538_Project: Flood Inundation Prediction under Extreme Rainfall Events in Hamilton City

Gongfan Zhang (1640121)

Question: When extreme rainfall occurs, where will flooding happen, and where can people go?

1. Introduction

As the climate continues to change, the frequency of extreme weather events is increasing. According to the Ministry for the Environment (2020), "Extreme weather events such as storms, heatwaves, and heavy rainfall are likely to become more frequent and intense. Large increases in extreme rainfall are expected across the country."

Given this trend, Hamilton City is also likely to face more frequent and intense rainfall events. This raises an important question: **When extreme rainfall occurs, where will flooding happen, and where can people go?**

To address this, I decided to use GIS applications to predict potential flood inundation areas in Hamilton City under extreme rainfall scenarios, estimate inundation extent and water depth, and identify suitable shelter locations for emergency response.

2. Previous research

In response to the problem posed in the introduction, a review of relevant literature was conducted to identify suitable modeling frameworks and spatial analysis techniques.

For example, Zheng et al. (2018) and Cohen et al. (2018) emphasize the integration of hydrological tools such as Flow Direction and Flow Accumulation with surface characteristics like land cover and soil drainage. Similarly, Tran and Nguyen (2025) present a clear workflow for combining high-resolution DEMs, land use, and soil data for flood inundation modeling. In this project, land use and soil weights were assigned based on Manning's n values referenced from Soliman et al. (2022).

In the modeling process, specific methodological choices were also guided by literature. Cohen et al. (2018) demonstrated that a 10-meter resolution DEM provides sufficient accuracy for large-scale flood depth estimation using remote sensing and terrain analysis, and recommend using Focal Statistics to smooth the resulting water depth surface.

To classify floodwater depth, a 0.5-meter interval was adopted following the approach of Tran and Nguyen (2025), who used this classification to improve flood visualization and interpretation in agricultural flood assessments.

For shelter planning, Kar and Hodgson (2008) highlighted that public facilities such as schools and hospitals are often selected as dual-purpose emergency shelters.

Together, these studies form the methodological backbone of this project, supporting the design and implementation of a GIS-based flood inundation areas assessment model from hydrological simulation to visualization and emergency planning.

3. My analysis process

There are several analysis steps in this project, and further details can be found in the accompanying Python scripts.

3.1 Download original datasets

This includes the Hamilton City boundary, DEM, soil, landcover, rainfall, and facility datasets. Detailed sources for each dataset are provided in *appendix 1: Dataset Sources*.

3.2 Pre-process datasets

The Hamilton City boundary is extracted and set as the analysis area for this project. The DEM is resampled from a 1-meter to a 10-meter resolution to improve processing speed and reduce noise during hydrological analysis. The soil and landcover datasets are reclassified into weight categories and merged into a single **composite weight layer** for use in the subsequent Flow Accumulation process. Details of this process are described in *appendix 2: Weight Assignment*.

3.3 Simulate surface water flow:

Apply hydrological tools to the pre-processed DEM in the following order: **Fill**, **Flow Direction**, and **Flow Accumulation** (The weight layer mentioned in the previous step is used to simulate surface drainage).

3.4 Estimate water depth and filter values:

Simulated water depth was estimated by **multiplying** rainfall intensity values (1-hour duration for 10-year, 20-year, and 50-year return periods) with the accumulated flow surface. (This approach assumes that rainfall volume is proportionally distributed along the flow accumulation surface, enabling estimation of potential water depth.)

Inundation areas are identified where accumulated rainfall depth exceeds terrain elevation. To avoid overestimation caused by local outliers, the **90th percentile (PCT_90)** is used as a threshold to filter extreme values.

3.5 Simulate water spread:

Apply **focal statistics** to smooth the water depth raster, using a combination of radius (3 or 5) and statistical method (Mean or Median) to simulate realistic water dispersion. Then, Areas with water depth **less than 0.1 meters** were removed (considered non-waterlogged) to produce the smoothed flood depth layer.

3.6 Extract connected flood regions and calculate area:

Use **Region Group** to identify connected inundation regions. Convert the result to a vector layer using **Raster to Polygon**, calculate the area of each region, and then filter out regions **smaller than 1000 square meters** (considered non-waterlogged) to obtain the final inundation layer.

3.7 Generate final depth layer and contour lines:

The smoothed depth layer was **extracted** using the final inundation extent, and contour lines at **0.5-meter intervals** were then generated to delineate areas of significant water depth.

