ICFM 10
Adapting to Global Change:
Innovative Approaches to Flood Management and Resilience
May 20-22, 2026 | Western University, London, Ontario, Canada

Estimating Flood Damages in Data-Sparse Regions: Challenges, Uncertainties, and Strategies for Efficient Field Data Collection

Nathan Valsangkar[1] and Ben Humphreys[2]
Northwest Hydraulic Consultants Ltd. (NHC), North Vancouver, BC, Canada

E-mail[1]: NValsangkar@nhcwater.com
E-mail[2]: BHumphreys@nhcwater.com

ABSTRACT

Flood risk assessments are powerful tools for supporting flood management decisions. These assessments often rely on flood damage estimates to quantify potential economic losses, evaluate mitigation strategies, and understand community vulnerability. However, accurate flood damage estimation requires detailed information on flood hazards and the characteristics of exposed assets. In data-sparse regions, practitioners must often rely on simplifying assumptions or undertake field data collection for site-specific analysis. Both have limitations: simplifying assumptions may introduce uncertainty, while data collection in remote communities poses financial and logistical challenges. A critical evaluation of uncertainty sources is needed to guide decisions on balancing study accuracy and data collection costs. This study conducts an uncertainty analysis of flood damage estimates for a First Nation community in Northern BC using the open-source NRCan CanFlood software. The analysis relies on flood damage estimates from a highly detailed building dataset for the community, which serves as a reference case for relatively low uncertainty. To assess the impact of data limitations, we remove layers of information from the detailed dataset, replace them with simplifying assumptions, and quantify the associated change in damage estimates. We also assess, through sub-sampling of the building dataset, how uncertainty may change with differing community sizes from 5 to 100 buildings. Uncertainty ranges are compared to other sources of uncertainty such as hydrology, hydraulics, and the use of depth-damage vulnerability curves. The findings highlight key uncertainty sources affecting flood damage estimates and inform efficient field data collection strategies to balance cost against acceptable levels of study uncertainty. These insights can support more reliable flood risk assessments in data-sparse regions.

KEYWORDS: Flood, risk, CanFlood, uncertainty, sensitivity, vulnerability, Nass, Nisga'a

1 INTRODUCTION

The shift toward risk-based flood management has elevated the importance of quantitative flood risk models. These models integrate four key components to estimate flood risk, whether at the event level (e.g., 200-year flood) or annualized: flood hazard (primarily inundation depth), exposure (e.g., asset inventories), asset value, and vulnerability (e.g., depth–damage relationships). Each component introduces uncertainty, which propagates through the analysis and accumulates in the final damage estimates (de Moel & Aerts, 2011). Understanding and managing these uncertainties is critical for producing reliable risk assessments and supporting decision making.

While flood risk uncertainties have been examined in prior studies (e.g., Ward et al., 2011), a bibliometric review by Diez-Herrero and Garrote (2020) found that only one-sixth of published flood risk assessments examine uncertainty as a key topic. Existing research largely focuses on urban or regional scales with high-quality data. In British Columbia (BC), recent policy shifts aim to extend flood risk assessments to remote and Indigenous communities (BC, 2024), where data scarcity is common. These

assessments often depend on assumptions or limited datasets, which can significantly affect study outcomes (Rrokaj et al., 2025). Collecting additional data may be costly or infeasible. Identifying dominant uncertainty sources in small-scale, data-limited contexts is essential to guide efficient data collection strategies and use of reasonable assumptions.

The purpose of this study is to determine whether the detail of a building asset inventory can be reduced to significantly lower data collection requirements for future studies. The question arose after a detailed building asset inventory was developed for a flood risk assessment of Nisg̲a'a villages in northwest BC. The objectives of this study are to determine whether simplified building assumptions are sufficient for the calculation of annualized flood risk, whether these yield results that are within acceptable uncertainty bounds, and how this uncertainty compares to other sources of uncertainty inherent in the risk assessment.

2 STUDY AREA AND DATA SOURCES

The study area consists of the villages of Lax̲galts'ap and Gitwinksihlkw in the Nass River Valley, in Northwestern BC, Canada. Figure 1 shows the village locations, as well as the distribution of buildings and floodplain extents. Gitwinksihlkw is subject to riverine flooding from the Nass River, has a population of 207 people (as of the 2021 census), and contains 70 buildings. Lax̲galts'ap is subject to riverine flooding from both the Nass River and Greenville Creek, has a population of 248, and contains 186 buildings. The buildings in both villages are primarily residential, with some being commercial and institutional (of both the Nisg̲a'a Lisims Government and Village governments).

