Natural Resources Canada | Ressources naturelles Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 8402

Inventory models for regional-scale natural hazards risk assessment

S.K. Ploeger, M.-J. Nollet, M. Sawada, and A. Abo El Ezz

2018

Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 8402

Inventory models for regional-scale natural hazards risk assessment

S.K. Ploeger[1], M.-J. Nollet[2], M. Sawada[1], and A. Abo El Ezz[3]

[1] University of Ottawa, 60 University Private, Ottawa, Ontario K1N 6N5
[2] École de technologie supérieure, 1100, rue Notre-Dame Ouest, Montréal, Quebec H3C 1K3
[3] Geological Survey of Canada, 490, rue de la Couronne, Québec, Quebec G1K 9A9

2018

© Her Majesty the Queen in Right of Canada, as represented by the Minister of Natural Resources, 2018

Information contained in this publication or product may be reproduced, in part or in whole, and by any means, for personal or public non-commercial purposes, without charge or further permission, unless otherwise specified.
You are asked to:
- exercise due diligence in ensuring the accuracy of the materials reproduced;
- indicate the complete title of the materials reproduced, and the name of the author organization; and
- indicate that the reproduction is a copy of an official work that is published by Natural Resources Canada (NRCan) and that the reproduction has not been produced in affiliation with, or with the endorsement of, NRCan.

Commercial reproduction and distribution is prohibited except with written permission from NRCan. For more information, contact NRCan at nrcan.copyrightdroitdauteur.rncan@canada.ca.

Permanent link: https://doi.org/10.4095/308352

This publication is available for free download through GEOSCAN (http://geoscan.nrcan.gc.ca/).

Recommended citation
Ploeger, S.K., Nollet, M.-J., Sawada, M., and Abo El Ezz, A., 2018. Inventory models for regional-scale natural hazards risk assessment; Geological Survey of Canada, Open File 8402, 43 p. https://doi.org/10.4095/308352

Publications in this series have not been edited; they are released as submitted by the author.

ABSTRACT

This document presents the results of research efforts aimed to develop inventory models of the demography and general building stock for urban centers and rural communities across eastern Canada. The inventory models are intended for use in the rapid seismic risk assessment tool ER^2 (for Rapid Risk Evaluation). Research and development of the inventory models were carried out jointly by the University of Ottawa and École de technologie supérieure (ETS) Montréal. These were part of the larger Prompt Evaluation of Seismic Risk project (PESR CSSP-2016-CP-2283), led by National Resources Canada (NRCan) for the Canadian Safety and Security Program (CSSP) managed by Defence Research and Development Canada (DRDC) Centre for Security Science (CSS) and Public Safety Canada. The ER^2 tool informs the public safety community and emergency management decision makers with information on various aspects of seismic risk. It consists of two software components for two distinct types of use. The first focusses on near real-time risk analyses following a major earthquake event, while the second component supports various risk assessment initiatives for scenario-based risk analyses.

The inventory models were developed across a study area extending from the Greater Toronto Area (Ontario) to Quebec City (Quebec). The study boundary encompasses approximately 123,455 km^2, 6,398 census units (census tracts and dissemination areas) and over 4.2 million buildings. Detailed inventories for building types and occupancy classes were conducted for 12 municipalities ranging from rural communities to large urban centres in both Ontario and Quebec. Since these inventories cover only about 0.6% of the study area (km^2) or 4.8% of the total number of buildings, a procedure had to be developed to extrapolate representative building information from the detailed inventories to be applied across the remaining census units in the study area. First, the procedure started with an estimation on the number of buildings for each census unit based on available geospatial datasets; demographic information was also collected. Second, each census unit was identified by an IoX class code that accounts its population (by size and as populated or non-populated), land use (as residential or commercial, and related densification) and average age (before or after 1960). All census units within the study area was represented in total by 45 different IoX classes. Third, the distribution of building characteristics within the detailed inventory were summarized by IoX code which provide a reasonable representation of the actual construction practices. Within the inventoried buildings, around 94% are residential buildings, 93% are wood constructions and around 74% were build after 1960. Distribution of building characteristics differ depending on the size of the community, its main land use and the average year of construction.

In order to estimate the potential economic losses from earthquake scenarios, average square footage, and replacement and content values in dollar terms were associated for various occupancy classes in each default IoX class. To further estimate potential social losses (injuries, fatalities, shelter needs), demography distribution models were built considering three common times of the day: 2am (nighttime), 2pm (daytime) and 5pm (commuting time), accompanied with respective residential, working and commuting population estimations. A two-tiered approach was used to address population distributions. These tiers represent a 'rural' model where it is assumed that the 'workers' work within their census unit and the 'urban' model where it is assumed that most of the 'workers' commute to another census unit.

RÉSUMÉ

Ce document présente les résultats de travaux de recherche visant à développer des modèles d'inventaires de bâtiments et de démographie des centres urbains et de communautés rurales pour l'est du Canada. Ces modèles seront utilisés dans l'outil d'évaluation rapide du risque ER^2. Les travaux de recherche et le développement des inventaires ont été réalisés conjointement par l'Université d'Ottawa et l'École de technologie supérieure (ÉTS) à Montréal. Ils s'intègrent au projet intitulé "Prompt Evaluation of Seismic Risk" (PESR CSSP-2016-CP-2283) mené par Ressources naturelles Canada (NRCan) dans le cadre du programme canadien pour la sûreté et la sécurité (PCSS) de Recherche et développement pour la défense Canada (RDDC), en partenariat avec Sécurité publique Canada. L'outil ER^2 vise à informer les intervenants en sécurité publique et en gestion des mesures d'urgence sur les différents aspects du risque sismique. Il comprend deux modules à usages distincts. Le premier vise l'estimation des pertes et dommages immédiatement après un évènement sismique. Le second permet de réaliser différents scénarios de risques sismiques.

La zone d'étude du projet s'étend du grand Toronto (Ontario) à la ville de Québec (Québec) et couvre une superficie d'environ 123,455 km^2. Elle comprend 398 unités de recensement (secteurs de recensement et aires de dissémination) et plus de 4.2 millions de bâtiments. Des inventaires détaillés de bâtiments relevant les classes de bâtiments et d'occupation ont été réalisés pour 12 municipalités de tailles et vocations différentes, petites municipalités rurales et grands centres urbains de l'Ontario et du Québec. Cependant, ces inventaires ne couvrent que 0.6% de la zone d'étude (km^2) ou 4.8% du nombre total de bâtiments. Ainsi, une procédure a été développée pour extrapoler les informations recueillies par les inventaires détaillés pour construire des inventaires par défaut qui peuvent être appliqués à l'ensemble des autres unités de recensement de la zone d'étude. En premier lieu, le nombre de bâtiments de chaque unité de recensement est estimé et des données démographiques sont extraites des bases de données géo-spatiales disponibles. En second lieu, chaque unité de recensement est identifiée par un code de classe IoX qui considère la population (par taille et densité de population), l'utilisation du sol (résidentielle ou commerciale et selon la densification urbaine) et l'âge moyen du bâti (avant ou après 1960). Un total de 45 classes IoX différentes ont été définies pour représenter toutes les unités de recensement de la zone d'étude. En troisième lieu, la distribution des caractéristiques des bâtiments (classes de bâtiments et d'occupation) relevées par les inventaires détaillés sont extrapolées et résumées par un code IoX donnant une représentation raisonnable des pratiques constructives. Globalement, les inventaires détaillés ont établi que près de 94% des bâtiments sont à usage résidentiel, environ 93% sont construits en bois et 74% ont été construits après 1960. La distribution des caractéristiques des bâtiments (occupation, matériau, âge) dépend cependant de la taille de la municipalité, l'utilisation principale du sol et l'âge moyen du bâti.

Les inventaires par défaut (classes IoX) incluent aussi les superficies type de plancher et les coûts de remplacement en fonction des classes d'occupation pour l'estimation des pertes économiques issues de différents scénarios sismiques. L'estimation des pertes sociales (blessures, décès, besoins en abris temporaires) s'appuie sur une estimation de la population résidente, au travail ou en déplacement pour trois périodes de la journée : 2am (nuit), 2pm (jour) et 5pm (période de déplacements). Les distributions de population ont été définies par deux modèles : un modèle « rural » supposant que les travailleurs oeuvrent dans leur unité de recensement respective et un modèle « urbain » supposant que la plupart des travailleurs se déplacent vers une autre unité de recensement que celle de leur résidence.

TABLE OF CONTENTS

Abstract i
Résumé ii
Table of contents iii
List of tables iv
List of figures iv
1. Introduction 1
1.1. Context of the project 1
1.2. Seismic risk assessment tool (ER^2) 1
1.3. Objective of the project 2
2. Building inventory 3
2.1. Detailed inventory 4
2.2. Inferred inventories 7
2.2.1. Number of buildings 7
2.2.2. Community open datasets 8
2.3. Building distributions 9
2.3.1. Inference of size by population –IoS (**##**XXXX) 10
2.3.2. Inference of land use - IoL (XX**###**X) 11
2.3.3. Inference of age – IoA (XXXXX**#**) 12
2.3.4. IoX Classes 13
3. Economic inventory 16
3.1. Square footage 16
3.2. Replacement and content costs 17
4. Population distribution 19
4.1. Rural population model 19
4.2. Urban population model 21
4.2.1. EPOI surface 21
4.2.2. Population surface 23
5. Conclusion and recommandations 24
References 25
Appendix A : Building and occupancy classes 27
Appendix B : Summary of building inventories 29
Appendix C: Toolset to predict land use 30
Appendix D: Examples of distribution of building characteristics according to IoX codes 33
Appendix E: Replacement cost according to occupancy class and Iox Code (community size) 36
Appendix F: Skew normal function to calculate the distance weights from each census unit to all other census 37

LIST OF TABLES

Table 1 : Summary of building inventory from Urban RAT 5
Table 2: Summary building inventory from walking survey and IMR. 5
Table 3: Summary of communities with open datasets used to determine the number of buildings within their respective jurisdictions. 9
Table 4 : Summary of IoX Classes codification 13
Table 5 : Summary of the number of census units that have data and no data available 14
Table 6: Summary of communities, their IoX size and the number of buildings that have a square footage within their inventory 16
Table 7: Comparison of the three replacement cost tables in US$. Note that RES 1 is represented by an average 2-storey home. M & S (2014) represent Marshall and Swift 17
Table 8: Summary of the Census datasets that were used and how they were coded. 20
Table 9: Summary of the rural population model. 20
Table 10: Summarizes the translations of the EPOI NAICS to the ER^2 occupancy classes. 22

LIST OF FIGURES

Figure 1: The study area boundaries in eastern Canada as well as earthquake events. The eastern extension from Quebec City to Chicoutimi is planned to be added in the next phase of the project. 3
Figure 2: Distribution of buildings between occupancy classes and construction material for some of the inventoried communities. 6

INVENTORY MODELS FOR REGIONAL-SCALE NATURAL HAZARDS RISK ASSESSMENT

S.K. Ploeger, M.-J. Nollet, M. Sawada, and A. Abo El Ezz

This document presents the adopted approach and results of the research activity aimed at the development of an inventory database to be used for rapid seismic risk assessment. This activity was carried out jointly at University of Ottawa and École de technologie supérieure (ETS) Montréal as a part of a larger research effort on the Prompt Evaluation of Seismic Risk (PESR) project led by National Resources Canada (NRCan) in accordance with the Canadian Safety and Security Program (CSSP) managed through Defence Research and Development Canada (DRDC) Centre for Security Science (CSS).