3.8 Determine shelter locations:

Use the final flood extent to perform an **Erase** operation on the facilities layer (which includes schools and hospitals) to identify public facilities located in non-inundation areas as candidate shelters.

3.9 Data extraction:

From the result of each analysis condition, extract the following indicators: **total inundation area**, **number of inundation regions**, **average water depth**, and **maximum water depth**.

4. Results and Analysis

The following table shows the extracted data。

Table 1: Summary of Simulated Inundation Characteristics under Different Rainfall Scenarios and Focal Statistics Conditions

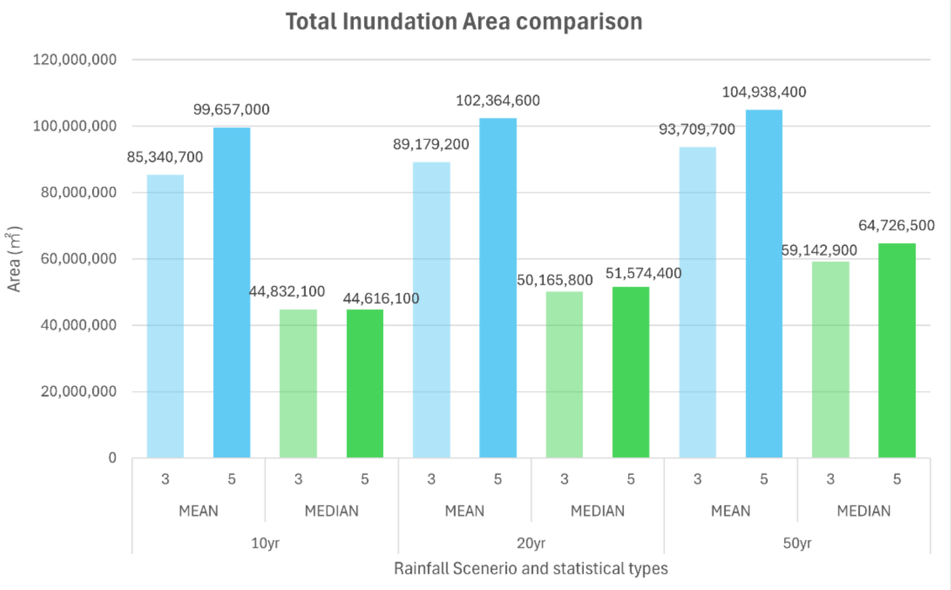
Rainfall Scenario	Radius	Type	Maximum Depth (m)	Total Area (㎡)	Region Count	Average Depth (m)
10yr	3	MEAN	2.14	85,340,700	1,351	1.12
10yr	3	MEDIAN	2.14	44,832,100	3,107	0.61
10yr	5	MEAN	1.99	99,657,000	262	1.03
10yr	5	MEDIAN	2.14	44,616,100	1,145	0.29
20yr	3	MEAN	2.48	89,179,200	1,968	1.16
20yr	3	MEDIAN	2.48	50,165,800	2,859	0.77
20yr	5	MEAN	2.42	102,364,600	668	1.09
20yr	5	MEDIAN	2.48	51,574,400	998	0.39
50yr	3	MEAN	2.96	93,709,700	2,099	1.22
50yr	3	MEDIAN	2.96	59,142,900	2,348	1.09
50yr	5	MEAN	2.88	104,938,400	1,170	1.13
50yr	5	MEDIAN	2.96	64,726,500	703	0.65

Note: Rainfall Scenario = rainfall return period (10-, 20-, 50-year in 1 hour duration); Radius = neighborhood size in cells in Focal Statistics; Type = smoothing method (MEAN or MEDIAN) in Focal Statistics; Maximum Depth = highest depth (m); Total Area = area with depth ⩾ 0.1 m (m^2); Region Count = number of inundation regions; Average Depth = mean depth (m).

Next, I analyze the relationship between different focal statistics approaches and the resulting inundation characteristics.

Analysis 1: Relationship between total inundation area and focal statistics under different rainfall scenarios

Figure 1. Total Inundation Area under Different Focal Statistics Methods



Note: Blue bars represent the **MEAN** method, and green bars represent the **MEDIAN** method. Lighter colors indicate a radius of **3**, while darker colors indicate a radius of **5**.