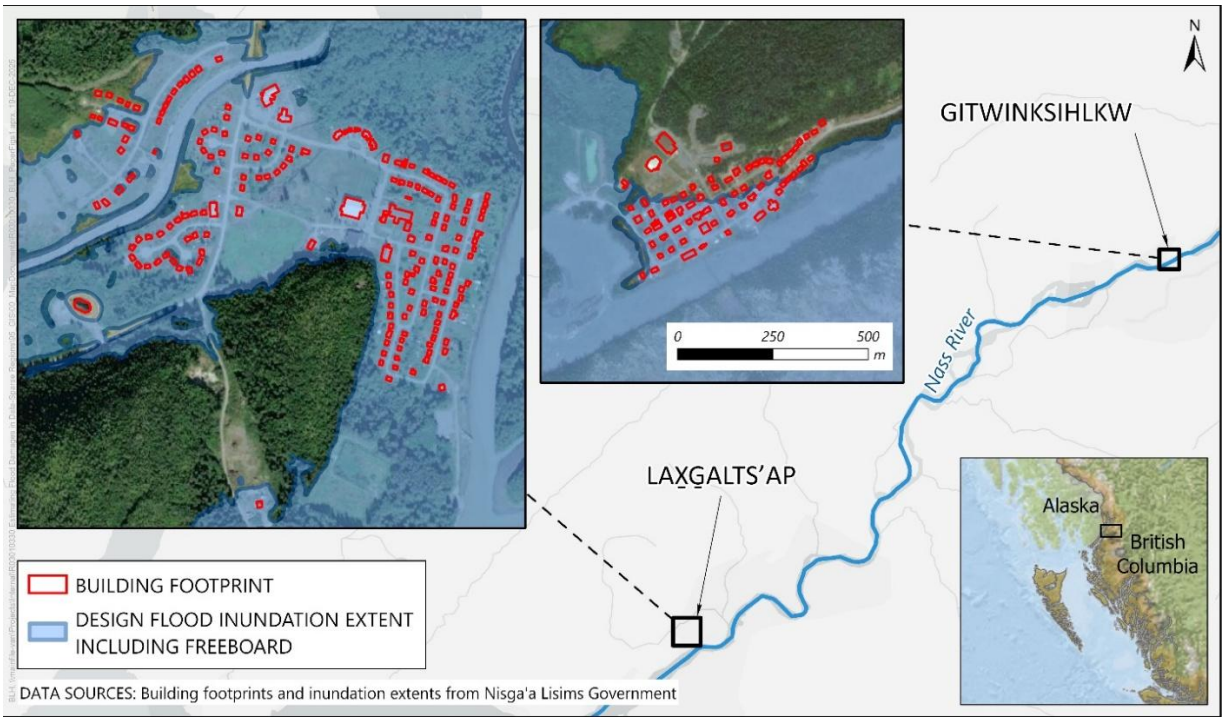

Figure 1: Location of Nisg̲a'a Villages within study area.

2.1 Flood Hazards

This study focuses on flood hazards and risks primarily associated with the Nass River. The Nass River originates in the Skeena Mountains and has the third largest watershed entirely contained in British Columbia, draining an area of roughly 20,500 km^2 (Cascadia, 1997). Peak flood flows typically occur in June due to spring freshet; however, the largest floods have been generated by extreme fall precipitation. Detailed flood hazard data for Gitwinksihlkw and Lax̲galts'ap was developed from local hydrologic and floodplain mapping studies completed for the Nisg̲a'a Lisims Government (NHC, 2022). Those studies

used gauge flood frequency analysis (93-year record period, spanning back to 1917) to estimate Nass River discharges at the two villages for return periods up to 500 years. Greenhouse gas emission scenarios from CMIP6, SSP5-8.5 were also modelled to assess increased flood discharges for the mid- (2041-2070) and end-of-century (2071-2100) climate periods. Flood discharges were input to a calibrated two-dimensional numerical model (HEC-RAS 6.1) to simulate floodplain hydraulics. Key outputs utilized included floodplain limits and raster files of flood elevation, depth, and velocity. For the present study, we used hydraulic model outputs from the end-of-century climate period as those produced the greatest inundation extents and a larger sample size of buildings for uncertainty analysis. Return periods evaluated included the 5-, 10-, 25-, 50-, 100-, 200-, and 500-year events.