1. INTRODUCTION

1.1. Context of the project

The PESR project gathers expertise from academia and government at federal, provincial/territorial (P/T) and municipal levels to collectively develop standardized methods and tools for natural hazard (earthquakes, floods, etc.) risk assessment and to promote its understanding, acceptance and widespread usage. To attain these objectives and to meet the pressing need for Canadian municipalities and other levels of governments to perform multi-hazard risk assessment, a few existing risk assessment tools were considered for adaptation (Nastev and Todorov 2013, Nastev 2014). However, it was concluded that these tools could not be fully adapted for the Canadian hazard and exposure settings. The reasons were twofold. First, some of the methods could not be directly applied in this context because of the lack of required information (e.g., flood discharge vs. frequency regressions). Second, the option for external contractors was costly and inefficient.

As a result, the focus was shifted to the development of a user-friendly purpose-built tool referred to as Rapid Risk Evaluation (ER^2). This newly developed tool is equipped with step-by-step prompts and national out-of-the-box capacity which can be run by a simple 'push of the button'. The objective is to provide the public safety community with a better understanding of the local vulnerability to natural hazards (earthquakes and floods) and with the capability to run otherwise complex risk scenarios. This facilitates risk identification and risk analysis for various aspects of emergency management decision making (PSC 2012).

1.2. Seismic risk assessment tool (ER^2)

Seismic risk assessment at urban or regional scales involves the assessment of seismic hazard, the local inventory of assets at risk and its respective vulnerability. The seismic risk assessment module of ER^2 has been developed to include both shakemap and vulnerability computation capacity (Nastev et al. 2015). ER^2 consists of two software components for two distinct types of use. The first focusses on near real-time risk analysis following a major earthquake event. This component fills the current gap in the federal government's ability to automatically generate and display potential impacts of major earthquakes, informing the greater emergency management and public safety community in near-real time. This is

achieved by a continuous connection to the national and local seismograph networks where the spatial distribution of seismic parameters and its attenuation is calibrated against acquired real-time data immediately following an earthquake event. Damage as well as economic and social losses are then generated based on the calibrated shakemap.

The second planned component of ER^2 will provide further capability for scenario-based risk analyses. A web-based platform with national coverage and comprehensive databases is planned to be offered to the non-expert public safety community. In this component, 129 building classes and 33 occupancy classes are considered for computation of physical damage as well as economic and social losses. The tool also calculates the probability of the building components, its structural system, non-structural acceleration-sensitive content and non-structural drift-sensitive content, being in each of the five potential damage states (none, slight, moderate, extensive and complete). Based on these probabilities, indoor casualties at four severity levels (minor, serious, life-threatening and fatal) and economic losses sustained by the building components and contents are calculated. The mean damage factor (MDF), defined as a fraction of the replacement cost, and the coefficient of variation (COV) are also calculated. In the current version, time of the day is not explicitly considered. Considering that this parameter is important for the computation of social losses, it will be included in a later phase of the tool's development.

1.3. Objective of the project

In view of performing near real-time or scenario-based seismic risk analyses with ER^2, different databases that support the ER^2 application should be available for large, moderate and small-sized urban centers as well as rural communities in Canada. This requires creating a building inventory that adopts the predefined 129 building classes and 33 occupancy classes.

The objective of this joint research project is to develop a database / inventory of the general building stock to be used with the ER^2 tool. The primary goal was to:

1) Collect actual building inventory data for several communities, and
2) Generate a default building inventory for the entire study area.

The secondary goal was to generate economic and population distributions for the study area to support the indoor casualty and economic loss computation modules of ER^2.

2. Building inventory

Inventory of the assets at risk is a critical component of the risk assessment process in which the existing buildings and structures with similar design and loss characteristics can be grouped into a set of representative distributions of structural and occupancy classes. Due to rapid population and infrastructure changes in growing communities, the inventory must be updated on a regular basis. A building Inventory of all structures can be generated at a local scale by sidewalk surveys, at an urban scale by interpreting municipal property databases and at a regional scale by the interpretation of Statistics Canada's national coverage demography and property variables. In terms of structural building type information, the first approach relies heavily on the surveyor's experience but generates a more detailed inventory, while the other two approaches are faster and can be automated but generate a more approximate inventory. New methods and tools for rapid site specific assessment of building structural parameters and occupancies have recently been developed as efficient and effective alternatives. The first method, developed at University of Ottawa, is a urban rapid building-by-building inventory tool (Urban RAT) which takes advantage of GIS platforms to create detailed building inventories (Ploeger et al. 2014). The second method, developed at ETS, uses inference matrix relations (IMR) between available municipal data and building types to create aggregated building inventories (Nollet et al. 2012). The building inventory research was conducted by the University of Ottawa team (Dr. Mike Sawada, Dr. Kate Ploeger, and several undergraduate students) and ETS team (Dr. Marie-José Nollet, Dr. Ahmad Abo El Ezz, and several graduate students).

The study area extending from the Greater Toronto Area in the province of Ontario to Quebec City in the province of Quebec is shown in Figure 1. The study boundary contains 6,398 census units, which covers a geographic area of approximately 123,455 km^2. This area also encompasses three of Canada's largest cities, Toronto, Montreal and Ottawa.

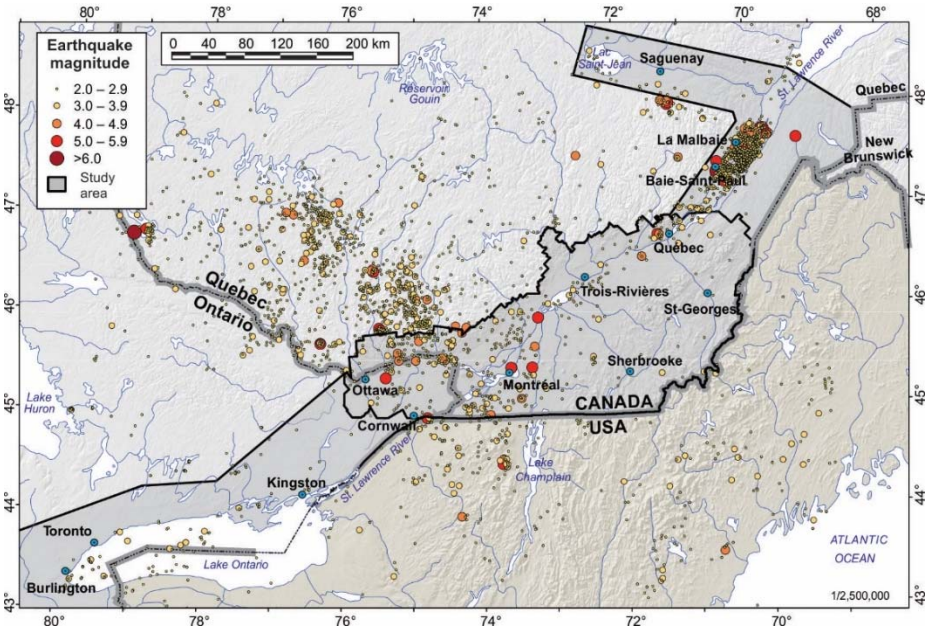


Figure 1: The study area boundaries in eastern Canada as well as earthquake events. The eastern extension from Quebec City to Chicoutimi is planned to be added in the next phase of the project.

Assembling a building inventory within such a large area is inherently challenging. It is estimated that the project area contains four to five million 'occupied' buildings, that is, a building with sustained occupants (e.g. single family dwelling, office building, gas station). In this project, basic data on inventory characteristics were accomplished using two approaches. The first approach was to conduct actual inventories within various communities, and the second approach was to leverage and compile datasets from the acquired inventories into default datasets.

2.1. Detailed inventory

The University of Ottawa research team led the project in the acquisition of building inventory. However, the University of Ottawa was already performing inventories of communities before the commencement of this project.

The detailed building inventory of the City of Ottawa began in 2011. Over the course of six years (2011-2016), a team of ten undergraduate students enrolled in the civil engineering program, mainly through the co-op program, completed an inventory of about 174,950 buildings. A summary of the building inventory is presented in Table 1. All inventory data was collected using Urban RAT. For more information on that tool refer to Ploeger et al. (2014).

Communities were selected based on criteria that reflected their seismic hazard; in an effort to be proactive with communities at potential risk. These criteria were:

- Earthquake hazard - peak ground acceleration ≥ 0.35 based on Adams and Halchuk (2003)
- Flood hazard – Designated flood areas from Environment Canada
- Soft sediments – being located on clay and/or marine sediments
- Sediment depth – being located on sediments with a depth ≥ 20 m
- Availability of supplemental information (e.g. Underwriters Bureau Insurance maps)

Flood hazard information was included to consider the eventual modelling of flood hazard in ER^2. Additionally, a region with high earthquake and flood hazard may also be vulnerable to landslides given a set of conditions.

The ETS research team conducted a building inventory of Quebec City and Boucherville. Over the course of 5 years (2012-2016), a team of 3 undergraduate and 2 graduate students enrolled in the Construction Engineering program completed an inventory of about 29,250 buildings. A summary of the building inventory is presented in Table 2. Inventory data for Boucherville was collected using municipal property database and a walking survey. For the district La Cité-Limoilou in Québec City, inventory data was collected using an inference matrix relations (IMR) correlating the probable building class to occupancy, number of stories, and year of construction. The methodology was validated against a limited walking survey engineering plans and drawings. For more information on this method refer to Nollet et al. (2012).

Table 1 : Summary of building inventory from Urban RAT

Community	Province	Population*	Size	Number of Buildings
Grand-Saint-Esprit	QC	471	Rural	237
Joliette	QC	19,621	Small	3,108
Lochaber	QC	409	Rural	242
Ottawa	ON	883,391	Large	165,760
Pakenham (Mississippi Mills)	ON	12,385	Small	582
Quebec City**	QC	516,622	Large	2,537**
Rigaud	QC	7,346	Small	1,008
Saint Gilbert	QC	282	Rural	151
Saint-Jacques-de-Leeds	QC	711	Rural	157
Shawville	QC	1,664	Small	871
Toronto	ON	2,614,406	Large	299
Total		**4,057,308**		**174,952**

*Population data from Statistics Canada[1]

**The surveyed buildings with Urban RAT in Quebec City are included in the inventory (Table 2).

Table 2: Summary building inventory from walking survey and IMR.