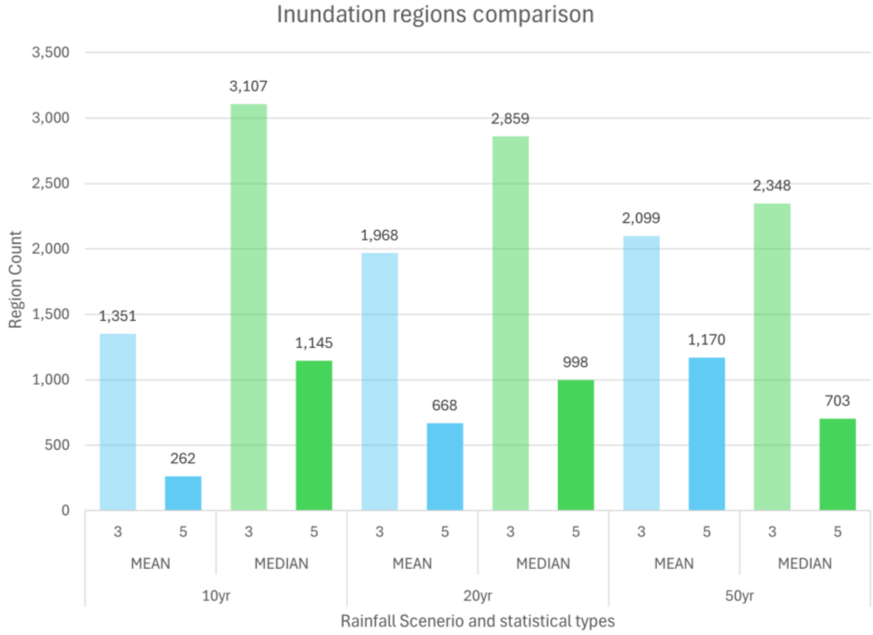
In general, as the rainfall amount increases, the total inundation area also gradually increases. For example, under the MEAN statistical method with radius = 5, the flood area increases from **99,657,000 m²** (10-year) to **104,938,400 m²** (50-year).

Under the same rainfall scenario, the total flood area calculated using the MEAN type is significantly larger than that using MEDIAN. For instance, under the 20-year scenario with radius = 3, the MEAN method results in **89,179,200 m²**, whereas the MEDIAN method yields only **50,165,800 m²**.

When comparing smoothing radius, a radius of 5 tends to produce a slightly larger flood area than radius 3. For example, under the 50-year MEAN condition, radius = 5 gives **104,938,400 m²**, slightly larger than **93,709,700 m²** with radius = 3. However, the difference between MEAN and MEDIAN remains much greater than that between different radii.

Analysis 2: Relationship between number of inundation regions and focal statistics under different rainfall scenarios

Figure 2: Number of Inundation Regions under Different Focal Statistics Methods



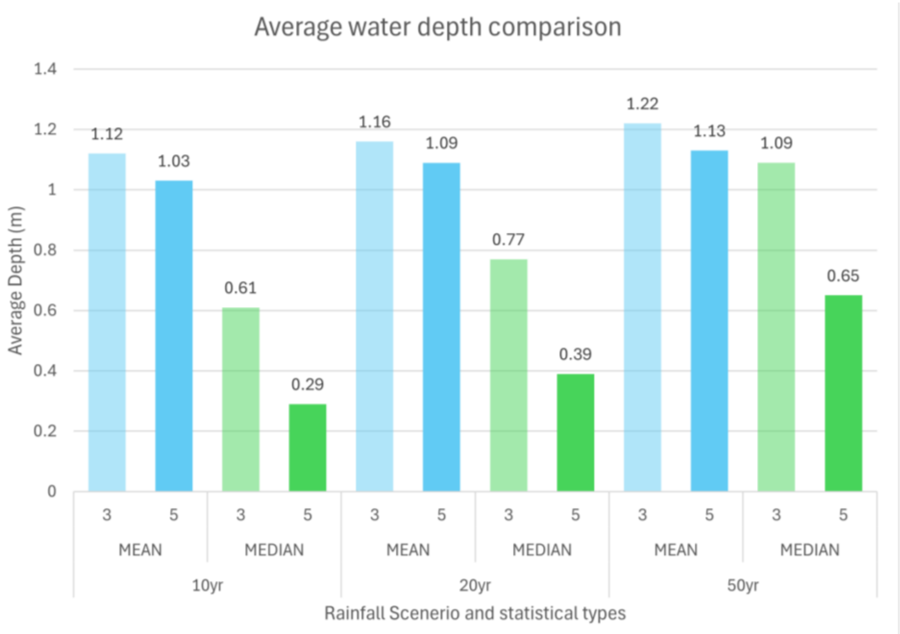
Note: Blue bars represent the **MEAN** method, and green bars represent the **MEDIAN** method. Lighter colors indicate a radius of **3**, while darker colors indicate a radius of **5**.

