Hydrodynamic Simulation and Geospatial Data Architecture for Urban Flood Prediction in the National Capital Territory of Delhi: A Computational Framework for AI-Driven Resilience

The National Capital Territory (NCT) of Delhi occupies a uniquely precarious hydrological position, functioning as a high-density urban agglomeration situated within the semi-arid northern plains of India, yet subject to the extreme seasonal dynamics of the Southwest Monsoon and the regulated discharge of the Yamuna River. Developing an artificial intelligence (AI) system for drainage and waterlogging prediction requires a multi-scalar data architecture that integrates high-resolution topographic surfaces, complex meteorological disaggregation, and the decaying hydraulic efficiency of aging infrastructure. This report provides an exhaustive technical analysis of the parameters required to generate a synthetic but realistic 10,000-row dataset for such an AI engine, grounded in a 250m × 250m grid-cell resolution. By synthesizing meteorological records, soil mechanics, and urban morphological theories, the following analysis establishes the logical and physical constraints necessary for robust predictive modeling of Delhi's urban flood landscape.[1]

Geospatial Foundations and Topographic Modeling

The spatial domain for the Delhi flood prediction engine is strictly bounded between the latitudes of 28.2°N to 29.0°N and the longitudes of 76.85°E to 77.65°E, covering the core metropolitan area and its satellite cities, including Gurgaon, Noida, and Ghaziabad.[1] Within this domain, the elevation profiles vary from a minimum of approximately 213 meters above mean sea level in the southern floodplains to over 305 meters along the rocky outcrops of the Delhi Ridge.[5] The topography is not a uniform plane; rather, it is divided into three distinct physiographic regions: the Yamuna floodplains, the Ridge (an extension of the Aravalli Range), and the expansive alluvial plains.[5]

The master slope of Delhi exhibits a gentle southerly gradient, which is the primary driver for the southerly flow of the Yamuna River.[7] However, significant localized reversals occur. East of the Yamuna, the general slope is westerly, directed toward the river channel.[7] In the western sectors, the Ridge acts as a local watershed divide, separating the drainage into two primary catchments: the eastern region, which drains directly into the Yamuna, and the western region, which converges into the Najafgarh Drain system.[5] For the purpose of the 250m grid-cell simulation, each cell must be assigned a slope_percent derived from these regional master slopes, where a lower slope percentage (e.g., <0.5%) directly correlates with an increased probability of waterlogging during high-intensity events.[9]

Table 1: Regional Topographic and Geomorphological Characteristics

Region	Elevation Range (m)	Typical Slope (%)	Drainage Basin	Soil Character
Yamuna Floodplains	213 - 215	0.1 - 0.3	Shahadra / Alipur	Fertile Alluvium
Delhi Ridge	230 - 305	2.5 - 5.0	Kushak-Barrapullah	Rocky / Quartzite
Alluvial Plains	218 - 228	0.5 - 1.2	Najafgarh	Sandy Silt / Loam
Peripheral North	214 - 218	0.2 - 0.5	Alipur	Silty Clay

5

Meteorological Drivers and Precipitation Disaggregation

Delhi's hydrological cycle is dominated by the Southwest Monsoon, which historically arrives around late June and withdraws by late September.[7] In the 2025 monsoon season, the onset occurred on June 29th, resulting in a seasonal total of 902.6 mm at the Safdarjung observatory, representing an 41% departure from the long-period average of 640.4 mm.[11] For a predictive AI to function, daily rainfall data must be disaggregated into sub-hourly intervals. The MCIMD (Modified Chowdhury Indian Meteorological Department) method provides the mathematical basis for this downscaling [12]:

$$R_t = R_{24} \cdot \left(\frac{t}{24}\right)^n + C$$

where R_t represents the rainfall depth for a duration t (in hours), R_{24} is the daily rainfall accumulation, n is an exponential constant typically calibrated to $1/3$, and C is a constant reflecting local urban atmospheric conditions.[12] The simulation engine must ensure that for every grid cell, the rainfall values obey the physical law of temporal accumulation:

$rain\_15min \le rain\_1h \le rain\_3h \le rain\_6h \le rain\_24h$.