Community	Province	Population*	Size	Number of Buildings
Boucherville	QC	40,753	Medium	12,897
Quebec City (La Cité-Limoilou)**	QC	516,622	Large	16,358*
Total number of buildings		**557,375**		**29,255**

*Population data from Statistics Canada[1]

**This includes the 2,537 surveyed buildings by Urban Rat (Table 1)

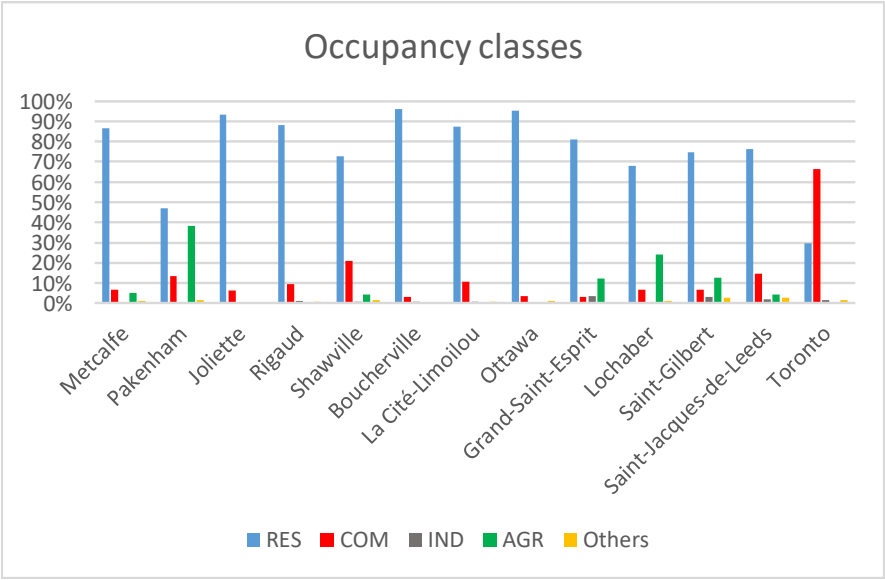
Information data collected are mandatory building characteristics required for the ER[2] inventory:

- Building class as defined in Table A-1 (***Appendix A***): Basic information on the building's structure based on FEMA-P-154 (ATC 2015) and HAZUS standards (FEMA 2012a, b), as well as (NRC 1992) (e.g. C1 is a concrete moment frame building).
- Occupancy class as defined in Table A-2 (***Appendix A***): Basic information on the use of the building based on HAZUS standards (e.g. RES1 is a single-family dwelling).
- Building height: the number of storeys.
- Year of construction: this is a reference to the building code that was in place at the time of construction (e.g. a building constructed in 1940 is considered pre-code).

Appendix B presents a summary of the detailed inventories of the communities in Table 1 and Table 2. Overall, around 94% are residential buildings, 93% are wood constructions and around 74% were build

[1] https://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo049a-eng.cfm

after 1960. Distribution of building characteristics differ depending on the size of the community, its main land use and the average year of construction. Figure 2 shows the distribution of buildings between occupancy classes and construction material for some of the inventoried communities. In small rural communities such as Grand-Saint-Esprit, Lochaber and Saint-Gilbert, 68% to 81% of the buildings are residential while 12% to 24% are agricultural. Wood buildings represent 81% to 86% of the constructions in those communities. In mostly residential communities (87% and more), wood construction represent about 94%-95% of the buildings. In La Cité-Limoilou, 86% of the buildings were built before 1960 explaining the larger proportion of masonry construction (18%) and lower proportion of wood construction (74%) for this mostly residential community. Only three census units in downtown Toronto were sampled to be inclusive of skyscrapers. This is well represented by the large proportion of steel and concrete buildings, 31% and 55% respectively, compare to the other communities.



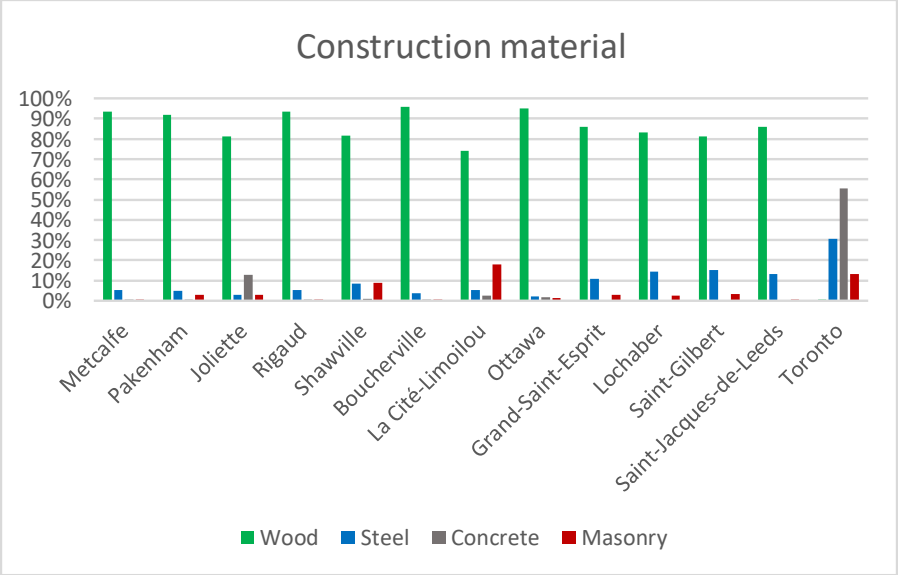


Figure 2: Distribution of buildings between occupancy classes and construction material for some of the inventoried communities.

2.2. Inferred inventories

Generally, most loss estimation programs have to infer their building inventories. HAZUS (FEMA 2012b), for example, derives its inventory from Census information, governmental agency data and/or statistical procedures, however, there is very little, if any, data regarding the structure of the building (Ploeger et al. 2013). HAZUS Canada, in particular, only includes aggregated residential building information and demographics based on Canada's 2011 Census (Ulmi et al. 2014).

In this project, the actual inventories (i.e. Table 1 and 2) cover roughly 0.6% of the study area (km^2) or 4.8% of the total number of buildings, hence, a procedure was devised to infer the building inventory based on available data. This procedure has two main components; the first is an estimate of the number of buildings per census unit, and the second is to determine a distribution of building characteristics that can reasonably represent the census unit. When combined the number of buildings can be applied to the distribution to avail an inventory tailored to the characteristics of that census unit.

2.2.1. Number of buildings

Available datasets were explored via the Geographic, Statistical and Government Information Centre (herein referred to as the GSG library) at the University of Ottawa. Three datasets were selected to calculate the number of buildings across the majority of the study area (i.e. base dataset), they are:

- The Land Information Ontario 'Building as Symbol' (BUILDSYM) and 'Building to Scale' (BUILDSCA) geospatial datasets;
- The Base de données topographiques du Quebec (BDTQ) 'bâtiments' (batim_p) dataset; and
- DMTI Spatial Inc's 'Building Points' dataset

The primary dataset used for the Ontario portion of the study area was the "Building as Symbol" (BUILDSYM) dataset produced by the Ontario Ministry of Natural Resources (Land Information Ontario). Its original publish date was 1977-01-01 with the last revision in 2011-11-10. The dataset represents "landmarks (buildings and structures) that are permanent in nature" and defined as "a structure with permanent walls and roofed construction" (Government of Ontario 2012a). This point dataset was derived using aerial photography. There are approximately 1,026,251 buildings in this dataset within the Ontario study boundary.

However, structures that are not represented in this dataset include (1) mostly small structures less than 50 m^2, and (2) structures with one dimension over pre-defined lengths at specific source scales.

The secondary dataset used for the Ontario portion of the study area was the "Building to Scale" (BUILDSCA) dataset which was also produced by the Ontario Ministry of Natural Resources (Land Information Ontario). Its original publish date was 1977-01-01 with the last revision in 2010-10-19. This dataset represents buildings that "have one side larger than 50 metres for 1:20,000 scale data or one side larger than 30 metres for 1:10,000 data" (Government of Ontario 2012b). There are approximately 78,879 buildings in this dataset within the Ontario study boundary.

To represent the Ontario portion of the study area, the BUILDSYM and BUILDSCA were combined for a total of some 1,108,951 buildings. All overlapping buildings (i.e. duplications) were omitted.

There are three major limitations to this combined dataset. First, it overestimates the number of buildings in rural areas as most buildings were inventoried from aerial photographs. To briefly elaborate, in rural regions, residential properties often contain multiple buildings including car garages and sheds that are not likely occupied throughout the day or night. Second, it underestimates the number of buildings in urban areas. To briefly elaborate, in downtown areas, several individual structures could be inadvertently interpreted as one building footprint. Furthermore, rapid development, especially in suburban regions, may not be captured within these datasets as they were constructed after the LIO data was acquired. Third, the City of Toronto building inventory is not represented in the dataset.

The primary dataset used for the Quebec portion of the study area was the Base de données topographiques du Québec (BDTQ) buildings (batim_p) dataset produced by the Ministère des Resources naturelles et faune du Québec. This dataset is set up in a grid/tile format that covers most of the populated regions of Québec; each tile covering roughly 279 km^2. Approximately 356 coverage tiles were required to represent the study area. A total of 1,867,793 buildings in this dataset are represented in the Quebec study area.

A secondary dataset was used to maximize coverage of the building count. In the CanMap Suite, DMTI Spatial Inc. has a national dataset called 'Building Points' which defines permanent walled and roofed structures. A total of 175,872 buildings for both Ontario and Quebec were included in the overall dataset; all overlapping points with other datasets were omitted.

The combination of the above four edited datasets created the base dataset which consisted of over 2.9 million buildings.

2.2.2. Community open datasets

In order to address some of the above limitations, various community open data websites were consulted for useful information. Thirteen communities and/or regions had freely available datasets on their websites that provided information on the number of buildings within their jurisdiction (Table 3). These datasets are fairly recent and mostly consist of address points and civic addresses.

Table 3: Summary of communities with open datasets used to determine the number of buildings within their respective jurisdictions.

Community/Region	Province	Type of data	Number of Buildings	Number of census units
Burlington	ON	Building footprints	44,498	43
Durham Region	ON	Civic Addresses	223,654	178
Gatineau	QC	Adresses d'immeubles	92,490	56
Hamilton	ON	Address points	247,011	23
Kingston	ON	Civic Address	62,461	30
Laval	QC	Adresse Civique	132,305	81
Longueuil	QC	Adresses	77,860	53
Montreal	QC	Adresse ponctuelle	335,297	540
Oakville	ON	Address Points	64,754	40
Ottawa	QC	Address Points	367,855	22
Peel Region	ON	Address points	411,830	232
Toronto	ON	Address Points	522,622	544
York region	ON	Address Points	333,160	196
Total			**2,915,797**	**2,038**

Subsequently, buildings from the Section 2.2.1 databases were removed from the base inventory and updated to include buildings from the open datasets which represent 31.8% of the census units in the study area.

There are some limitations when using the above open datasets. First, there is likely an overestimation of buildings as many datasets represent addresses within the community. For example, a row of townhouses might be one structure but represented as six addresses. Second, some parcels of land in the rural regions of a city may be represented as a point, when in fact, no occupied buildings exist within that space.

In total, it is estimated that there are 4,230,118 buildings within the entire study area.

2.3. Building distributions

The premise of this project's default inventory is that census units often share similar characteristics. If an actual building inventory (represented as building classes and occupancy classes) shares similarities with a census unit that does not have an inventory, then it is reasonable to assume that these similarities would be reflected in the distribution of the building inventory. For example, a suburb in Toronto will likely have similar building characteristics as a suburb in Ottawa.

In order to better assign building distributions to each census unit, a 6 digits code (XXXXXX) identified IoX Class was assigned to each census unit based on certain key characteristics. They are:

- Inference of the community size by population: *IoS (**##**XXXX)* – Field 1 and 2;
- Inference of land use in the census unit: *IoL (XX**###**X)* – Field 3, 4 and 5;
- Inference of the age of buildings within the census unit: *IoA (XXXXX**#**)* – Field 6.

The following is a discussion on how each of the above characteristics were coded in order to assign similar characteristics (see Table 4).

2.3.1. Inference of size by population –IoS (##XXXX)

Census units across the study area were classified into four distinct groups based on the population of the community in which they are located in. The premise for this inference is that the census units, although located in different geographic areas, may share similarities between them. For example, a census unit in a rural area in eastern Ontario will likely share similar characteristics as one in central Ontario, and not with a large community such as Toronto. The key datasets required in this step are Census data (population, area) and data that reflects business locations (e.g .DMTI Spatial Inc's Enhanced Points of Interest (EPOIs)). Available datasets were explored via the GSG library and the aforementioned EPOIs dataset was selected in addition to Census data.

