65*, 113–120. [CrossRef]
24. Martínez-Gomariz, E.; Gómez, M.; Russo, B. Experimental study of the stability of pedestrians exposed to urban pluvial flooding. *Nat. Hazards* **2016**, *82*, 1259–1278. [CrossRef]
25. Krumm, J.; Haag-Wanka, I.; Leistert, H.; Hänsler, A.; Steinbrich, A.; Schmit, M.; Weiler, M. *Konzeption, Ermittlung und Anwendungsmöglichkeiten des Sturzflutindexes (SFI)*; FreiDok plus, Universitätsbibliothek Freiburg: Freiburg im Breisgau, Germany, 2024. [CrossRef]
26. Smeets, E.; Weterings, R. *Environmental Indicators: Typology and Overview*; Technical Report No. 25; European Environment Agency: Copenhagen, Denmark, 1999.
27. Fang, X.; Lu, L.; Li, Y.; Hong, Y. A Driver-Pressure-State-Impact-Response study for urban transport resilience under extreme rainfall-flood conditions. *Transp. Res. Part D Transp. Environ.* **2023**, *121*, 103819. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.