2.2 Building Inventory

The detailed building asset inventory dataset consisted of attributed building footprint polygons in a GIS format. The polygon boundaries were derived from full-feature lidar data, with manual modifications based on imagery and photogrammetry-derived point cloud data where applicable (see description below, and overview of the point cloud for the Village of Gitwinksihlkw in Figure 2). Manual modifications included the removal of outbuildings (e.g., sheds), and the addition and removal of buildings that were not correctly represented or did not exist at the time of the lidar data collection.

Figure 2: Overview of photogrammetry-derived point cloud for the Village of Gitwinksihlkw.

Open-source 3D imagery such as Google Street View is often used for building attribution (Ho et al., 2024) but is unavailable for large portions of this remote study area. To address this, we used a combination of drone- and mobile ground camera-based imagery to develop a detailed 3D model of the two villages, georeferenced using lidar tie points. The drone imagery provides a good overview of the study area and coverage of the building roofs, streets, and some building faces, but not significant detail. The ground-based imagery provides greater detail of the building faces, but lacks overall areal context and makes 3D construction difficult. Integrating the datasets mitigated these limitations and produced a complete 3D model.

The building footprints were attributed with data from a variety of open and closed data sources, including BC Assessment data, lidar, and the photogrammetry-derived point cloud. Main attributes include building type (e.g., single family home, mobile home, commercial), floor area, foundation type (e.g., crawlspace, basement, slab), first floor elevation, and adjacent ground elevation. The process of collecting data and attributing the buildings to this level of detail requires a significant amount of time and effort. The opportunity to use simplifying assumptions would greatly reduce the level of effort required for this stage of flood risk assessment projects.

3 METHODOLOGY

The uncertainty analysis focused on tangible direct damages to building structures and contents. A baseline risk assessment used the full building inventory and hazard datasets, calculating both event-level and annualized risk. Next, sensitivity runs explored the impact of simplified building inventory assumptions on the risk results. Finally, the influence of sample size was explored by resampling the baseline and sensitivity scenario results to determine the threshold at which a lesser number of buildings cause simplifying assumptions to introduce large magnitudes of uncertainty.

3.1 Using CanFlood for Risk Analysis

Natural Resources Canada's CanFlood V1.2.2 flood risk modelling toolbox was utilized for the building risk modelling (NRCan, 2024). CanFlood is an open source QGIS plugin that contains tools to automate modelling tasks and utilizes a standardized data format to enhance reproducibility and future expansion of the assessment. Asset data inventories and water surface elevation rasters are input to determine the exposure and risk of the structures in the study area.

Building depth-damage vulnerability functions were taken from previous flood risk studies in BC's Lower Mainland (IBI Group, 2020) and Okanagan Valley (Arcadis IBI Group and NHC, 2023). The curves represent the replacement value of flood damages, rather than the depreciated value, and include minimum, maximum, and "typical" damage values based on flood depth. The base curves were scaled up by 25% to account for inflation and regional differences in construction costs (NRCan, 2021). Despite being derived from other regions, the curves provide a reasonable approximation of potential flood damages considering the existing building stock in the Nass Valley. Previous work by NHC (unpublished, confidential data) compared these vulnerability curves to actual flood damages in another BC Indigenous community, and found that the minimum and maximum curves generally enveloped the reported flood damages across eight homes. Vulnerability functions were applied to each asset to estimate losses associated with building structures and contents. Structural losses include repairs to basements, common areas, garages, and main floors for residential and non-residential buildings. Content losses cover household items and equipment or stock for non-residential properties. Contents were not verified in the field and rely on curve-based estimates. Both structural and content damage estimates are highly sensitive to small changes in water surface elevation, as illustrated in Figure 3.