The first step is to classify each census unit by the population of the community. According to Statistics Canada, population centres are grouped into four classes (see https://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo049a-eng.cfm); a fourth class was added to distinguish between small population centres and rural regions (< 1,000):

- Rural regions (< 1,000)
- Small population centres (1,000 to 29,999)
- Medium population centres (30,000 to 99,999)
- Large population centres (> 100,000)

To accomplish this task, all census units were dissolved according to their CSDNAME (Census subdivision name). During this task, the population field was also summed to get the total population for each CSDNAME. The 22,750 dissemination areas were dissolved into 811 Census subdivision names. Note that dissemination areas were used as they cover the entire study area.

The first field for size of population (IoSClass1) was added and each of the 811 Census subdivision names were classified as either:

- Rural population (< 1,000)
- Small population centre (1,000 to 29,999)
- Medium population centre (30,000 to 99,999)
- Large population centre (≥ 100,000)

When census subdivision names and populations were incomplete at boundary edges, they were manually input. In the IoX Class, this field is at the first position as **#**XXXXX.

A second step was undertaken to further distinguish between populated regions and truly rural regions. For example, some Census subdivisions (cities) have a geographic area of over 1,000 km^2 but have a few small communities located within their jurisdiction. Data acquired in the previous step was spatially joined back to the dissemination area Shapefile. A second field for size of population (IoSClass2) was added to label truly rural dissemination areas and/or green spaces.

1. Rural regions and/or green spaces (population < 1000 and population density < 400 per km^2 (with EPOI exception see below); https://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo049a-eng.cfm)

2. Populated (population > 1000 and population density > 400; population < 1000 and population density < 400 and EPOI total >100 (this stipulation captures census units like Pearson Airport)

In the IoX Class, this field is in second position as X**#**XXXX.

Note that EPOIs within census tracts were treated differently. Instead of using an EPOI threshold of 100 (that is, there are 100 businesses located in this census unit), it was changed to a threshold of 200 or an EPOI density of 50. These changes were made to reflect the difference in geographic area between census tracts and dissemination areas.

*2.3.2. Inference of land use - IoL (XX**###**X)*

Similarly to the inference of population codes, this section also acknowledges the similarities of the land use within census units. For example, a suburban residential area in Ottawa will be similar to one in Montreal, that is, an area dominated by wood-frame single-family dwellings. A downtown core, on the other hand, would most likely be dominated by a mixture of buildings of all ages, heights and building types, especially numerous high-rise concrete buildings.

The first step is to determine the similarities between groups of building classes which are most likely to have similar vulnerability (Nollet et al. 2018). These were determined to be:

- Rural and/or green spaces
- Predominately residential (with or without a small commercial strip)
- Commercial/industrial park (with little to no residential)
- Commercial downtown (typical of medium-size community or suburban downtown areas)
- Dense commercial downtown (typical large urban centre downtown areas)
- Predominately heritage buildings
- Predominately skyscrapers

Note that Inference of age will further distinguish between the above groups.

The second step is to classify the general land use for each dissemination area. A field for land use (IoLClass1) was added and each unit was classified as either

- Rural and/or green space which all have IoS 'X1XXXX' designations.
- Predominately residential
- Predominately commercial

In the IoX Class, this field is in the third position as XX**#**XXX.

A model was developed to address how a census unit can be coded as either residential or commercial based on available data. Given a set of random variables, $S_i \in \{1, \ldots, n\}$ within each census unit, to what degree can these be used to predict landuse within each census unit? There are a number of classification algorithms that can address this question including, Bayesian inference, support vector machines (SVM) and the more recently developed deep neural networks.

In order to predict land use as a function of multiple variables (e.g. EPOIs, zoning), a toolset was developed using the deep learning library MxNet in the R language as well as the machine learning library e1071. For example, a variable set might consist of a variable (S_1) to represent the population density within each census unit (S_2) to represent the density of EPOIs and any other relevant variables ($S_{3,\ldots,n}$). Using the

selected variables and one outcome (i.e. commercial, residential) for a subset of the census units, a prediction of the outcome at all census units can be deduced ($F_{1,\ldots,n}$). For an example of how the model is applied refer to ***Appendix C***.

For census tracts (ct), proportions by geographic area for each dissemination area (da) within a ct was used to assign the IoLClass1. For example, if ct1 contained five das (total area of 10km^2) where three das (area of 6km^2) had an IoLClass1 of 2 and two das (area of 4km^2) had an IoLClass1 of 3, then the IoLClass1 for the ct would be 2.

A third step was undertaken to further distinguish between commercial classes (COM). This was deduced by population density; the premise being that population density reflects some building characteristics. For example, a COM census unit with a population density of 15,000 is likely to contain several multi-family dwellings (apartment buildings) that are concrete high-rises. A field for land use (IoLClass2) was added and each COM unit was classified as either

- No further designation
- Likely a commercial/industrial park (pop density 0-1000)
- Likely a commercial downtown/strip (pop density 1000-10000)
- Likely a dense commercial/residential downtown (pop density > 10,000)

<u>In the IoX Class, this field is in forth position as XXX**#**XX.</u>

Note that a third IoL class was planned to reflect census units with other unique building characteristics such as heritage buildings and skyscrapers. This has not been implemented and remains in the <u>IoX Class as 'X', in fifth position as XXXX**X**X.</u>

*2.3.3. Inference of age – IoA (XXXXX**#**)*

Many buildings have unique characteristics associated with the year of construction. For example, a census unit with the majority of buildings constructed before 1960 will likely have URM within the unit. Additionally, the performance of structures will likely be poorer when compared to areas constructed after 1960 due to introduction of seismic provisions in the building codes and standards. Data from Census Canada's 2011 National Household Survey (i.e. total number of occupied dwellings by periods of construction) was utilized.

A field for age (IoAClass1) was added and classified as either:

- Proportion of buildings constructed before 1960
- Proportion of building constructed on or after 1960

<u>In the IoX Class, this field is in sixth position as XXXXX**#**.</u>

Given the above code, an IoX Class of 3221X2 has the following characteristics:

- Medium-size population centre (IoSClass1 = 3)
- Populated region within the community (IoSClass2 = 2)
- Predominately residential (IosLClass1 = 2)
- With no further designation (IosLClass1 = 1)
- With most buildings constructed after 1960 (IoAClass = 2)

Table 4 present a summary of the IoX Class and codification of the key characteristics.

Table 4 : Summary of IoX Classes codification

IoX Class	Characteristic	Categories
#XXXXX	IoSClass1	1. Rural population (< 1,000) 2. Small population centre (1,000 to 29,999) 3. Medium population centre (30,000 to 99,999) 4. Large population centre (≥ 100,000)
X#XXXX	IoSClass2	1. Rural regions and/or green spaces 2. Populated
XX#XXX	IoLClass1	1. Rural and/or green space which all have IoS 'X1XXXX' designations 2. Predominately residential 3. Predominately commercial
XXX#XX	IoLClass2 For COM census unit	1. No further designation 2. Likely a commercial/industrial park (pop density 0-1000) 3. Likely a commercial downtown/strip (pop density 1000-10000) 4. Likely a dense commercial/residential downtown (pop density > 10,000)
XXXX#X	IoLClass3	For unique building characteristics such as heritage buildings and skyscrapers (not defined yet)
XXXXX#	IoAClass1	1. Proportion of buildings constructed before 1960 2. Proportion of building constructed on or after 1960

2.3.4. IoX Classes

The study area is comprised of 6,398 census units; 2,859 are census tracts (ct) and 3,539 are dissemination areas (da). Dissemination areas are a smaller unit than the census tract. However there are inherent representation issues when using dissemination areas in many populated regions as some dissemination areas only contain a handful of buildings which would it be difficult to represent using a default distribution. Therefore, census tracts were used where possible (i.e. 15 census metropolitan areas).

The study area is represented by 45 IoX classes; refer to Table 5. Building distributions were assigned to each census unit from the summaries of actual inventories (see Apendix B). For example, 236 actual building inventories exist to represent the 1111X1 IoX class. These inventories can be used to create a distribution template that can be applied to all census units with the 1111X1 designation. There are, however, 17 IoX classes (5.7% of the study area) that do not have available data within that IoX class. In these circumstances, other IoX distributions were used as a template to represent the missing IoX classes. Distributions and templates were matched based on expert opinion. In Table 5, census units that have no data to construct a representative building distribution were supplemented by those that did.

The actual inventory was applied to census units when data was available.

Table 5 : Summary of the number of census units that have data and no data available

IoX Class	Available	No data*	Represented by*
1111X1	236		
1111X2	124		
1211X1		1	1111X1
1211X2		1	1111X2
1221X2		1	2111X1
1232X1		1	2111X1
1232X2		1	2111X1
1233X1		1	2111X1
2111X1	503		
2111X2	1367		
2211X1		4	4111X1
2211X2		9	4211X2
2221X1	274		
2221X2	331		
2224X2		1	4224X2
2232X1	51		
2232X2	171		
2233X1	53		
2233X2	78		
3111X1		43	4111X1
3111X2	173		
3211X1		1	4211X1
3211X2		6	4211X2
3221X1		182	2221X1
3221X2	345		
3224X1		1	4224X1
3224X2		2	4224X2
3232X1	19		
3232X2		76	2232X2
3233X1		35	2233X1
3233X2	110		
4111X1	18		
4111X2	63		
4211X1	2		
4211X2	4		
4221X1	355		
4221X2	782		

IoX Class	Available	No data*	Represented by*
4224X1	138		
4224X2	59		
4232X1	19		
4232X2	83		
4233X1	150		
4233X2	453		
4234X1	16		
4234X2	55		

*IoX Classes with no data available, were supplemented by those that did.

Appendix D presents examples of the distribution of building characteristics for different IoS Class.

3. ECONOMIC INVENTORY

One of the key elements to earthquake loss estimation models is to provide an overview of economic-related losses, including the replacement cost of buildings as well as content loss. As with most models, these estimates are based on a combination of square footage, the building costs per square foot and amount of damage that the building sustained. The later is estimated from the probability of reaching each damage state (minor, moderate, extensive and complete) obtained from the vulnerability model.

3.1. Square footage

Loss estimation programs generally implement a representative value for square footage, e.g. Means (2002). In this project, square footage was determined by using average square footage values of each occupancy class from the actual inventory. All IoX community sizes were represented in the summary table (i.e. IoX starting with 1 for IoSClass1 -rural, 2 for IoSClass2 –small, etc; refer to Table 4). IoX1 and IoX2 share the same values, and IoX3 and IoX4 also share the same values. The average square footage values for both rural and small communities were calculated from 1,408 buildings with square footage; refer to Table 6. All occupancies were represented except for COM7, COM9, IND3, IND4, IND5, RES2 and RES5. These were supplemented with the Iox3 and IoX 4 values (***Appendix E***).

The average square footage values for both medium and large communities were represented by 28,184 buildings with square footage (Table 6). All occupancies were represented. Although the Toronto inventory contained square footage, they were not considered in these calculations. The Toronto census units were specifically selected to represent units with skyscrapers, therefore, square footage would likely overestimate the values in all other large communities.

Table 6: Summary of communities, their IoX size and the number of buildings that have a square footage within their inventory

Community	IoX Class	Number of buildings with ft^2
Lochaber	IoX1	9
Grand-Saint-Esprit	IoX1	5
Saint-Gilbert	IoX1	7
Saint-Jacques-de-Leeds	IoX1	2
Joliette	IoX2	28
Rigaud	IoX2	6
Pakenham	IoX2	507
Shawville	IoX2	844
Boucherville	IoX3	275
Ottawa	IoX4	26,624
Quebec City	IoX4	1,285

3.2. Replacement and content costs

There are few generic tables available that represent replacement cost valuations for the various occupancy classes. HAZUS, for example, uses the Means (2002) and the Means (2006) tables for its earthquake and flood models, respectively. However, Marshall and Swift (2014) for average buildings were used in this model. These represent not only the most recent valuations but also the Canadian estimates. Table 7 presents a comparison of the three replacement cost tables.