Overall, the number of inundation regions calculated using the **Mean** method increases as rainfall increases, while the number of regions under the **Median** method decreases. Under the same rainfall scenario, **Median** produces significantly more flood regions than **Mean**. Additionally, increasing the radius leads to a noticeable **reduction in region count** for both methods.

However, under the extreme **50-year rainfall scenario**, an exception emerged: **MEAN smoothing with radius = 5** resulted in more inundation regions (1,170) than **MEDIAN** (703). This counterintuitive outcome suggests that MEAN smoothing may raise marginal pixel values above the 0.1m threshold, preserving small isolated regions. In contrast, MEDIAN may suppress these edge values, leading to their elimination. This highlights the model's high sensitivity to minor variations in flood depth under extreme rainfall conditions.

Analysis 3: Relationship between average flood depth and focal statistics under different rainfall scenarios

Figure 3: Average Water Depth under Different Focal Statistics Methods



Note: Blue bars represent the **MEAN** method, and green bars represent the **MEDIAN** method. Lighter colors indicate a radius of **3**, while darker colors indicate a radius of **5**.

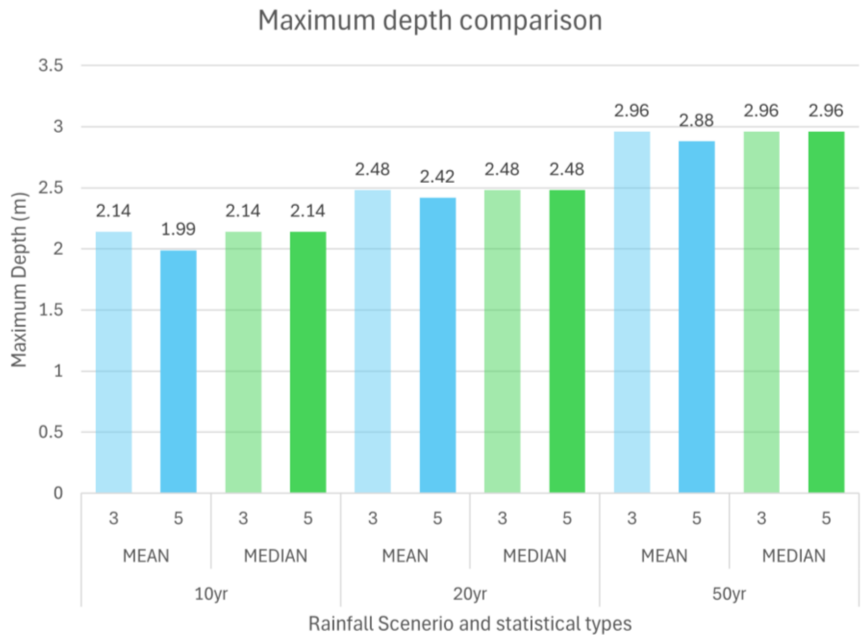
In general, average flood depth increases as rainfall increases, showing a positive correlation. For example, under the MEAN method with radius = 3, the average depth increases from **1.12 m** (10-year) to **1.22 m** (50-year).

Under the same rainfall scenario, the MEDIAN method consistently yields a significantly lower average depth compared to the MEAN method. For instance, in the 20-year scenario with radius = 3, MEAN gives **1.16 m**, while MEDIAN gives only **0.77 m**.

Furthermore, a radius of 5 tends to produce slightly lower average depths than radius = 3. For example, under the 10-year MEAN condition, radius = 3 gives **1.12 m**, while radius = 5 gives **1.03 m**.

Analysis 4: Relationship between maximum flood depth and focal statistics under different rainfall scenarios

Figure 4: Maximum Water Depth under Different Focal Statistics Methods



Note: Blue bars represent the **MEAN** method, and green bars represent the **MEDIAN** method. Lighter colors indicate a radius of **3**, while darker colors indicate a radius of **5**.