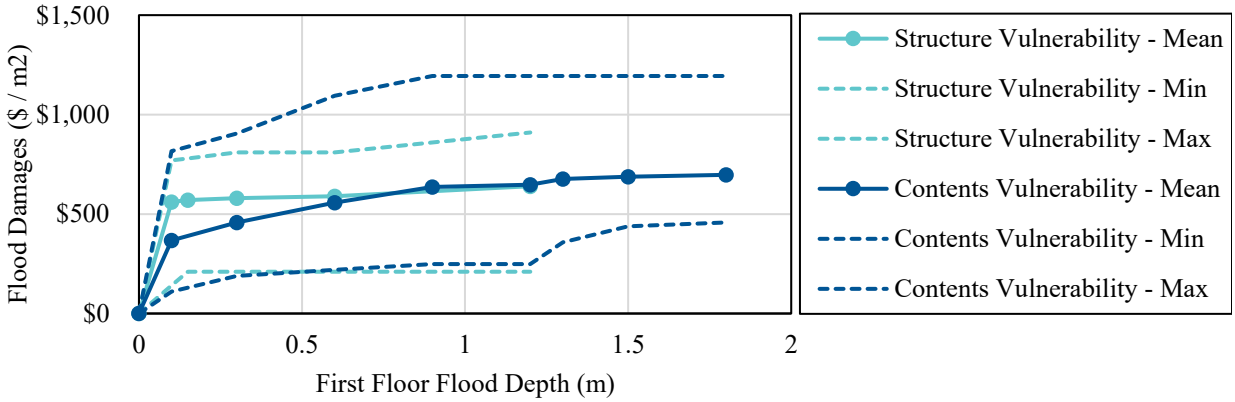

Figure 3: Example of structural and contents damage curves for single-family residential buildings

3.2 Baseline Risk Assessment

The baseline risk assessment utilized the fully detailed building asset inventory (Section 2.2), vulnerability curves (Section 3.1), and the set of seven flood hazard return periods (Section 2.1). Damages were calculated for each return period and integrated over the damage–frequency distribution to estimate expected annualized damage (EAD), assuming zero damage for the 2-year event and a maximum of 120% of the 500-year damages for the upper tail.

3.3 Sensitivity Scenarios

Eight sensitivity scenarios evaluated how simplified assumptions regarding the building inventory influence flood risk estimates and associated uncertainty. The analysis emphasized quantifying the magnitude of uncertainty introduced by these assumptions in comparison to other contributing sources of uncertainty.

Scenario 1. Building type (e.g., single family, commercial, institutional) was set from zoning data, and structures that visually appeared to be outbuildings (e.g., sheds) were removed from the inventory. Building main floor elevations were set from the following assumptions: all single-family dwellings have basements, and the first floor is 0.6 m above adjacent ground; all mobile homes are 0.6 m above adjacent ground; all other buildings are slab-on-grade, with the first floor 0.2 m above adjacent ground. This represents a scenario where the building type and footprint are known, but where the main floor elevation and foundation status are unknown. The 0.6 m assumption is based on engineering judgement and typical stair heights to the front door.

Scenario 2. All buildings are set as single family dwellings with basements, and the first floor is 0.6 m above adjacent ground. This represents a further simplification of Scenario 1, where the building type or zoning information are unknown.

Scenario 3. Building main floor elevations were set as per Scenario 1, but foundation conditions (basement vs. crawlspace vs. slab) and building type (main building vs. outbuilding) were updated based on BC Assessment data, where available. This represents a scenario where building information is available for the majority of buildings, other than the main floor elevation.

Scenario 4. Hydrologic uncertainty was assessed by recalculating baseline risk using the 90% confidence limits from the Nass River flood frequency analysis (NHC, 2022). To minimize additional hydraulic runs, scenarios adjusted the return periods for given discharges when computing EAD. For example, the 100-year discharge at Gitwinksihlkw has a 90% confidence range of 44-233 years. This represents a lower estimate of the hydrologic uncertainty, which is further influenced by gauge measurement error, choice of statistical distribution, and climate change effects.

Scenario 5. Hydraulic uncertainty was assessed by uniformly adjusting flood depths in CanFlood by ±0.3 m, based on hydraulic model calibration errors (0.1-0.4 m) identified for the June 2021 freshet (NHC, 2022). This approach does not capture the spatial variability in hydraulic uncertainty and may underestimate damage, as higher depths could expand flood extents and the number of impacted properties beyond the baseline scenario.

Scenario 6. Uncertainty in the building vulnerability curves was evaluated by applying the minimum and maximum curve values contained in the CanFlood inventory, rather than the mean value used for the baseline risk scenario. This is a simplified approach to evaluating uncertainty. A more fulsome assessment would require the development of site-specific curves, or a comparison of estimated damages to actual damages from recent floods. However, local data to support either approach is not available.

Scenario 7. Uncertainty in the EAD calculation was first assessed by reducing the number of computed return periods from seven (5, 10, 25, 50, 100, 200, 500 year) to four (10, 50, 100, 500 year). Extrapolation of the risk curve tails was done under the same assumptions as the baseline assessment.