The 2025 season recorded 1 "Very Heavy" rainfall day (115.6 to 204.4 mm), 10 "Heavy" days (64.5 to 115.5 mm), and 36 "Moderate" days.[11] For instance, on July 30th, 2025, the Ridge observatory recorded 129.8 mm of rain, an event that would likely overwhelm any drainage infrastructure designed for the 1976 standard of 50 mm in 24 hours.[6]

Table 2: Observed Precipitation Events in Delhi (2025 Monsoon)

Date	Observatory	Rainfall (mm)	Classification
July 1, 2025	Safdarjung	117.2	Very Heavy
July 7, 2025	Najafgarh	95.0	Heavy
July 10, 2025	Dhansa	107.6	Heavy
July 30, 2025	Ridge	129.8	Very Heavy
August 21, 2025	Ayanagar	145.0	Very Heavy

11

Hydrodynamics of the Yamuna River and Backflow Risks

The Yamuna River is the primary drainage artery of Delhi, yet it poses a dual risk: direct riverine flooding of the floodplains and the "backflow" effect on the city's internal drainage network.[6] The water level at the Old Railway Bridge (ORB) is largely a function of discharge from the Hathnikund Barrage, located 240 km upstream in Haryana.[14] The transit time for water released at Hathnikund to reach Delhi is approximately 36 to 50 hours, providing a critical window for predictive AI alerts.[15]

The risk indices for the Yamuna are defined by specific elevation markers:

- **Warning Level**: 204.50 meters.[15]
- **Danger Level**: 205.33 meters.[16]
- **Evacuation Level**: 206.00 meters.[16]

During extreme events, such as the record peak of 208.66 meters in July 2023, the river level exceeds the discharge outlets of major city drains.[6] This reversal of the hydraulic gradient means that even if city pumps are operational, the sheer volume of river water prevents the gravity-fed discharge of urban runoff, leading to extensive waterlogging in low-lying areas like Civil Lines, Yamuna Bazar, and the Monastery Market.[16] For the AI dataset, the yamuna_risk_index must be modeled as a non-linear function of the distance to the river and the current river stage relative to the danger mark.[15]

Urban Morphological Parameters and Infiltration Decay

The transformation of Delhi from a series of settlements into a megacity has fundamentally altered its surface permeability. Urbanization is represented by the growth of impervious surfaces, with a 36.49% increase in built-up area observed between 2005 and 2016.[20] In densely populated wards such as Old Delhi (density: 30.3 thousand/km²) and the North-East district (density: 29.5 thousand/km²), the impervious_ratio often exceeds 0.90, leaving minimal area for natural infiltration.[7]

The simulation of soil infiltration follows Horton’s Infiltration Model, which accounts for the exponential decay of a soil's capacity to absorb water as it reaches saturation [22]:

$$f_p = f_c + (f_0 - f_c) e^{-kt}$$

where f_p is the infiltration capacity at time t, f_c is the final equilibrium capacity, f_0 is the initial capacity, and k is the decay constant.[24] In Delhi, the soil types range from sandy silt in the floodplains (Group A/B) to silty clays in the peripheral zones (Group C/D).[5]

Table 3: Typical Horton Model Parameters by SCS Soil Group

Soil Group	Initial Rate f0 (mm/hr)	Final Rate fc (mm/hr)	Decay Constant k (1/hr)
Group A (Sand/Gravel)	250	25.4	2.0
Group B (Sandy Loam)	200	12.7	2.0
Group C (Clay Loam)	125	6.3	2.0

Group D (Clay/Silty)	76	2.5	2.0

[22]

The soil_absorption_rate in the synthetic dataset must reflect these parameters, with a significant reduction in f_0 for cells located in highly compacted urban areas, as rural-to-urban transfer of these values is generally advised against due to compaction and vegetation differences.[22]