Maximum flood depth also increases with increasing rainfall, showing a clear positive correlation. For example, under the MEAN method with radius = 3, the maximum depth rises from 2.14 m (10-year) to 2.96 m (50-year).

Under the same rainfall scenario, when radius = 3, there is almost no difference between MEAN and MEDIAN. For instance, in the 20-year scenario with radius = 3, both MEAN and MEDIAN result in a maximum depth of 2.48 m.

However, when radius = 5, the maximum depth under the MEDIAN method tends to be slightly higher. For example, in the 10-year scenario, MEAN with radius = 5 gives 1.99 m, while MEDIAN gives 2.14 m. A similar trend appears in the 50-year scenario, where MEAN with radius = 5 results in 2.88 m, and MEDIAN reaches 2.96 m.

4.5 Summary

The analysis results show that as rainfall increases, all four indicators—total flood area, number of flood regions, average water depth, and maximum water depth—generally increase, which aligns with expected hydrological patterns.

The choice of statistical method has a significant impact on the results: using MEAN for smoothing tends to produce higher average depths and larger flood areas, while MEDIAN often results in more fragmented but more numerous inundation regions. Changes in radius also have a clear effect on region connectivity—radius = 5 helps reduce the number of regions and improves spatial continuity compared to radius = 3.

Notably, under the extreme 50-year rainfall scenario, an unexpected trend appeared: the number of inundation regions extracted using MEAN + radius = 5 was actually higher than with MEDIAN, indicating the model's high sensitivity to small changes in water depth near threshold boundaries. This also highlights the influence of smoothing methods on threshold behavior.

Considering spatial continuity, clarity of region delineation, and consistency across scenarios, the visualizations in the following sections will prioritize the results generated using Focal Statistics with radius = 5 and MEAN. This approach offers a more accurate and integrated representation of urban flood risk patterns under different rainfall scenarios.