Scenario 8. Uncertainty in the EAD calculation was further assessed by modifying assumptions around the upper risk curve tail. Rather than setting the tail value to 120% of the 500-year flood damages, additional tests were conducted at 150% and 200%.

3.4 Effect of Building Sample Size

For sensitivity Scenarios 1-3, applying simplified assumptions to a small set of buildings could introduce proportionally greater uncertainty than when applied to a larger dataset. This effect is most pronounced for first-floor elevation, which strongly influences flood depth and resulting damage estimates. In smaller samples, elevation assumptions are more likely to skew results, whereas with larger samples the effect could be smoothed out over the building inventory. To evaluate the influence of sample size, 1,000 randomized building subsets were generated, with replacement, for sample sizes of 5, 10, 20, 50, and 100

buildings. For each subset, percentage differences in damage estimates were calculated relative to the baseline scenario for sensitivity Scenarios 1-3. Percentile statistics were then computed to characterize the distribution of errors.

4 RESULTS AND DISCUSSION

In the baseline risk scenario, for the 500-year return period, there were 93 buildings exposed to flooding at Laxgalts'ap, and 38 at Gitwinksihlkw. Integration over the risk curve yielded an EAD of $551,197. Table 1 compares the sensitivity scenarios' EAD to the baseline risk scenario, in absolute and percentage terms. Figure 4 illustrates the event-level risk curves for the baseline scenario and sensitivity Scenarios 1-3.

Simplifying assumptions in Scenarios 1-3 produced errors of about ±20% in total EAD. Scenario 1 underestimated damages due to assumed floor elevations (0.6 m above grade) being, on average, higher than surveyed values (Figure 4). Scenario 2 overestimated damages because all buildings were treated as single-family homes, which have greater vulnerability than mobile or commercial structures. Scenario 3 improved accuracy in Laxgalts'ap by using BC Assessment data for foundation types; however, it introduced large errors in Gitwinksihlkw, where BC Assessment data contained errors such as missing buildings, or misclassification of main structures as outbuildings.

Table 1: Comparison of sensitivity scenarios EAD to baseline risk

Scenario	Flood Damages EAD			Difference from Baseline Risk		
	Laxgalts'ap	Gitwinksihlkw	Total	Laxgalts'ap	Gitwinksihlkw	Total
Baseline	$475,633	$75,565	$551,197	-	-	-
1	$410,375	$65,135	$475,510	-14%	-14%	-14%
2	$559,466	$89,929	$649,396	18%	19%	18%
3	$445,849	$14,243	$460,092	-6%	-81%	-17%
4 (lower)	$287,521	$36,030	$323,551	-40%	-52%	-41%
4 (upper)	$801,761	$170,281	$972,042	69%	125%	76%
5 (lower)	$342,594	$48,983	$391,577	-28%	-35%	-29%
5 (upper)	$629,855	$100,352	$730,207	32%	33%	32%
6 (lower)	$190,914	$34,987	$225,900	-60%	-54%	-59%
6 (upper)	$768,254	$123,903	$892,156	62%	64%	62%
7	$637,112	$84,835	$721,947	34%	12%	31%
8 (150%)	$481,779	$78,136	$559,915	1%	3%	2%
8 (200%)	$492,023	$82,422	$574,445	3%	9%	4%

Uncertainty from hydrologic (Scenario 4), hydraulic (Scenario 5), and vulnerability curve (Scenario 6) inputs was substantially greater than that from the simplifying assumptions of Scenarios 1-3. Hydrologic uncertainty was particularly influential due to the wide confidence intervals for high-return-period floods extrapolated from the 93-year gauge record. This finding is likely applicable to other remote locations in BC where gauge records are often short or non-existent. Vulnerability curve uncertainty was approximately twice that of hydraulic uncertainty and comparable to hydrologic uncertainty. Developing site-specific vulnerability curves could reduce this error source but would not significantly reduce overall EAD uncertainty because of the underlying hydrology and hydraulics.

A reduction in the number of modelled flood return periods from 7 to 4 (Scenario 7) increased the overall EAD by 31% compared to the baseline risk. This result is similar to previous work by Ward et al (2011) on the Meuse River, Netherlands, where the use of only three return periods overestimated EAD by 33% to 100%. Guidance from Messner et al (2007) based on a literature review recommends using at least three and preferably six flood return periods. Because EAD is computed over discrete return periods, even

the use of seven return periods could result in overestimation of EAD; however, it is not possible to quantify that error without completing additional hydraulic model simulations.