Drainage Infrastructure and Systematic Bottlenecks

The NCT of Delhi is divided into five major drainage basins: Najafgarh, Alipur, Shahadra, Kushak-Barrapullah, and Mehrauli.[5] The network includes 426.55 km of natural drainage lines and 3,311.54 km of engineered storm water drains.[28] However, the efficiency of this network is severely compromised by several systemic factors. A primary issue is that Delhi's drainage system, largely designed in 1976, was built to handle a rainfall intensity of only 50 mm in 24 hours.[6]

The drain_blockage_risk factor in the AI model is influenced by four primary stressors:

1. **Sewage Contamination**: The discharge of sewage into storm drains is common, as the Delhi Jal Board sometimes punctures sewer lines during blockages to prevent local overflows, inadvertently causing siltation in the storm network.[28]
2. **Solid Waste Accumulation**: Poor municipal waste management leads to the accumulation of plastic and other debris at drain inlets.[9]
3. **Encroachments**: Illegal construction over drains prevents routine maintenance and reduces cross-sectional capacity.[6]
4. **Sedimentation**: Studies using UAV bathymetry on the Ghazipur Drain revealed that sedimentation and solid waste accumulation have reduced channel capacity by as much as 25%.[9]

For the synthetic dataset, drain_density must be higher in the planned areas of central Delhi and lower in the peripheral zones where unauthorized colonies have emerged without formal drainage planning.[3]

Table 4: Drainage Basins of Delhi and Primary Outfalls

Basin Name	Primary Drain	Catchment Character	Vulnerability
Najafgarh Basin	Najafgarh Drain	Western Delhi /	Sahibi River

		Rural	Overspill
Shahadra Basin	Ghazipur Drain	Eastern Delhi / Dense	Siltation / High Waste
Barrapullah Basin	Kushak Nallah	Central / South Delhi	Encroachment
Alipur Basin	Alipur Drain	North Delhi / Semi-Rural	Low Gradient
Mehrauli Basin	Local Nallahs	South Delhi / Hilly	Flash Flooding

5

The Inverted Compact City Paradox and Socio-Technical Risks

A critical insight for the "Delhi Drainage & Waterlogging Prediction AI" is the concept of the "inverted compact" city. Contrary to conventional urban models where density is highest at the core, Delhi's population density is significantly higher in its peripheral zones than in its urban center (Lutyens' Delhi).[3] Densities in some peripheral clusters range as high as 136,385 persons per km², while the core remains relatively low at 2,884 persons per km².[3]

This spatial arrangement creates a massive disparity in drainage provision. The core areas, with their low population density and high-quality colonial-era planning, have surplus drainage capacity. In contrast, the peripheral zones—where 30% of Delhi's population lives in the Trans-Yamuna area alone—face uncoordinated development and failing infrastructure.[3] Consequently, the dataset must reflect a flood_frequency that is higher in these peripheral, high-density, low-drainage-density wards.[3]

Administrative Zoning and Ward Delimitation

As of the 2025-2026 period, the Municipal Corporation of Delhi (MCD) consists of 250 wards categorized into 12 administrative zones.[29] These wards serve as the primary unit for municipal flood reporting and response. For the 250m grid simulation, each cell is mapped to a ward_name and zone to allow the AI to generate actionable insights for specific municipal teams.[29]

Table 5: Sample MCD Wards and Population Metrics for Risk Modeling

Ward No.	Ward Name	Zone	Total Voters (approx.)	Polling Stations
1	Minto Road	City-SP	28,761	52
33	Rohini	Rohini	155,011	190
48	Uttam Nagar	West	144,163	164
101	Narela	Narela	65,650	68
104	Burari	Civil Lines	107,957	123
109	Darya Ganj	City-SP	42,108	48

[29]

The citizen_reports_count and avg_reported_depth_cm are modeled as social sensors. Areas with higher voter density and population (e.g., Ward 33 - Rohini) are likely to generate more reports for a given flood depth compared to low-density industrial or rural wards.[29]

