

# AMALGAMATION (ETL) OF GOVERNMENT DATASETS TO SUPPORT FLOOD RISK ASSESSMENT

Heather McGrath

Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada

## ABSTRACT

In this project, a new ArcGIS Toolbox comprised of models built using default ESRI tools and custom python scripts has been created which automates the process of amalgamating datasets from different levels of government into a single dataset which can be imported into Hazus (an extension for ArcGIS Desktop) for flood risk estimation. This Toolbox is designed to provide speed, accuracy, and repeatability of iterative tasks such as querying, reclassification, and transference of values via extract transform, and load (ETL) to fit into the Hazus data schema. Designed and tested on data in the Fredericton area, this tool can be easily adapted for use by other communities to facilitate the creation of Hazus acceptable feature classes.

### OBJECTIVE:

Develop simplified methodology to aggregate datasets from multi-levels of government into a single dataset, while querying and reclassifying to meet the strict data schema standards of Hazus-MH Flood model.

### METHODS:

Combine existing ESRI tools with custom python scripts to automate the process of combining datasets, querying and reclassification of data.

### RESULTS:

Specialized toolbox for importing New Brunswick parcel data along with local level building footprints into format acceptable for use in Hazus-MH

### CONCLUSIONS:

A toolbox which facilitates automated amalgamation and reclassification of data. As government datasets are updated, this toolbox supports facilitation of updating and reimporting data into Hazus-MH.

## BACKGROUND

Flooding is the most expensive natural disaster in Canada. Providing tools which allow government officials, GIS specialists, emergency managers, and first responders tools capabilities to visualize risk across a community is a key requirement. Hazus-MH has been selected as a tool for use, though the process to get local data into Hazus is error prone and time consuming.



Location of Fredericton, NB in eastern Canada



Images from 2008 flood in Fredericton, NB. The second worst flood of the Saint John River in record. Water levels reached 1.86m above flood stage. Across the province, over \$23 million in damages, with approximately \$6 million in damages in Fredericton. Flooding was attributed to melting of record snow-fall from 2007-2008 winter and heavy rainfall.



Image Source: Miao, et al., 2008, Early Warning and Mapping for Flood Disasters, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B4, Beijing 2008

## CONCLUSIONS

To facilitate extract, transform, and load (ETL), a custom toolbox comprised of existing ArcGIS tools and custom python scripts are run on datasets from multi-levels of government to create a singular dataset which is used in the assessment of risk due to natural disasters (specifically floods).

Designed for City of Fredericton, the model may be easily adapted for use in other municipalities in New Brunswick or across the country, or extended for use in risk estimation of other categories of natural disasters. This toolbox helps facilitate the quick amalgamation and reclassification of data to support flood risk estimation studies through ETL functions. This automation reduces the time and manual supervision of repetitive tasks.

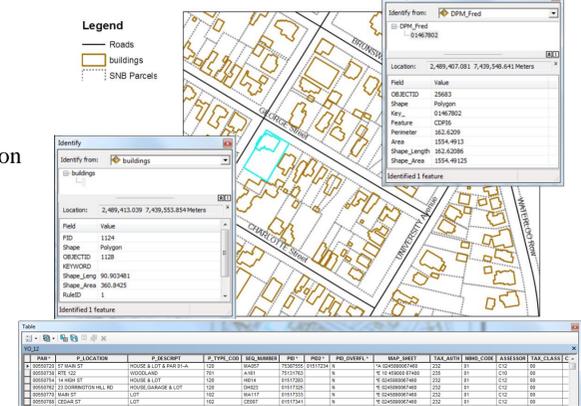
## MOTIVATION

Given the propensity for flooding in NB, providing government officials, GIS specialists, emergency managers, and first responders tools for risk analysis and capabilities to visualize risk across a community is a key requirement.

Create singular dataset which can be used for community level risk assessments, and potential economic and societal losses due to flood events using GIS.

- Combining available open data: Service New Brunswick, the City of Fredericton and the Government of Canada
- To a format which may be imported to Hazus-MultiHazard (Hazus) for flood loss estimation
- Amalgamation and conversion include:
  - Extract, transform, and load (ETL)
  - Merging, queries, reclassification and transferring values
  - Iterative tasks

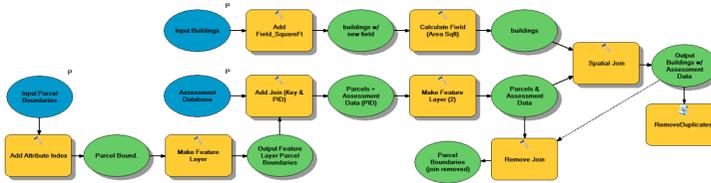
### AVAILABLE DATA



## METHODOLOGY

### (a) Combine Datasets

- Combine spatial and non-spatial data into single point feature class - through attribute (Key and PID fields) and spatial join