4.6 Visualization of Inundation Maps for Different Rainfall Scenarios (Radius = 5, MEAN)

Figure 5. Inundation areas under **10-year** rainfall return period

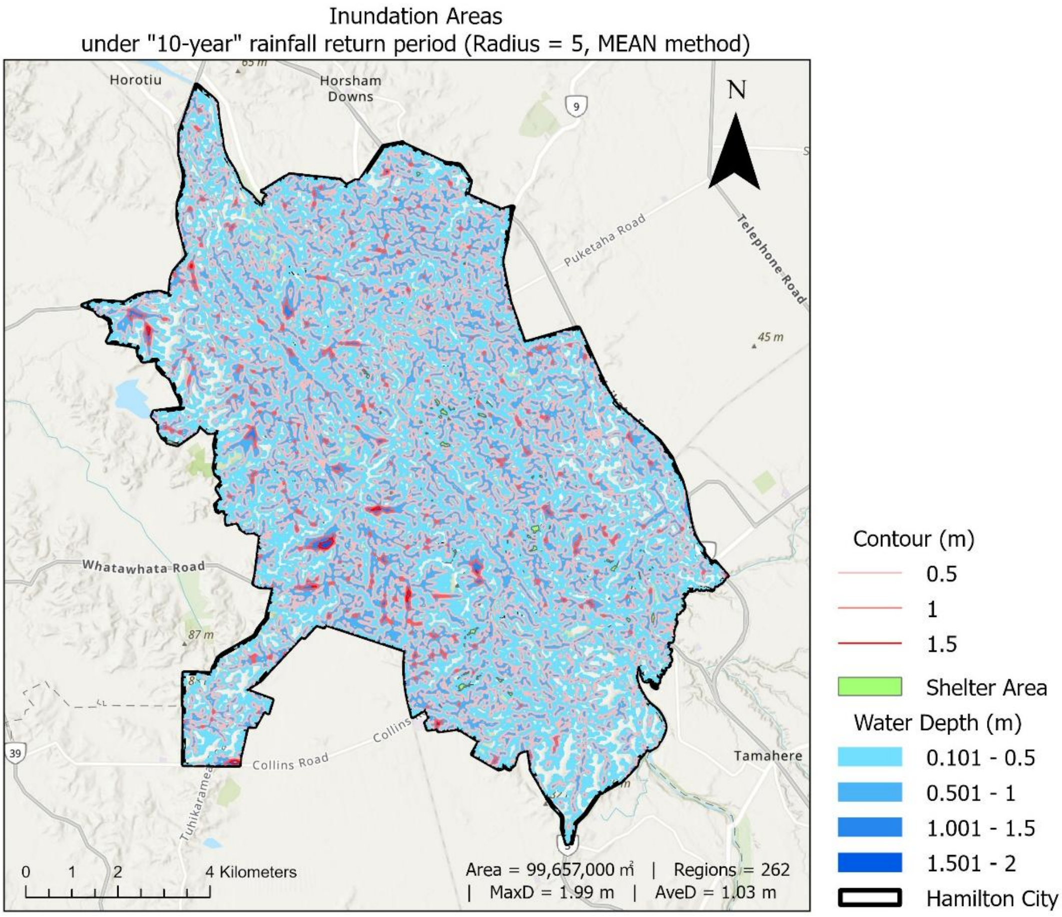


Figure 6. Inundation areas under **20-year** rainfall return period

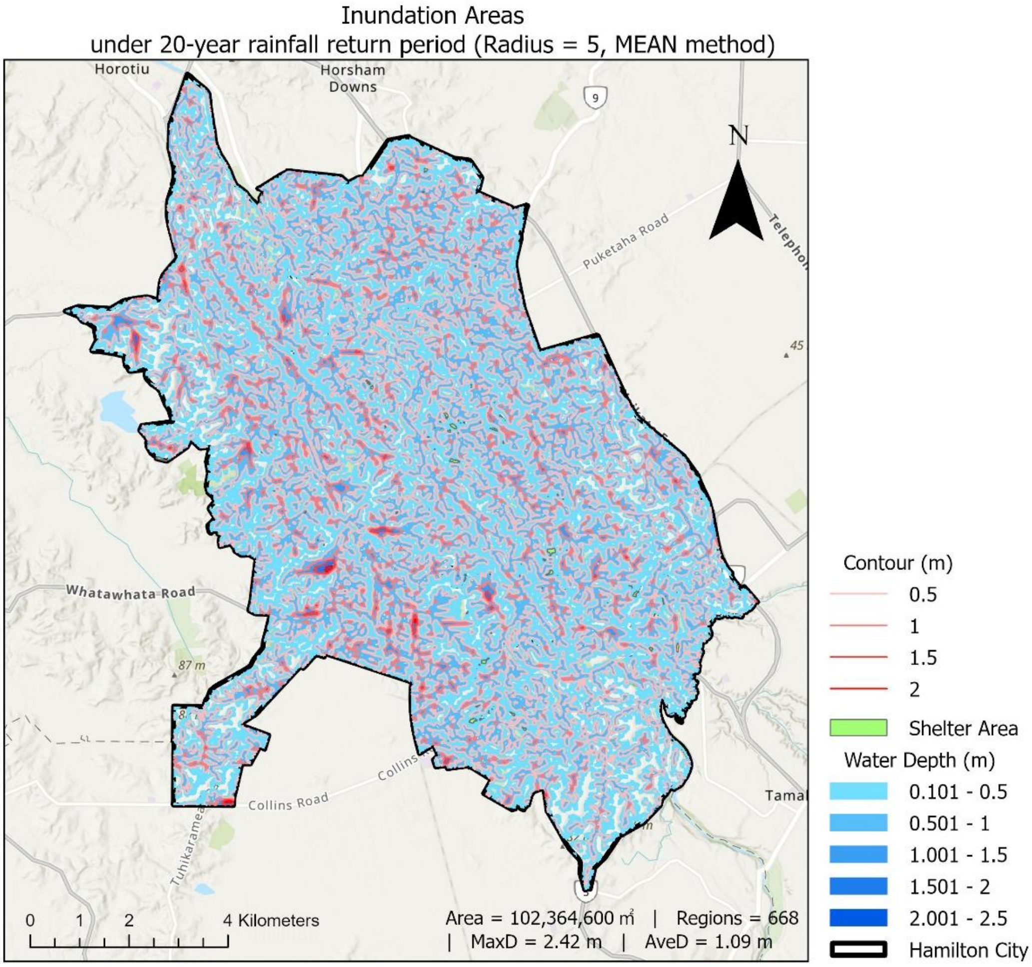
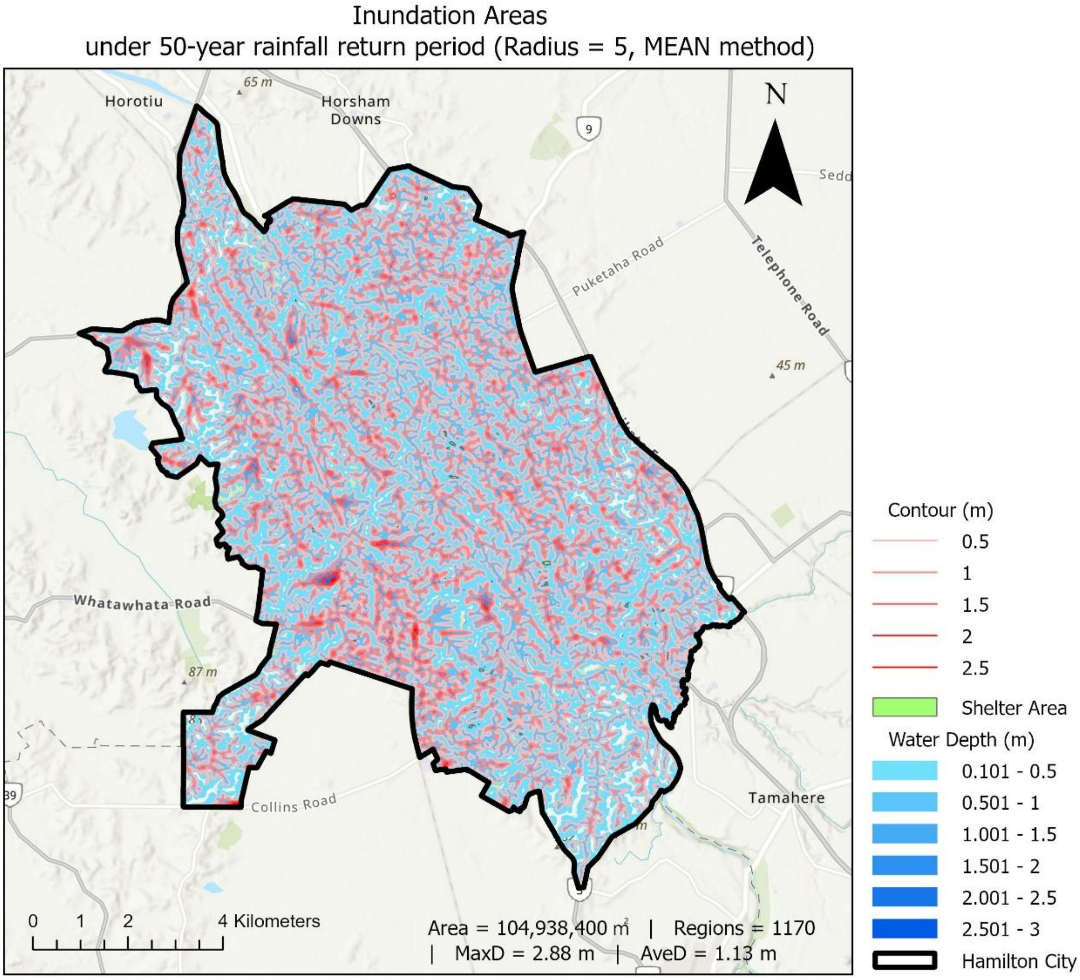


Figure 7. Inundation areas under **50-year** rainfall return period



5. Limitations of the Analytical Methods

This project has several methodological limitations:

1> Weight assignment for soil and landcover layers was based solely on literature references, without performing any sensitivity analysis to evaluate the robustness of the assigned values.

2> Simulation of water flow accumulation did not consider urban drainage systems (e.g., stormwater pipes, culverts, and so on), which may significantly influence real-world flow behavior.

3> Shelter site selection was conducted without incorporating accessibility analysis, such as proximity to roads, population distribution, or walking distances.

6. Future work

Based on the limitations, future research can focus on the following directions to enhance the model's accuracy, robustness, and practical applicability:

1> Improve terrain resolution: Integrate higher-resolution topographic data, such as 1-

meter LiDAR-derived DEM, to achieve more precise surface representation and enable comparative analysis of model outcomes.

2> Incorporate dynamic hydrological processes: Introduce time-varying rainfall inputs to better simulate real-world flood behavior, and conduct model validation using historical flood records to ensure reliability.