Dataset Generation Logic: Hydrological and Physical Constraints

To satisfy the realism constraints of the simulation, the relationship between input variables (rainfall, topography) and output labels (flood_occurred) is governed by a multi-variate logical function. A grid cell is flagged as flood_occurred = 1 if it meets a combination of the following criteria:

1. **Pluvial Saturation**: rain_1h_mm > 40mm AND impervious_ratio > 0.85 AND drain_capacity_score < 2. This represents flash flooding due to high-intensity rain on non-absorbent surfaces with poor drainage.[2]
2. **Topographic Accumulation**: elevation_m is at the 10th percentile for the basin AND flow_accumulation > 5000 AND rain_6h_mm > 60mm. This represents water pooling in low-lying natural sinks.[5]
3. **Riverine Backflow**: distance_to_yamuna_m < 2000 AND yamuna_level_m > yamuna_danger_level_m (205.33) AND rain_24h_mm > 30mm. This represents the scenario where high river levels prevent local drainage.[14]

4. **Infrastructure Failure**: drain_blockage_risk > 0.8 AND rain_3h_mm > 35mm. This accounts for floods caused by choked drains even in areas with moderate topography.[9]

Conversely, the model ensures that slope_percent has an inverse relationship with flooding probability, as steeper terrain facilitates faster runoff into the drainage network, provided the downstream capacity exists.[7]

Advanced Analytical Insights: Second and Third Order Effects

The Synchronization of Upstream Release and Local Storms

The most catastrophic flood scenarios in Delhi are not caused by local rain alone, but by the synchronization of heavy local precipitation with the arrival of a peak discharge wave from the Hathnikund Barrage. Since the lag time is 48 hours, a "perfect storm" occurs when a heavy rainfall event hits Delhi exactly as the river reaches the 206m mark.[14] This synchronization overwhelms the temporal buffer of the city's drainage network, leading to max_flood_depth_cm values that can exceed 100 cm in low-lying residential clusters.[6]

The Impact of Wetland Loss on Peak Flow Velocity

The decline of water bodies from 800 to 600 has a third-order effect on the flow_accumulation and peak flow velocity.[5] In a balanced system, wetlands act as shock absorbers that increase the time-of-concentration (t_c) for a basin. The loss of these wetlands, particularly in the South and North-West districts, has led to a "sharper" hydrograph where floodwaters reach their maximum depth much faster than they did four decades ago, reducing the lead time available for evacuation and emergency response.[20]

Socio-Economic Feedback Loops in Data Reporting

The citizen_reports_count variable introduces a fascinating socio-economic dimension to the dataset. In middle-to-high income wards like South Extension or Vasant Vihar, there is often a higher frequency of reporting for lower flood depths compared to low-income informal settlements where residents may have a higher threshold for disruption or less access to digital reporting platforms.[3] This reporting bias must be accounted for in the AI training to ensure that resources are not disproportionately diverted from the most vulnerable zones simply because they are "under-reporting" their distress.

Conclusion and Strategic Recommendations for AI Integration

The synthetic dataset generated through this hydrological simulation framework provides a high-fidelity environment for the development of "Delhi Drainage & Waterlogging Prediction

AI." By integrating the 125-year history of monsoon onset, the specific elevation thresholds of the Yamuna, and the 250m grid-level granularity of Delhi's ward structure, the model moves beyond simple pattern recognition to true physics-based prediction.[11]

To achieve operational success, the following architectural paradigms must be adopted:

1. **Transition to LID**: The simulation proves that reliance on grey infrastructure (drains) is insufficient. The AI should prioritize areas for Low Impact Development (LID), such as permeable pavements and infiltration trenches, based on the Horton model's f_c values for each cell.[22]
2. **Dynamic Thresholding**: Rather than a static "danger mark," the AI must use dynamic thresholds that consider the drain_blockage_risk and yamuna_risk_index in tandem, recognizing that a "low" river level can still cause flooding if drains are 25% choked by sediment.[9]
3. **Multi-Basin Coordination**: Management must move from ward-level response to basin-level coordination. As the five drainage basins of Delhi are hydraulically linked via the Yamuna, an overflow in the Najafgarh basin can influence the backflow dynamics in the Shahadra basin.[5]

By applying these insights, the NCT of Delhi can leverage AI to transform from a city that reacts to seasonal disasters into a resilient urban system capable of navigating the complex hydrological challenges of the 21st century.

