# FLOOD HAZARD MAPPING AI/ML POC DATASET

Alex Hoover, EIT, created this dataset on August 7th, 2025, for use as a proof-of-concept dataset only. All input and output data are calculated from a small subsection of the Cambie-Heather (CAH) InfoWorks ICM model using Ruby scripting. The code repository is located here: <https://tfs.vancouver.ca/COV/EngIntegratedSewerDrainagePlanning/_git/ICMScripts>. Please request access from ISDP staff if not accessible.

*Figure 1: POC Model*

A map of a city

AI-generated content may be incorrect.

*Figure 2: POC Model Location*

A map with a red circle

AI-generated content may be incorrect.

### Dataset Files

**Input 1 – Rainfall**: Contains 135 different design storm CSV files which represent the rainfall time-series input into the InfoWorks ICM model. Each rainfall input will correspond to an HGL output 1:1. The CSV files are in a specific ICM format; the relevant rainfall information is shown below. Units are in **Intensity (mm/hr).**

*Figure 3: Design Storm CSV Format*

A white grid with black text

AI-generated content may be incorrect.

**Input 2 – Catchment Characteristics:** Contains catchment characteristics for each node in the model. In the POC model, there are 185 nodes; therefore, there are 185 rows of data in the provided CSV file. Only one CSV file is provided as **the catchment characteristics are constant for all model runs in this dataset.**

Explanation of parameters:

* Depth\_to\_invert: Distance from ground elevation to pipe invert. Calculated as flood\_level – chamber\_floor.
* Upstream\_pipe\_count: Number of pipes that are upstream of the manhole.
* Upstream\_pipe\_total\_length: Total summed length of pipes that are upstream of the manhole.
* Upstream\_pipe\_weighted\_avg\_diameter:
* Upstream\_pipe\_weighted\_avg\_gradient:
* Upstream\_storm\_impervious\_area: Total summed area of impervious catchment area for upstream storm catchments.
* Upstream\_storm\_pervious\_area: Total summed area of pervious catchment area for upstream storm catchments.
* Upstream\_storm\_total\_area: Total summed catchment area for upstream storm catchments. Should be equal to upstream\_storm\_impervious\_area + upstream\_storm\_pervious\_area.
* Upstream\_sanitary\_population: Total summed population attributed to the upstream sanitary catchments.
* Stream order: Summed count of headwater nodes upstream of the manhole. Used to describe the shape of the catchment.

*Figure 4: Stream Order Explanation*

A diagram of a stream order

AI-generated content may be incorrect.

Units of parameters:

* Depth\_to\_invert: meters
* Upstream\_pipe\_count: count
* Upstream\_pipe\_total\_length: meters
* Upstream\_pipe\_weighted\_avg\_diameter: millimeters
* Upstream\_pipe\_weighted\_avg\_gradient: m/m
* Upstream\_storm\_impervious\_area: hectares
* Upstream\_storm\_pervious\_area: hectares
* Upstream\_storm\_total\_area: hectares
* Upstream\_sanitary\_population: number of people
* Stream order: count

**Output 1 – Flood Depths:** Contains maximum flood depth relative to the ground level of the manhole chamber. This is technically not HGL, rather HGL – Ground Elevation at Manhole, as this should be easier to predict. Negative values mean the water level is below ground and positive values mean the water level is above ground level (surface flooding). There should be 184 rows of data in each file (not 185), as there are no results for the outfall node. **Units are in meters.**

Each file follows a specific naming scheme: “sim\_{SIM NAME}\_results.csv. InfoWorks defines the SIM NAME as “{NETWORK SCENARIO NAME} {RAINFALL NAME}.” Since the network scenario is constant throughout (POC\_Area\_v1), one can match the input rainfall files with the output water level files easily. For example, output file “sim\_POC\_Area\_v1 COV2018\_2050High\_2y1h\_Z1N\_AES\_results.csv” corresponds to “COV2018\_2050High\_2y1h\_Z1N\_AES.csv” rainfall.