Table 7: Comparison of the three replacement cost tables in US$. Note that RES 1 is represented by an average 2-storey home. M & S (2014) represent Marshall and Swift

HAZUS Occupancy Class Description		Means (2002)	Means (2006)	M & S (2014)
RES1	Single family dwelling	79.29	90.15	111.31
RES2	Manufactured housing	30.90	35.75	64.97
RES3A	Multi family dwelling (Duplex)	67.24	79.84	86.08
RES3B	Multi family dwelling (3-4 units)	73.08	86.60	86.08
RES3C	Multi family dwelling (5-9 units)	125.63	154.31	86.08
RES3D	Multi family dwelling (10-19 units)	112.73	137.67	86.08
RES3E	Multi family dwelling (20-49 units)	108.86	135.39	112.96
RES3F	Multi family dwelling (+50 units)	111.69	131.93	112.96
RES4	Temporary lodging	104.63	132.52	159.99
RES5	Institutional dormitory	118.82	150.96	125.39
RES6	Nursing home	104.62	126.95	157.55
COM1	Retail trade	71.54	82.63	128.66
COM2	Wholesale trade	61.91	75.95	58.58
COM3	Personal and repair services	86.81	102.34	65.83
COM4	Prof./ Tech./ Business services	98.96	133.43	129.78
COM5	Banks	153.97	191.53	195.42
COM6	Hospital	144.60	224.29	250.23
COM7	Medical office/ clinic	129.82	164.18	163.35
COM8	Entertainment & recreation	137.02	170.51	136.77
COM9	Theaters	102.35	122.05	121.85
COM10	Parking	34.78	43.72	63.16
IND1	Heavy	73.82	88.28	134.60
IND2	Light	61.91	75.95	58.58
IND3	Food/ drugs/ chemicals	119.51	145.07	211.98
IND4	Metals/ minerals processing	119.51	145.07	211.98
IND5	High technology	119.51	145.07	211.98

HAZUS Occupancy Class Description		Means (2002)	Means (2006)	M & S (2014)
IND6	Construction	61.91	75.95	58.58
AGR1	Agriculture	61.91	75.95	58.58
REL1	Church	114.08	138.57	157.62
GOV1	General services	90.30	107.28	153.40
GOV2	Emergency response	136.10	166.59	160.68
EDU1	Schools/ libraries	92.80	115.31	161.91
EDU2	Colleges/ universities	114.68	144.73	176.29

Content costs implement the default values for the HAZUS model (see HAZUS table 3.10 (FEMA 2012b).

4. POPULATION DISTRIBUTION

In order to estimate earthquake casualties, a population model must be considered. Earthquake scenarios, much like loss estimation programs for other natural hazards, typically consider three times of the day which reflect the distribution of the general population throughout the day. These times are 2am (nighttime), 2pm (daytime) and 5pm (commuting time).

- The time of 2am reflects where the majority of population are located during all evening to morning hours. This population is most easily represented as the census population with the assumption that most people are located within their census units and in their dwellings at this hour. In HAZUS Canada, for example, the total number of persons in private households is considered (Ulmi et al. 2014).
- The time of 2pm reflects where a significant portion of the population is located during the weekday mornings and afternoons. This population is more difficult to model. This data may be derived from detailed census data such as those that are related to the labour force. In HAZUS Canada, for example, the number of people considered to be 'working' was derived from the Census-National Housing Survey (NHS) regarding those who work in the commercial and/or industrial sectors. There are similar procedures for estimating students in schools. A major limitation with this traditional approach is that it does not account for those who commute outside their home census units. Therefore, in many large urban centres, the daytime population may not be adequately represented.
- The time of 5pm reflects that a significant number of persons may be neither at work nor at home.

In loss estimation models such as HAZUS and HAZUS Canada, the aforementioned populations can be further assigned using default distribution relationships to place persons inside or outside of a building as well as at home or in 'working sectors' within the census unit; refer to Table 13.2 in the HAZUS-MH 2.1 earthquake model technical manual (FEMA 2012b) and Ulmi et al., 2014 (i.e. HAZUS Canada user and technical manual).

Given the large size of the study area (123,455 km^2 and almost 6,400 census units), a two-tiered approach was used to address population distributions. These tiers represent a 'rural' model where it is assumed that the 'workers' work within their census unit and the 'urban' model where it is assumed that most of the 'workers' commute to another census unit.

In this project, the three common times of day were considered as well as residential, working, school and commuting populations.

4.1. Rural population model

The rural population model encompasses areas that do not have an associated census tract. Therefore it is important to consider that, in this model, the term 'rural' population can include IoX community sizes up to medium population centre (i.e. IoX starting with 1, 2 and 3; refer to Table 4). There are 3,539 census units designated in the 'rural' population model; this is 55.3% of the study area's census units and 74.4% of the geographic area.

The rural population model utilizes only census data from the 2011 Census including the National Housing Survey (NHS). The key data it uses is: total population, children ages 0-15 years and the labour force classes (employed, unemployed, not in labour force (LFS)) (Table 8).

Table 8: Summary of the Census datasets that were used and how they were coded.

Census/NHS	Dataset	Coded
Census	Total population	2AM RES
NHS	In the LFS - Unemployed	[At home]
NHS	Out of the LFS	
NHS	In the LFS - Employed	[At work]
Census	Persons aged 0-15 years	[Kids all]

Using this data, relationships were applied to distribute the population into residential (RES), non-residential (NRES) and school (EDU) settings. Note that NRES populations are those working in or visiting commercial, industrial, government, agricultural and educational buildings. Both the RES and NRES populations are inclusive of indoor and outdoor settings. Table 9 summarizes each dataset and the relationships that were applied. The relationships were derived from the HAZUS 'default relationships for estimating population distribution' table (see Table 13.2 in the HAZUS earthquake model technical manual (FEMA 2012b)).

Table 9: Summary of the rural population model.

Model times*	Relationships	Notes
2AM RES	Total population *1.0	Used for both rural and urban models Does not consider persons who work during these hours
2PM RES	[At home] * 0.75	
2PM NRES	[At work]	
	[At home] * 0.25	
2PM EDU	[Kids all]	Considers all children to be in a school or educational setting
5PM RES	[At home] * 0.75	
	[Kids all] * 0.75	
5PM NRES	[At work] * 0.5	
	[At home] * 0.25	
	[Kids all] * 0.25	
5PM Commute	[At work] * 0.5	Used for both rural and urban models

* Note that RES refers to the residential population (those located in their dwelling on their property), and NRES refers to persons not located on a residential property.

4.2. Urban population model

The urban population model encompasses areas that have an associated census tract. The term 'urban' population model can include all IoX community sizes, however, the majority are IoX4 (i.e. large population centre; refer to Table 4). There are 2,859 census units designated for the 'urban' population model; this is 44.7% of the study area's census units and 25.6% of the geographic area. The communities included in the population model were: Barrie, Belleville, Drummondville, Granby, Hamilton, Kingston, Montreal, Oshawa, Ottawa, Gatineau, Peterborough, Québec City (La Cité-Limoilou), Saint-Jean-sur-Richelieu, Sherbrooke, Toronto and Trois-Rivières.

The urban population model utilizes both census data from the 2011 Census including the National Housing Survey as well as geospatial data that represents commercial activity (e.g. DMTI Spatial Inc's EPOI dataset). The urban model is comprised of two parts. The first part includes the geographic distribution of the EPOI dataset throughout each of the aforementioned communities; herein called the EPOI surface. The second part includes the geographic distribution of the population throughout the same community; herein called the population surface.

4.2.1. EPOI surface

The commercial activity dataset that was used was DMTI Spatial Inc's Enhanced Points of Interest (EPOIs) which is a national database of businesses, landmarks and recreational facilities. The EPOI dataset was cleaned to eliminate excessive overlapping points. For example, it was common to have hundreds or thousands of identical points registered at a city hall. When overlapping points were unreasonable, all but one were omitted.

The first step was to code the EPOI dataset to be compatible with the ER^2 framework. The EPOI dataset included fields that represented the points' North American Industry Classification System' (NAICS).

a) Derive the industry sector using the classification structure provided by Statistics Canada (see: http://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=307532). This can easily be coded into the ER^2 framework occupancy classes (Table 10). Note that NAICS codes regarding (1) 'Health care and social assistance' represented both COM 6 and 7, (2) 'Arts, entertainment and recreation' represented both COM 8 and 9, (3) 'Manufacturing' represented IND 1, 2, 3 and 5, (4) 'Public administration' represented both GOV 1 and 2, and (5) 'Educational services' represented both EDU 1 and 2.

The second step was to generate the EPOI surface within each community.

a) Extract EPOIs for each occupancy class and spatial join these to each community to include the number of EPOIs per census unit and the total number of EPOIs per community. The latter can be accomplished using a dissolve function on the community identifier (in this case the census CMA).
b) Calculate the distribution of each occupancy class throughout the community. This is accomplished by dividing the number of EPOIs of that occupancy class within the census unit by the number of all EPOIs of that occupancy in the community. For example, if census unit A has 55 COM1 points and there are 953 COM1 points within the community, then the proportion of COM1s in census unit A is 0.0577 (55/935).

This surface provides an overview of the distribution of all occupancy classes within a community. These surfaces should provide an indication of downtown cores, industrial sectors, etc.

Table 10: Summarizes the translations of the EPOI NAICS to the ER^2 occupancy classes.

ER^2	NAICS	
Occupancy Class	*Code*	*Sector*
COM1	45,45	Retail trade
COM2	41, 42	Wholesale trade
	48, 49	Transportation and warehousing
COM3	81	Other services (except public administration)
COM4	51	Information and cultural industries
	53	Real estate and rental and leasing
	54	Professional, scientific and technical services
	55	Management of companies and enterprises
	56	Administrative and support, waste management and remediation services
	22	Utilities
COM5	52	Finance and insurance
COM6	62	Health care and social assistance
COM7		
COM8	71	Arts, entertainment and recreation
	72	Accommodation and food services
COM9		
COM10		
IND1	31, 32, 33	Manufacturing
IND2		
IND3		
IND4	21	Mining, quarrying, and oil and gas extraction
IND5		
IND6	23	Construction
GOV1	91, 92	Public administration
GOV2		
EDU1	61	Educational services
EDU2		
AGR	11	Agriculture, forestry, fishing and hunting
REL		

4.2.2. Population surface

The population surface represents the 2PM population distribution in populated urban areas. This surface attempts to capture the RES, COM, IND, GOV, AGR, and EDU populations. Note that the 2AM RES, 2PM RES, 5PM RES and 5PM commuting populations follow the same procedure as the rural model.

The 5PM COM, IND, GOV, EDU and AGR models follow a similar procedure to that of the rural model in that these working populations are divided by two.

The 2PM COM, IND, GOV, EDU and AGR models are more evolved. The urban population model utilizes the EPOI surface and the same 2011 census and NHS data as the rural model but also includes median commuting time in minutes, the working force NAICS sectors (NHS) as well as a geospatial road network dataset (e.g. DMTI Spatial Inc's CanMap Route Logistics Network layer).