Works cited

1. Delhi - District Boundaries - UrbanEmissions.Info, accessed January 10, 2026, https://urbanemissions.info/delhi-air-quality-forecasts/delhi-districts/
2. Identification of Critical Hotspots in Urban Drainage Networks using MIKE URBAN, accessed January 10, 2026, https://www.chijournal.org/H539
3. Impact of urban density on human well-being and sustainable development in Delhi, India, accessed January 10, 2026, https://pmc.ncbi.nlm.nih.gov/articles/PMC12485110/
4. NCR Constituent Areas - National Capital Region Planning Board, accessed January 10, 2026, https://ncrpb.nic.in/ncrconstituent.html
5. Floods in Delhi: causes and challenges, accessed January 10, 2026, https://www.toi.no/getfile.php/1348352-1530708215/Publikasjoner/Flooding_Delhi.pdf
6. Evaluating Drainage Infrastructure and Flood Resilience in Delhi city - International Journal of Humanities Social Science and Management (IJHSSM), accessed January 10, 2026, https://ijhssm.org/issue_dcp/Urban%20Drainage%20Management%20Evaluating%20Drainage%20Infrastructure%20and%20Flood%20Resilience%20in%20Delhi%20city.pdf
7. Urban Floods: Case Study of Delhi - NIDM, accessed January 10, 2026, https://nidm.gov.in/journal/PDF/Journal/Journal20091/Journal20091b.pdf

8. Natural Urban Drainage - Development Alternatives, accessed January 10, 2026, https://devalt.org/newsletter/sep98/of_7.htm
9. High-Resolution Waterlogging Mapping Along Ghazipur Drain in Delhi: A UAV-Based Bathymetric Analysis Approach, accessed January 10, 2026, https://neptjournal.com/upload-images/(43)B-4316.pdf
10. (PDF) Morphometric Analysis of Yamuna River Basin in Delhi Region: Implications to Hydrological Hazards - ResearchGate, accessed January 10, 2026, https://www.researchgate.net/publication/363295449_Morphometric_Analysis_of_Yamuna_River_Basin_in_Delhi_Region_Implications_to_Hydrological_Hazards
11. Government of India Ministry of Earth Sciences India Meteorological ..., accessed January 10, 2026, https://mausam.imd.gov.in/newdelhi/mcdata/seasonal_report.pdf
12. Modeling Rainfall Intensity-Duration-Frequency (IDF) and Establishing Climate Change Existence in Uyo-Nigeria Using Non-Stationary Approach - Scirp.org., accessed January 10, 2026, https://www.scirp.org/journal/paperinformation?paperid=124986
13. It has been published in some print media that Delhi has not got rainfall after the advance of monsoon. It is to inform that Del - India Meteorological Department, accessed January 10, 2026, https://internal.imd.gov.in/press_release/20220713_pr_1722.pdf
14. Floods | Delhi Disaster Management Authority, accessed January 10, 2026, https://ddma.delhi.gov.in/ddma/floods
15. Yamuna water level crosses warning mark in Delhi, hits 204.60 metre | India News, accessed January 10, 2026, https://www.business-standard.com/india-news/yamuna-crosses-warning-mark-in-delhi-as-barrage-water-release-raises-level-125081700738_1.html
16. Yamuna's Water Above Danger Mark, But Level On The Decline | Delhi News, accessed January 10, 2026, https://timesofindia.indiatimes.com/city/delhi/yamunas-water-above-danger-mark-but-level-on-the-decline/articleshow/123393151.cms
17. Yamuna crosses danger mark in Delhi, flows above 205.33 metres - The Tribune, accessed January 10, 2026, https://www.tribuneindia.com/news/delhi/yamuna-crosses-danger-mark-in-delhi-flows-above-205-33-metres/
18. Yamuna Flows Below Danger Mark, Evacuation Threshold Still at 206 Metres - Newsonair, accessed January 10, 2026, https://www.newsonair.gov.in/yamuna-flows-below-danger-mark-evacuation-threshold-still-at-206-metres/
19. Delhi: Yamuna crosses danger mark; residents in low | DD News, accessed January 10, 2026, https://ddnews.gov.in/en/delhi-yamuna-crosses-danger-mark-residents-in-low-lying-areas-shifted-to-safer-locations/
20. Predicting impact of urbanization on water resources in megacity Delhi - ResearchGate, accessed January 10, 2026, https://www.researchgate.net/publication/343674704_Predicting_impact_of_urba