3> Conduct sensitivity analysis: Evaluate the influence of weight assignments for land cover and soil layers through sensitivity testing, to refine parameter settings and improve model stability.

4> Enhance emergency planning: Integrate network-based accessibility analysis for shelter site selection, improving the realism and responsiveness of evacuation planning.

Appendix 1: Dataset Sources

Index	Name	Source	URL
1	DEM_Hamilton_Area	LINZ Data Service	https://data.linz.govt.nz/layer/117092-waikato-hamilton-lidar-1m-dem-2023/
2	Rainfall	National Environmental Data Center	https://hirds.niwa.co.nz/
3	Hamilton City Boundary	Stats NZ Geographic Data Service	https://datafinder.stats.govt.nz/layer/120963-territorial-authority-2025/
4	Soil	LRIS（Land Resource Information System）	https://lris.scinfo.org.nz/layer/48066-nzlri-soil/
5	Landcover	LRIS（Land Resource Information System）	https://lris.scinfo.org.nz/layer/104400-lcdb-v50-land-cover-database-version-50-mainland-new-zealand/
6	Facilities	LINZ Data Service	https://data.linz.govt.nz/layer/105588-nz-facilities/

Appendix 2: Weight Assignment

Weight layer	Category	Weight value
Soil layer	Well drained	0.2
	Moderately well drained	0.3
	Imperfectly drained	0.5
	Poorly drained	0.6
	Very poorly drained	1
Landcover layer	Lake or Pond	0.2
	River	0.2
	Indigenous Forest	0.3
	Exotic Forest	0.3
	Deciduous Hardwoods	0.3
	Broadleaved Indigenous Hardwoods	0.3
	High Producing Exotic Grassland	0.5
	Low Producing Grassland	0.5
	Short-rotation Cropland	0.5
	Orchard, Vineyard or Other Perennial Crop	0.5
	Urban Parkland/Open Space	0.6
	Herbaceous Freshwater Vegetation	0.6
	Transport Infrastructure	0.6
	Built-up Area (settlement)	1
	Surface Mine or Dump	1
	Gorse and/or Broom	0.6
	Manuka and/or Kanuka	0.5

Reference:

Ministry for the Environment. (2020). *National Climate Change Risk Assessment for Aotearoa New Zealand: Main report – Arotakenga Tūraru mō te Huringa Āhuarangi o Āotearoa: Pūrongo whakatōpū*. Wellington: Ministry for the Environment. Retrieved from https://environment.govt.nz/assets/Publications/Files/national-climate-change-risk-assessment-main-report.pdf

Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., & Passalacqua, P. (2018). GeoFlood: Large-Scale Flood Inundation Mapping Based on High-Resolution Terrain Analysis. *Water Resources Research*, 54(12), 10013–10033. https://doi.org/10.1029/2018WR023457

Soliman, M., Morsy, M. M., & Radwan, H. G. (2022). Assessment of implementing land use/land cover (LULC) 2020-ESRI global maps in 2D flood modeling application. *Water*, *14*(23), 3963. https://doi.org/10.3390/w14233963

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall–runoff modeling, and doing hydrology backward. *Water Resources Research*, 45(2), W02429. https://doi.org/10.1029/2008WR006912

Cohen, S., Brakenridge, G. R., Kettner, A., Bates, B., Nelson, J., McDonald, R., Huang, Y. F., Munasinghe, D., & Zhang, J. (2018). *Estimating floodwater depths from flood inundation maps and topography*. Journal of the American Water Resources Association, 54(4), 847–858. https://onlinelibrary-wiley-com.ezproxy.waikato.ac.nz/doi/epdf/10.1111/1752-1688.12609

Tran, T. P., & Nguyen, D. L. (2025). *Flood Inundation Assessment on Agricultural Land: Integrating High Spatial Resolution Sentinel Data with LiDAR DEM*. *International Journal of Geoinformatics, 21*(3), 71–81. https://journals.sfu.ca/ijg/index.php/journal/article/view/3997/2113

Kar, B., & Hodgson, M. E. (2008). *A GIS-based model to determine site suitability of emergency evacuation shelters*. *Transactions in GIS, 12*(2), 227–248. https://onlinelibrary-wiley-com.ezproxy.waikato.ac.nz/doi/full/10.1111/j.1467-9671.2008.01097.x