The first step is to determine the distribution of those employed in the labour force by sector. Although there are values available in the NHS that represent the number of workers per sector by census unit, there was a slight discrepancy between these values and the value that represented those employed in the labour force. In order to be consistent with the rural population model (i.e. using the LFS dataset), a distribution of the sector data was created for the LFS-employed values.

- Translate the work sectors into the ER^2 occupancy classes (see Table 10).
- Calculate the distribution of the sector data per census unit. For example, if census unit A has 100 persons who work in COM1, 20 in COM8, and 13 in IND1, then the distribution is 0.75 for COM1 (100/133), 0.15 for COM8 (20/133) and 0.10 for IND1 (13/133).
- Apply the LFS-employed values to the distribution, this provides the LFS-employed population by sector per census unit. For example, is census unit A as 175 persons who are employed in the LFS, then there are ~131 who work in the COM1 sector (175*0.75), ~26 who work in the COM8 sector (175*0.15), and ~17 who work in the IND1 sector (175*0.10).

The second step was to build a model that reflects where a person works within that community based on the sector in which they work, the EPOI surface (where they are going to), and the median commuting time.

- Determine the centroid of each census unit.
- Determine the travel time from each census unit to all others using a network analysis based on posted speed limits and distance. In this project, DMTI Spatial Inc's CanMap Route Logistics Network Layer was used.
- Calculate the distance weights from each census unit to all others using a skew-normal function centred on the median travel time to work for that census unit; see ***Appendix F*** for the skew-normal functions.
- Multiply the relative occupancy class (sector) within each census unit for each occupancy class type by the distance weights to produce the final weighting of the LFS-employed population within the given census unit for distribution among all other census units.

This process has the effect of honouring the locally relevant distribution of sectors within the relevant range of travel for the working population in each census unit.

In the first edition of ER^2, the occupancy classes were aggregated into the main sectors, they are, COM, IND, GOV, AGR, and EDU. Future iterations of the population model can also introduce new features such as rings that indicate the downtown core of major cities.

5. CONCLUSION AND RECOMMANDATIONS

This report presented the development of inventory models for the general building stock to be used with the Rapid Risk Evaluation (ER^2) tool. ER^2 will provide the public safety community near real-time risk analysis following a major earthquake event and will also support scenario-based risk analyses.

The inventory database developed for the ER^2 application was inferred for large and moderate-sized urban centers as well as rural communities across the study area extending from Toronto to Quebec City. It contains 6,398 census units, which covers a geographic area of approximately 123,455km^2 and over 4.2 million buildings. Two approaches were used to construct this inventory. The first approach conducted actual building surverys within 12 communities of different sizes in Ontario and Quebec. The second approach leveraged and compiled other building-related datasets (described in Section 2.2). The combination of these datasets allowed for an estimation of the number of buildings across the study area to be calculated including. Each census unit is identified by one of 45 IoX class codes according to its population (by size and as populated or non-populated), land use (as residential or commercial, and related densification) and average age (before or after 1960). Using actual inventory information, building characteristics within sampled IoX classes were assembled and templated across the entire study area.

The economic inventory gives average square footage as well as replacement and content values in dollar terms; these are associated to various occupancy classes in each default IoX class. Demography distribution models were built considering three common times of the day: 2am (nighttime), 2pm (daytime) and 5pm (commuting time), accompanied with respective residential, working and commuting population estimations.

The default inventories will allow ER^2 to be used in all census units across the study area. It will provide a rapid estimate of damages in near real-time following earthquake events and for scenario-based risk assessments. To encourage end users to implement their own detailed inventory, ER^2 will be developed to allow for customization of the inventory dataset.

REFERENCES

Adams, J., and Halchuk, S. 2003. Fourth generation seismic hazard maps of Canada: values for over 650 Canadian localities intended for the 2005 National Building Code of Canada. Geological Survey of Canada. Open-File Report 4459.

Applied Technology Council (ATC). 2015. *Rapid Visual Screening of Buildings for Potential Seismic Hazards: A Handbook (FEMA P-154)*. 3rd ed.; Applied Technology Council: Redwood City, CA, USA.

Federal Emergency Management Agency (FEMA). 2012a. *HAZUS-MH 2.1 – Earthquake Model User Manual*. Federal Emergency Management Agency, National Institute of Building Science: Washington, D.C., USA, 863 p.

Federal Emergency Management Agency (FEMA). 2012b. *HAZUS-MH 2.1: Multi-Hazard Loss Estimation Methodology Earthquake Model; Technical Manual*. Federal Emergency Management Agency, National Institute of Building Science: Washington, DC, USA, 718p.

National Research Council (NRC). 1992. *Manual for Screening of Buildings for Seismic Investigation*; Institute for Research in Construction, National Research Council: Ottawa, ON, Canada.

Government of Ontario. 2012a. Land Information Ontario data description: Building as Symbol.

Government of Ontario. 2012b. Land Information Ontario data description: Building to Scale.

Nastev, M. and Todorov, N. 2013. "HAZUS: A standardized methodology for flood risk assessment in Canada." Water Quality Research Journal of Canada, 38(3), 223-231.

Nastev, M. 2014. "Adapting HAZUS for seismic risk assessment in Canada." Canadian Geotechnical Journal, 51(2), 217-222.

Nastev, N., Nollet, M.-J., Abo El Ezz, A., Smirnoff, A., Ploeger,S.K., McGrath, H., Sawada,M., Stefanakis,E., and Parent,M. 2015. "Methods and tools for natural hazard risk analysis in Eastern Canada – Use of knowledge to understand vulnerability and implement mitigation measures". *ASCE Natural Hazards Review Journal*, B4015002, 13 p.

Nollet, M.-J., Abo El Ezz, A., Surprenant, O., Smirnoff, A., Nastev, M. 2018. "Earthquake Magnitude and Shaking Intensity Dependent Fragility Functions for Rapid Risk Assessment of Buildings", *Geosciences*, 8(16), doi:10.3390/geosciences8010016

Nollet, M.-J., Désilets, C., Abo El Ezz, A. et Nastev, M. 2012. Approche méthodologique d'inventaire de bâtiments pour les études de risque sismique en milieu urbain / Ville de Québec, Arrondissement La Cité-Limoilou, Commission géologique du Canada, Dossier public DP7260, 93 p.

Ploeger, S.K., Baingo, D., Parenteau, M.-P., Nastev, M., Rivard, J.R., and Sawada, M. 2013. Seismic loss estimation for the National Capital Region: a discussion on HAZUS. 3rd CSCE Annual Conference, Montreal, Quebec, Canada, 29 May – 01 June.

Ploeger. S.K., Sawada, M., Elsabbagh, A., Saatcioglu, M., Nastev, M., and Rosetti, E. 2014. Urban RAT: a new tool for virtual and site specific rapid data collection for seismic risk assessment. Journal of Computing in Civil Engineering. doi: 10.1061/(ASCE)CP.1943-5487.0000472.

Public Safety Canada (PSC). (2012). “All Hazards Risk Assessment Methodology Guidelines 2012-2013.” Ottawa, ON, 〈https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ll-hzrds-ssssmnt/ll-hzrds-ssssmnt-eng.pdf 〉 (Feb. 15, 2015).

Ulmi, M. et al., 2014. HAZUS-MH 2.1 Canada user and technical manual: Earthquake module. Geological Survey of Canada Open File 7474. pp. 235.

APPENDIX A : BUILDING AND OCCUPANCY CLASSES

The following tables described building classes and occupation classes used in ER^2.

Table A-1: Building classes as defined in ER^2 and HAZUS ®

Building classes				
	No.	Stories	Label	Description
Wood structures	1		W1	Wood light frame (< 5000 pi2)
	2		W2	Wood commercial and industrial (> 5000 pi2)
Steel Frames	3 4 5	S1L S1M S1H	S1	Steel moment frame
	6 7 8	S2L S2M S2H	S2	Steel braced frame
	9		S3	Steel Light Frame
	10 11 12	S4L S4M S4H	S4	Steel frame with cast-in-place concrete shear walls
	13 14 15	S5L S5M S5H	S5	Steel frame with unreinforced masonry infill walls
Concrete structures	16 17 18	C1L C1M C1H	C1	Concrete moment frame
	19 20 21	C2L C2M C2H	C2	Concrete shear walls
	22 23 24	C3L C3M C3H	C3	Concrete frame with unreinforced masonry infill walls
	25		PC1	Precast concrete tilt-up walls
	26 27 28	PC2L PC2M PC2H	PC2	Precast concrete frames with concrete shear walls
Masonry	29 30	RM1L RM1M	RM1	Reinforced masonry bearing walls with wood or metal deck diaphragm
	31 32 33	RM2L RM2M RM2H	RM2	Reinforced masonry bearing walls with precast concrete diaphragm
	34 35	URML URMM	URM	Unreinforced masonry bearing walls
	36	All	MH	Mobil Home

Table A-2: Occupation classes as defined in ER[2] and HAZUS ®

Codification	Occupancy class	Example descriptions as found in municipal database
	Residential	
RES1	Single family dwelling	House
RES2	Mobile home	Mobile home
RES3	Multi-family dwelling RES3A Duplex RES3B 3-4 Units RES3C 5-9 Units RES3D 10-19 Units RES3E 20-49 Units RES3F 50+ Units	Apartment/Condominium
RES4	Temporary lodging	Hotel/Motel/ *Daycare*
RES5	Institutional Dormitory	Group housing (military, college, jails)
RES6	Nursing home	CHSLD (Qc) / Retirement homes
	Commercial	
COM1	Retail Trade	Store/Hardware store/Grocery/Gas station/Car retailers
COM2	Wholesale Trade	Warehouse
COM3	Personal and Repair Services	Service station/Shop
COM4	Professional/Technical Services	Offices/Train station/Architects office/
COM5	Banks	
COM6	Hospital	
COM7	Medical Office/Clinic	
COM8	Entertainments & Recreation	Restaurants/Bars/Museum/Parks/Art galery
COM9	Theaters	
COM10	Parking	
	Industrial	
IND1	Heavy	Textile/Wood transformation/Glass
IND2	LIght	Furniture factory/Plastic/Leathe
IND3	Food/Drugs/Chemicals	Tabacco/Food
IND4	Metals/Minerals Processing	
IND5	High Technology	Electronic products/Computers
IND6	Construction	General and specialized building contractor offices
	Agriculture	
AGR1	Agriculture	Non applicable in urban area
	Religion/Non-profit	
REL1	Church/Non-profit	Church/Religious organisation
	Government	
GOV1	General services	Offices/Ministry/Postal Service
GOV2	Emergency response	Police station/Fire station
	Education	
EDU1	Grade schools	Preschool/Primary school/Highschool/Library
EDU2	Colleges/Universities	CEGEP/Professional school/Does not include group housing

APPENDIX B : SUMMARY OF BUILDING INVENTORIES

The following is a summary of the detailed inventories carried out by ***University of Ottawa (U of O)*** and ***École de technologie supérieure (ÉTS)*** as described in Section 2.1. The full inventories have data on all 36 building types, 28 occupancy classes, year of construction and more detailed engineering parameters.

This information was used to propose default inventories for similar communities.