Increasing the maximum damage for the tail of risk curve (Scenario 8) somewhat increased the EAD, though the change was substantially less than for other sources of uncertainty. The true uncertainty in tail assumptions could only be assessed by modelling additional flood scenarios beyond the 500-year event. In BC, provincial guidelines do recommend including additional return periods (1,000-year, 2,500-year) in locations of very high flood risk (EGBC, 2018). However, at those return periods, the degree of hydrologic uncertainty would be substantial and may not provide significant gains in EAD understanding, though there could be other gains in understanding event-based damages for specific discharges.

Overall, the sensitivity analysis indicates that simplifying assumptions introduce uncertainty, but their effect is less than uncertainties from the hydrologic, hydraulic, and vulnerability curve inputs, provided that EAD is integrated over an adequate number of return periods. Scenario 1 is particularly straightforward to apply if building types and ground elevations are known. However, this finding is specific to the local building stock and sample size and may not hold true in all cases.

An analysis of sample size effects (Figure 5) shows that uncertainty from Scenarios 1-3 is highly sensitive to the number of buildings. Small subsamples (n = 5, 10) produce uncertainty ranges comparable to or exceeding hydrologic, hydraulic, and vulnerability curve uncertainty. At intermediate sizes (n = 20, 50), uncertainty decreases substantially, and for large samples (n = 100) it becomes minimal. Scenarios 1 and 2 consistently introduce negative and positive bias, respectively, particularly for $n \geq 20$. These biases reflect local building stock and foundation height assumptions (see Section 3.3). In practice, the height assumptions could be "tuned" to reduce bias by reviewing photos of the typical building stock.

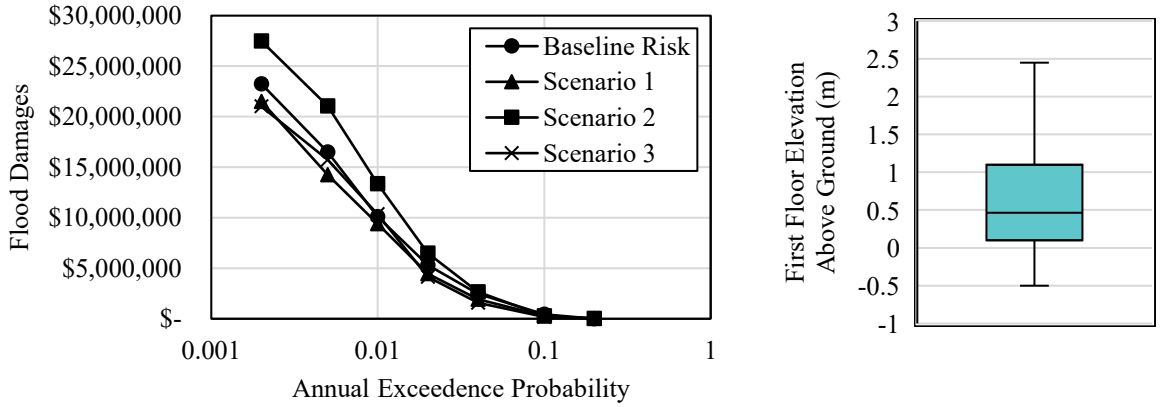

Figure 4: Left panel: Comparison of total flood risk curves for baseline and Scenarios 1-3. Right panel: First-floor elevation differences relative to adjacent ground for the baseline building inventory.