### (b) Delete Duplicates

- Building footprint data from the City of Fredericton includes all structures. (including sheds and auxiliary buildings found on properties). If left included - would lead to over-estimates of damage in Hazus
- Python script to search duplicate parcels PIDs and delete the record with the smallest area

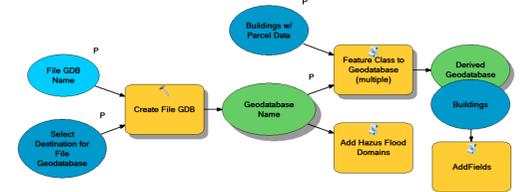
```

Building = arcpy.GetParametersAsText()
# Search variables
inPolyLayer = Building
rows = arcpy.SearchCursor(inPolyLayer)
field = "PID"
field2 = "AreaSqM"
dupList = []
for row in rows:
    val = row.getValue(field)
    dupList.append(val)
# Remove duplicate PID from iterator list
dupList = list(set(dupList))
# Start processing
for dup in dupList:
    query = "field2 <= " + str(dup)
    arcpy.SelectLayerByAttribute_management(inPolyLayer, "SUBSET_SELECTION", field)
    rows = arcpy.UpdateCursor(inPolyLayer, field2)
    for row in rows:
        if row.getValue(field2) <= dup:
            rows.deleteRow(row)
        else:
            rows.deleteRow(row)
    
```



### (c) Create Geodatabase & Import buildings

- Create File geodatabase and import the domains as per Hazus data schema: Import the buildings shape files, add fields (names, type, length, etc.) required for Hazus Flood



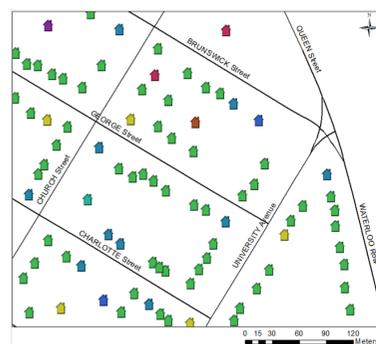
### (d) Query and Reclassify Attributes

- Hazus database design contains pre-configured domains for loss estimation, attributes are queried and the Hazus schema fields are populated with the appropriate coded values or properties

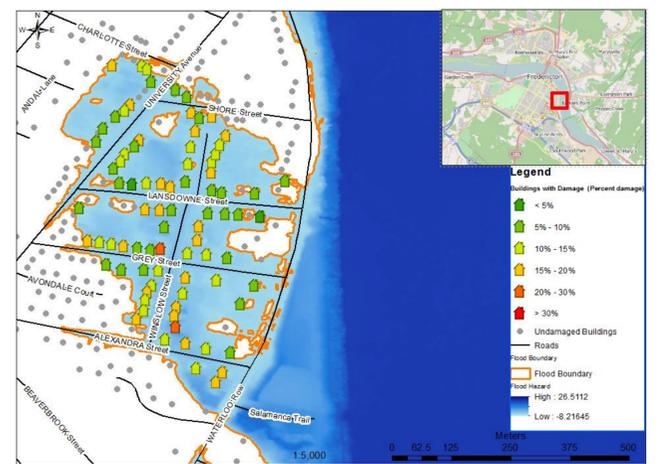
Hazus Category	Hazus Description	NB Code	NB Code Description	Property Description contains
RES1	Single Family Home	120	Residential Improved	House & Lot
RES2	Mobile Home	104, 105	Mobile Home Parks, Mobile/Mini Home w/ Land	Mini Home, Mobile & Lot
RES3A	Duplex 1-2 units	120	Residential Improved	Duplex
RES3C	Apartment 5-9 units	120	Residential Improved	Apt. Bld
COM1	Retail trade	202-207, 210, 214	Service stations, restaurant	

## RESULTS

The results of this project illustrate how separately maintained datasets from multiple organizations, collected and updated over different time periods can be quickly merged and reclassified into a single dataset. The compiled data may be used for building-by-building scenario computations or aggregated to estimate loss at the dissemination area (census block) level. Loss estimations from Hazus indicate ~ \$6million in damages to structures in Fredericton, based on the 2008 flood.



Data, after ETL, classified by Hazus Occupancy categories



Partial results from Hazus loss estimation using data derived via ETL models and python scripts