Table B-1: Distribution of buildings according to occupancy classes for the inventoried communities

Community	IoX Size	Number of buildings	RES	COM	IND	AGR	Others
Metcalfe[A]	IoX4[A]	914	86.65%	6.67%	0.22%	5.25%	1.20%
Pakenham	IoX2	582	46.91%	13.40%	0.00%	38.32%	1.37%
Joliette	IoX2	3,108	93.28%	6.24%	0.00%	0.00%	0.48%
Rigaud	IoX2	1,008	88.10%	9.42%	1.29%	0.50%	0.69%
Shawville	IoX2	871	72.90%	20.78%	0.57%	4.25%	1.49%
Boucherville	IoX3	12,897	95.94%	3.11%	0.74%	0.02%	0.18%
La Cité-Limoilou	IoX4	16,357	87.58%	10.64%	0.87%	0.00%	0.92%
Ottawa	IoX4	164,843	95.18%	3.51%	0.11%	0.19%	1.01%
Grand-Saint-Esprit	IoX1	237	81.01%	2.95%	3.38%	12.24%	0.42%
Lochaber	IoX1	242	68.18%	6.61%	0.00%	23.97%	1.24%
Saint-Gilbert	IoX1	151	74.83%	6.62%	3.31%	12.58%	2.65%
Saint-Jacques-de-Leeds	IoX1	157	76.43%	14.65%	1.91%	4.46%	2.55%
Toronto[B]	IoX4[B]	299	29.77%	66.56%	1.67%	0.33%	1.67%
Total		**201,666**	**94.10%**	**4.36%**	**0.23%**	**0.37%**	**0.94%**

[A] Note that Metcalfe is located within the City of Ottawa boundary

[B] Note that only three census units in downtown Toronto were sampled to be inclusive of skyscrapers

Table B-2: Distribution of buildings according to construction material for the inventoried communities

Community	IoX Size	Number of buildings	Wood	Steel	Concrete	Masonry
Metcalfe[A]	IoX4[A]	914	93.65%	5.25%	0.55%	0.55%
Pakenham	IoX2	582	92.10%	4.98%	0.17%	2.75%
Joliette	IoX2	3,108	81.21%	3.09%	12.64%	3.06%
Rigaud	IoX2	1,008	93.35%	5.36%	0.60%	0.69%
Shawville	IoX2	871	81.63%	8.38%	1.03%	8.96%
Boucherville	IoX3	12,897	95.86%	3.75%	0.19%	0.19%
La Cité-Limoilou	IoX4	16,357	74.30%	5.18%	2.40%	18.11%
Ottawa	IoX4	164,843	94.98%	2.01%	1.69%	1.31%
Grand-Saint-Esprit	IoX1	237	86.08%	10.97%	0.00%	2.95%
Lochaber	IoX1	242	83.06%	14.46%	0.00%	2.48%
Saint-Gilbert	IoX1	151	81.46%	15.23%	0.00%	3.31%
Saint-Jacques-de-Leeds	IoX1	157	85.99%	13.38%	0.00%	0.64%
Toronto[B]	IoX4[B]	299	0.33%	30.77%	55.52%	13.38%
Total		**201,666**	**92.89%**	**2.55%**	**1.88%**	**2.68%**

Table B-3: Distribution of buildings according to year of construction for the inventoried communities

Community	IoX Size	Number of buildings	Average year	Before 1960	After 1960
Metcalfe[A]	IoX4[A]	914	1972	24.40%	75.60%
Pakenham	IoX2	582	1943	97.59%	2.41%
Joliette	IoX2	3,108	1978	14.74%	85.26%
Rigaud	IoX2	1,008	1975	15.67%	84.33%
Shawville	IoX2	871	1949	87.83%	12.17%
Boucherville	IoX3	12,897	1978	10.40%	89.60%
La Cité-Limoilou	IoX4	16,357	1930	86.60%	13.40%
Ottawa	IoX4	164,843	1978	21.15%	78.85%
Grand-Saint-Esprit	IoX1	237	1956	62.87%	37.13%
Lochaber	IoX1	242	1955	57.44%	42.56%
Saint-Gilbert	IoX1	151	1952	72.19%	27.81%
Saint-Jacques-de-Leeds	IoX1	157	1962	49.04%	50.96%
Toronto[B]	IoX4[B]	299	1967	34.78%	65.55%
Total		**201,666**		**26.34%**	**73.66%**

APPENDIX C: TOOLSET TO PREDICT LAND USE

This is an example to demonstrate how the land use toolset is applied.

The recommended number of back propagation iterations is 1,000 to achieve overall prediction accuracies of 90% or greater.

In order to use R within ArcGIS, (1) the arcgisbinding library must be first installed (see https://r-arcgis.github.io/), (2) background geoprocessing must be enabled and (3) the MxNet deep learning library must be installed in 64 bit R using the following statements:

```
install.packages("drat", repos = "https://cran.rstudio.com")
drat:::addRepo("dmlc")
install.packages("mxnet")
```

A user interface was developed in ArcGIS by creating a custom toolbox and adding a script tool to that toolbox (Figure C-1).

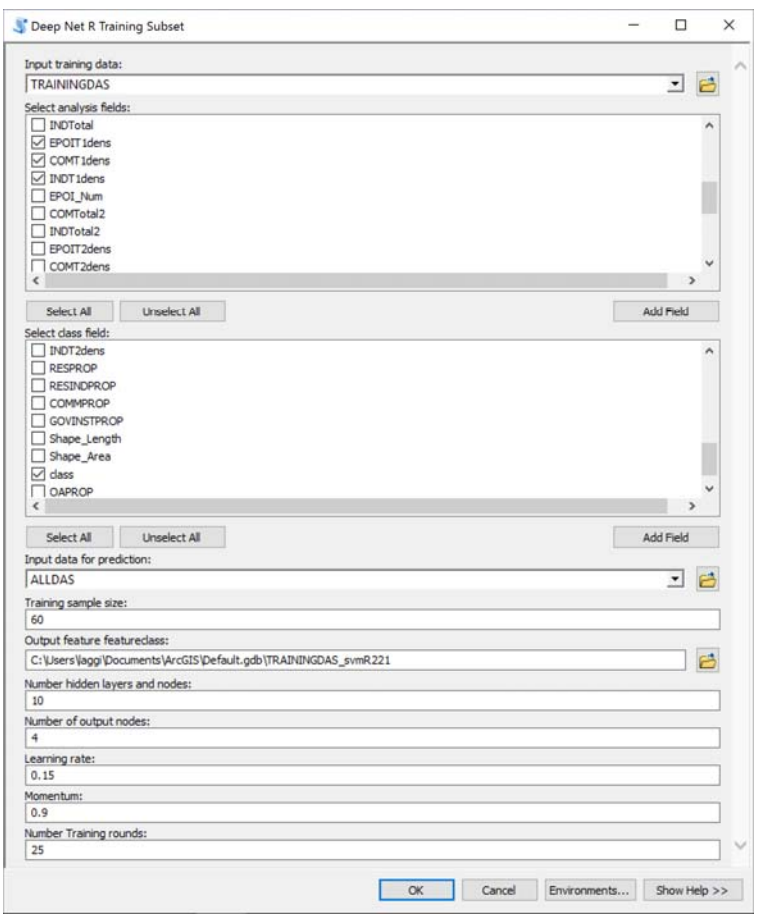


Figure C-1: Custom tool for deep neural network in R illustrating training data columns used as well as deep network hyperparamaters.

The interface in Figure C-1 allows the user to specify different hyperparameters that are critical for model accuracy. Moreover, in the above example, the number of hidden layers is specified as 10 perceptrons in the first hidden layer and four in the second. The number of output nodes must be equal to the number of classes. In the above example, there are four classes including commercial, industrial, mixed residential/commercial or residential.

For the next step, the R program was created and associated with the R tool:

```
tool_exec <- function(in_params, out_params)
{

  data.Normalization = function(x) {
    out = vector("numeric")
    ncols = dim(x)[2]
    for (i in 1:ncols) {
    k = x[, i]
    k = (k - min(k)) / (max(k) - min(k))
    out = cbind(out, k)
    }
    out
    }

  require(mxnet)
  require(mlbench)
  require(arcgisbinding)

  # GET PARAMETERS
  source_dataset = in_params[[1]]
  analysis_fields = in_params[[2]]
  source_class_field = in_params[[3]]
  data_predict = in_params[[4]]
  training_sample_size = strtoi(in_params[[5]])
  num_hidden_nodes = in_params[[6]]
  num_output_nodes = strtoi(in_params[[7]])
  learning_rate = as.numeric(in_params[[8]])
  momentum = as.numeric(in_params[[9]])
  number_training_rounds = strtoi(in_params[[10]])
  output_fc = out_params[[1]]

  num_hidden_nodes = strtoi(strsplit(num_hidden_nodes, " ")[[1]])

  trainingShapefile = arc.open(source_dataset)

  trainingData = arc.select(trainingShapefile, analysis_fields)

  numTrainingObs = length(trainingData[, 1])
  train.ind = sample(1:numTrainingObs, training_sample_size)
  train.x = data.matrix(trainingData[train.ind,])
  #train.x=scale(train.x)
  train.x = data.Normalization(train.x)
```

```
  y         =         arc.select(trainingShapefile,         source_class_field)
  y                              =                              y[train.ind,]
  train.y           =           as.numeric(factor(y))           -           1
  test.x          =          data.matrix(trainingData[-train.ind,          ])
  test.x                     =                     data.Normalization(test.x)
  #test.x=scale(test.x)
  y         =         arc.select(trainingShapefile,         source_class_field)
  y                              =                              y[-train.ind,]
  test.y           =           as.numeric(factor(y))           -           1

  mx.set.seed(0)
  model                              =                              mx.mlp(
  train.x,
  train.y,
  hidden_node                        =                        num_hidden_nodes,
  out_node                         =                        num_output_nodes,
  out_activation                           =                        "softmax",
  num.round                     =                     number_training_rounds,
  array.batch.size                                =                           15,
  learning.rate                        =                        learning_rate,
  momentum                             =                             momentum,
  eval.metric                        =                       mx.metric.accuracy
  )
  preds                =                predict(model,                test.x)
  pred.label           =           max.col(t(preds))           -           1
  confusionmat             =             table(pred.label,             test.y)
  print(confusionmat)

  testdata                      =                      arc.open(data_predict)
  testdatout         =         arc.select(testdata,         analysis_fields)
  testout.x                     =                     data.matrix(testdatout)
  testout.x                  =                  data.Normalization(testout.x)
  #testout.x=scale(testout.x)
  predfin              =              predict(model,              testout.x)
  pred.labelout          =          max.col(t(predfin))          -          1

  #                     #                     add                     field
  testdatout$class                =                as.character(pred.labelout)
  #                              #write                              output
  arc.write(output_fc,                                              testdatout)
}
```