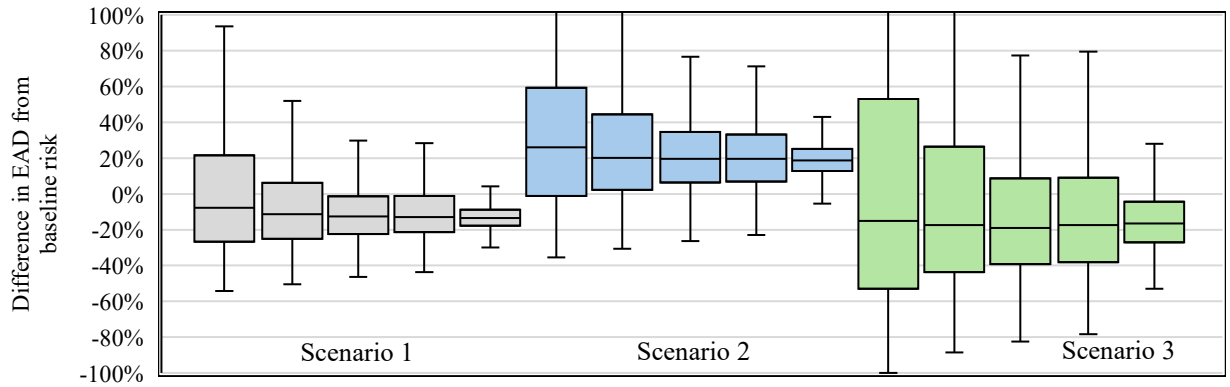

Figure 5: Effect of sample size on distribution of EAD differences between baseline and Scenarios 1-3

5 CONCLUSIONS

Flood damage estimates are inherently uncertain, with major influence from hydrologic and hydraulic variability and uncertainty, particularly in data-sparse regions where gauge records are short or non-existent, and model calibration is limited. Our findings indicate that while simplifying building inventory assumptions introduces error, its impact is generally smaller than uncertainties from hydrologic, hydraulic, and vulnerability curve inputs when annualized damages are computed over adequate numbers of return periods. For larger inventories (~50 buildings or more), generalized assumptions can provide acceptable accuracy and significantly reduce data collection effort. However, detailed building inventories remain critical when estimating structure-specific damages, designing lot-level mitigation measures, or addressing atypical building characteristics. The findings of this study support prioritizing hydrologic and hydraulic data quality while applying simplified asset assumptions judiciously to balance cost and reliability in remote or resource-limited contexts.

6 ACKNOWLEDGEMENTS

We gratefully acknowledge the Nisga'a Lisims Government, and in particular, David Cassidy, for agreeing to their data being used in this study. We would also like to recognize input of Julie Van de Valk, Robin Bourke, and Karl Chastko of Public Safety Canada for their valuable collaboration and contributions to the methodology and results analysis of this project. Finally, we thank Piotr Kuras of Northwest Hydraulic Consultants Ltd. for his review and input on this paper.

REFERENCES

Province of British Columbia (2024). From flood risk to resilience: a B.C. flood strategy to 2035, pp. 1-44.

Cascadia Natural Resources (1997). Water Survey of Canada stations pertaining to the Nass habitat capability model. Prepared for BC Ministry of Environment, Land, and Parks, pp. 1-23.

Diez-Herrero A. and Garrote J. (2020). Flood risk analysis and assessment, applications, and uncertainties: a bibliometric review. Water, 12(7), 2050.

Engineers and Geoscientists BC (2018). Legislated flood assessments in a changing climate in BC, version 2.1, pp. 1-175.

Ho et al. (2024). ELEV-VISION: automated lowest floor elevation estimation from segmenting street view images. ACM J. Comput. Sustain. Soc., 2(2), pp. 1-18.

IBI Group (2020). Flood risk assessment for BC's Lower Mainland. Prepared for Fraser Basin Council.

Arcadis IBI Group and NHC (2023). Flood risk mapping for the Okanagan Valley Watershed. Prepared for the Okanagan Basin Water Board Flood Risk Mapping for the Okanagan Valley Watershed – Okanagan Basin Water Board.

Messner F. et al (2007). Evaluating flood damages: guidance and recommendations on principles and methods. FLOODsite Report Number T09-06-01.

de Moel H. and Aerts, J.C.J.H. (2011). Effect of uncertainty in land use, damage models, and inundation depth on flood damage estimates. Nat. Hazards Earth Syst. Sci, 58, pp. 407-425.

Natural Resources Canada (2021). Federal flood damage estimation guidelines for buildings and infrastructure, Version 1.0.

Natural Resources Canada (2024). Documentation for CanFlood user manual.

NHC (2022). Adaptation of the Nisga'a Nation to the impacts of climate change: hydraulic modelling. Prepared for Nisga'a Lisims Government.

Rrokaj S. et al (2025). Flood risk assessment and participative process in the data-scarce Metuge district of Mozambique: an exportable approach. Int. J. Disaster Risk Reduct., 116, 105163

Ward P.J., de Moel, H. and Aerts, J.C.J.H. (2011). How are flood risk estimates affected by the choice of return periods? Nat. Hazards Earth Syst. Sci., 11, pp. 3181-3195