nization_on_water_resources_in_megacity_Delhi
21. Impervious surface growth and its inter-relationship with vegetation cover and land surface temperature in peri-urban areas of Delhi | Request PDF - ResearchGate, accessed January 10, 2026, https://www.researchgate.net/publication/349879482_Impervious_surface_growth_and_its_inter-relationship_with_vegetation_cover_and_land_surface_temperature_in_peri-urban_areas_of_Delhi
22. Help | Horton Infiltration Model | Autodesk, accessed January 10, 2026, https://help.autodesk.com/view/IWICMS/2024/ENU/?guid=GUID-3AD0EFBA-92CC-481E-8C3B-7502D51C0F19
23. 1 Horton's Model The change in moisture content in soil with elevation is described by a soil retention curve. The change in m - Jagadish Torlapati, accessed January 10, 2026, https://www.torlapati.com/resources/wre/HortonsModel.pdf
24. The infiltration process was thoroughly studied by Horton in the early 1930s [4).An outgrowth of his work, shown graphically in - MSU College of Engineering, accessed January 10, 2026, https://www.egr.msu.edu/classes/ce421/lishug/text%20book.pdf
25. Horton Infiltration Model - InfoWorks ICM Online Help, accessed January 10, 2026, https://help2.innovyze.com/infoworksicm/Content/HTML/ICM_ILCM/Horton_Infiltration_Model.htm
26. Low permeability soils - LID SWM Planning and Design Guide, accessed January 10, 2026, https://wiki.sustainabletechnologies.ca/wiki/Low_permeability_soils
27. Infiltration Properties – Scalgo, accessed January 10, 2026, https://scalgo.com/en-US/scalgo-live-documentation/soil-vegetation-atmosphere-properties/soil/properties
28. Drainage Master Plan for NCT of Delhi - Irrigation and Flood Control, accessed January 10, 2026, https://ifc.delhi.gov.in/sites/default/files/inline-files/main_report_dmp_version51.pdf
29. List and No. of Polling Stations and Total No. of Electors | State Election Commission, NCT of Delhi, accessed January 10, 2026, https://sec.delhi.gov.in/sec/list-and-no-polling-stations-and-total-no-electors
30. FINAL INDIVIDUAL MAPS OF NEWLY DELIMITED WARDS (241 - State Election Commission, NCT of Delhi, accessed January 10, 2026, https://sec.delhi.gov.in/sec/final-individual-maps-newly-delimited-wards-241-250-maps
31. Final List of Polling Stations- MCD Bye Election 2025 | State Election Commission, NCT of Delhi, accessed January 10, 2026, https://sec.delhi.gov.in/sec/final-list-polling-stations-mcd-bye-election-2025
32. Delhi Zones | PDF | New Delhi - Scribd, accessed January 10, 2026, https://www.scribd.com/document/872133015/Delhi-Zones
33. InfoDrainage pipe design - Design rainfall theory and developing IDF curves - Autodesk, accessed January 10, 2026,

https://www.autodesk.com/learn/ondemand/curated/infodrainage-pipe-design/4JCngVfRUr0fctBpfmXpaU
34. On the Rainfall Intensity–Duration–Frequency Curves, Partial-Area Effect and the Rational Method: Theory and the Engineering Practice - MDPI, accessed January 10, 2026, https://www.mdpi.com/2073-4441/12/10/2730