APPENDIX D: EXAMPLES OF DISTRIBUTION OF BUILDING CHARACTERISTICS ACCORDING TO IOX CODES

Table D-1: Distribution of building characteristics for IoX 1111X1 code

IoX Code	Occupancy Class	Building type-Code	Distribution
1111X1	GOV1	RM1L-m	0,00254453
1111X1	AGR1	S1L-l	0,03562341
1111X1	COM1	S1L-l	0,00763359
1111X1	COM4	S1L-l	0,00254453
1111X1	IND1	S1L-l	0,00763359
1111X1	AGR1	S1L-m	0,01017812
1111X1	IND1	S1L-m	0,00254453
1111X1	AGR1	S1L-p	0,00763359
1111X1	AGR1	S3-l	0,01272265
1111X1	COM3	S3-l	0,01017812
1111X1	COM4	S3-l	0,00508906
1111X1	AGR1	S3-m	0,00254453
1111X1	COM2	S3-m	0,00254453
1111X1	IND1	S3-m	0,00254453
1111X1	AGR1	S3-p	0,02544529
1111X1	COM3	S3-p	0,00763359
1111X1	COM3	S5L-l	0,00254453
1111X1	GOV1	S5L-p	0,00254453
1111X1	GOV1	URML-p	0,00254453
1111X1	REL1	URML-p	0,00254453
1111X1	RES1	URML-p	0,0178117
1111X1	REL1	URMM-p	0,00254453
1111X1	AGR1	W1-l	0,00254453
1111X1	COM1	W1-l	0,00254453
1111X1	COM8	W1-l	0,00254453
1111X1	RES1	W1-l	0,17557252
1111X1	AGR1	W1-m	0,00254453
1111X1	COM1	W1-m	0,00254453
1111X1	RES1	W1-m	0,05089059
1111X1	COM3	W1-p	0,00254453
1111X1	COM4	W1-p	0,00508906
1111X1	COM8	W1-p	0,00508906
1111X1	REL1	W1-p	0,00254453
1111X1	RES1	W1-p	0,45547074
1111X1	RES3A	W1-p	0,00763359
1111X1	AGR1	W2-l	0,01272265
1111X1	COM2	W2-l	0,00254453
1111X1	COM4	W2-l	0,00254453
1111X1	GOV1	W2-l	0,00254453
1111X1	AGR1	W2-m	0,00254453
1111X1	AGR1	W2-p	0,08142494
1111X1	COM3	W2-p	0,00254453

Table D-2: Distribution of building characteristics for IoX 2221X1 code

IoX Code	Occupancy Class	Building type-Code	Distribution
2221X1	COM1	C1L-l	0,00170068
2221X1	COM1	C2L-l	0,00170068
2221X1	COM1	C2L-p	0,005102041
2221X1	COM1	RM1L-p	0,006802721
2221X1	COM1	S1L-l	0,013605442
2221X1	COM1	S1L-p	0,003401361
2221X1	COM1	S5L-p	0,00170068
2221X1	COM1	W1-l	0,00170068
2221X1	COM1	W2-l	0,00170068
2221X1	COM2	S1L-l	0,003401361
2221X1	COM3	C2L-l	0,00170068
2221X1	COM3	C2L-p	0,005102041
2221X1	COM3	S1L-l	0,006802721
2221X1	COM3	S1L-p	0,003401361
2221X1	COM3	URML-p	0,003401361
2221X1	COM3	W1-l	0,00170068
2221X1	COM4	C2L-l	0,00170068
2221X1	COM4	C2L-p	0,00170068
2221X1	COM4	RM1L-p	0,00170068
2221X1	COM4	S1L-l	0,00170068
2221X1	COM4	S1L-p	0,003401361
2221X1	COM4	W1-p	0,00170068
2221X1	COM8	C2L-p	0,00170068
2221X1	COM8	S1L-l	0,003401361
2221X1	COM8	S1L-p	0,00170068
2221X1	COM8	W1-p	0,00170068
2221X1	EDU1	RM1L-p	0,00170068
2221X1	GOV1	C1L-l	0,00170068
2221X1	GOV1	S1L-l	0,00170068
2221X1	GOV1	URML-p	0,00170068
2221X1	GOV2	S1L-l	0,00170068
2221X1	RES1	W1-l	0,263605442
2221X1	RES1	W1-m	0,00170068
2221X1	RES1	W1-p	0,103741497
2221X1	RES2	W1-p	0,00170068
2221X1	RES3A	URML-p	0,008503401
2221X1	RES3A	W1-l	0,011904762
2221X1	RES3A	W1-p	0,039115646
2221X1	RES3B	C1L-l	0,013605442
2221X1	RES3B	URML-p	0,032312925
2221X1	RES3B	W1-p	0,105442177
2221X1	RES3C	C1L-p	0,030612245
2221X1	RES3C	RM1L-p	0,006802721
2221X1	RES3C	URML-p	0,034013605
2221X1	RES3C	URMM-p	0,005102041
2221X1	RES3C	W1-l	0,168367347
2221X1	RES3C	W1-p	0,056122449
2221X1	RES3C	W2-p	0,013605442
2221X1	RES3D	RM1L-p	0,00170068
2221X1	RES3D	URMM-p	0,00170068
2221X1	RES3D	W2-p	0,003401361
2221X1	RES3E	W2-p	0,003401361
2221X1	RES3F	C2L-l	0,00170068

Table D-3: Distribution of building characteristics for IoX 4111X1 code

IoX Code	Occupancy Class	Building type-Code	Distribution
4111X1	AGR1	S1L-p	0.057971
4111X1	AGR1	S3-p	0.021739
4111X1	COM1	C2L-l	0.003623
4111X1	COM1	S1L-l	0.007246
4111X1	COM2	S1L-p	0.003623
4111X1	COM2	S3-m	0.003623
4111X1	COM2	S3-p	0.054348
4111X1	COM2	W1-p	0.007246
4111X1	COM3	S1L-l	0.007246
4111X1	COM4	C2L-l	0.003623
4111X1	COM4	S3-l	0.231884
4111X1	COM8	C1L-l	0.007246
4111X1	COM8	S1L-l	0.01087
4111X1	EDU1	S1L-l	0.003623
4111X1	GOV1	C2L-l	0.003623
4111X1	GOV1	C2L-m	0.028986
4111X1	GOV1	S1L-p	0.014493
4111X1	GOV2	C2L-m	0.003623
4111X1	GOV2	S1L-m	0.003623
4111X1	GOV2	S4L-m	0.003623
4111X1	IND1	C2L-p	0.003623
4111X1	IND1	S1L-l	0.021739
4111X1	IND1	S1L-p	0.014493
4111X1	IND2	S1L-p	0.003623
4111X1	IND2	S3-m	0.003623
4111X1	IND6	S1L-l	0.003623
4111X1	REL1	S1L-l	0.01087
4111X1	REL1	W2-l	0.007246
4111X1	RES1	W1-l	0.126812
4111X1	RES1	W1-p	0.318841
4111X1	RES6	C2L-m	0.003623

APPENDIX E: REPLACEMENT COST ACCORDING TO OCCUPANCY CLASS AND IOX CODE (COMMUNITY SIZE)

Table E-1: Residential buildings

IoX Code	RES1	RES2	RES3A	RES3B	RES3C	RES3D	RES3E	RES3F	RES4	RES5	RES6
1XXXXX	111.31	64.97	86.08	86.08	86.08	86.08	112.96	112.96	159.99	125.39	157.55
2XXXXX	111.31	64.97	86.08	86.08	86.08	86.08	112.96	112.96	159.99	125.39	157.55
3XXXXX	111.31	64.97	86.08	86.08	86.08	86.08	112.96	112.96	159.99	125.39	157.55
4XXXXX	111.31	64.97	86.08	86.08	86.08	86.08	112.96	112.96	159.99	125.39	157.55

Table E-2: Commercial buildings

IoX Code	COM1	COM2	COM3	COM4	COM5	COM6	COM7	COM8	COM9	COM10
1XXXXX	128.66	58.58	65.83	129.78	195.42	250.23	163.35	136.77	121.85	63.16
2XXXXX	128.66	58.58	65.83	129.78	195.42	250.23	163.35	136.77	121.85	63.16
3XXXXX	128.66	58.58	65.83	129.78	195.42	250.23	163.35	136.77	121.85	63.16
4XXXXX	128.66	58.58	65.83	129.78	195.42	250.23	163.35	136.77	121.85	63.16

Table E-3: Industrial buildings

IoX Code	IND1	IND2	IND3	IND4	IND5	IND6
1XXXXX	134.6	58.58	211.98	211.98	211.98	58.58
2XXXXX	134.6	58.58	211.98	211.98	211.98	58.58
3XXXXX	134.6	58.58	211.98	211.98	211.98	58.58
4XXXXX	134.6	58.58	211.98	211.98	211.98	58.58

Table E-4: Agriculture, educational, government and religion buildings

IoX Code	AGR1	EDU1	EDU2	GOV1	GOV2	REL1
1XXXXX	58.58	161.91	176.29	153.4	160.68	157.62
2XXXXX	58.58	161.91	176.29	153.4	160.68	157.62
3XXXXX	58.58	161.91	176.29	153.4	160.68	157.62
4XXXXX	58.58	161.91	176.29	153.4	160.68	157.62

APPENDIX F: SKEW NORMAL FUNCTION TO CALCULATE THE DISTANCE WEIGHTS FROM EACH CENSUS UNIT TO ALL OTHER CENSUS

Use the skew-normal distribution for weigthing minutes travel time and distributing travel population to each of the census units.

```
# Skewnormal functions
# Create skewnormal function
erf = function(x) {
  2 * pnorm(x * sqrt(2)) - 1
}

pdfsk = function(x) {
  return ((1 / sqrt(2 * pi)) * exp(-x ^ 2 / 2))
}

cdfsk = function (x) {
  return ((1 + erf(x / sqrt(2))) / 2)
}

skew = function(x,
                e = 0,
                w = 1,
                a = 0)
                {
                t = (x - e) / w
                return (2 / w * pdfsk(t) * cdfsk(a * t))
                }


# Read in network travel data with CTUID - CTUID and Total_Minutes from netwo
rk analysis in arcgis
#ottawa4.csv, torontodists3.csv,barrydists.csv,belldists.csv,drumdists.csv,gr
andists.csv,hamiltondists.csv,kingstondists.csv,montrealdists.csv
#oshawadists.csv,gatineaudists.csv,quebecdists.csv,stjeandists.csv,sherbdists
.csv,troisdists.csv,peterdists.csv
x=read.csv("e:/att/quebecdists.csv",colClasses = c(rep("character",6),rep("nu
meric",2)))
x=x[,-c(1,3,4,5,6,8)]
# Split from-to centroids, probably want to u se block population weighted ce
ntroids for final run
library(stringr)
library(stringi)

## Warning: package 'stringi' was built under R version 3.3.3

splitctuids=str_split_fixed(x$Name,"-",2)
CTUID = stri_trim_both(splitctuids[,1])
CTUIDTO = stri_trim_both(splitctuids[,2])
```

```
x$CTUID=CTUID
x$CTUIDTO=CTUIDTO

# Read in census data containing  CTUID POPCOM MEDCOM  POP for all CTS in stu
dy                    area                    from                    Census
pp = read.csv("e:/att/censusdatall.csv",colClasses = c("character","numeric",
"numeric","numeric"))

#        Only        take        the        POPCOM        and        MEDCOM
pp=pp[,1:3]

# Census data full population for each CT for all CTS in study area from Cens
us
allpop = read.csv("e:/att/allpop.csv",colClasses = c("character","numeric"))



#   Merge   the   POPCOM   MEDCOM   from   pp   with   POP   from   allpop
pp=merge(allpop,pp,by="CTUID",all.x=TRUE)


# Read in kate's data with commercial distribution plus census populations fr
om            subset            from            Mike_Daytime_population_mode
#ott.csv,tor.csv,bar.csv,bell.csv,drum.csv,grand.csv,hamilton.csv,kingston.cs
v,montreal.csv,oshawa.csv,gatineau.csv,quebec.csv,stjean.csv,sherb.csv,trois.
csv
oo = read.csv("e:/att/quebec.csv",colClasses = c("character",rep("numeric",28
)))


# Join by CTUID to pp table so that the POP field is available with the MEDCO
M                                  and                                  POPCOM
pp=merge(pp,oo,by="CTUID",all.y=TRUE)

##############################
##             For             Montreal             remove             4620440
#pp                    =                    pp[-which(pp$CTUID=='4620440'),]
##############################

#      Make      any      NA      zero      for      commute      time      m
pp$MEDCOM[which(is.na(pp$MEDCOM))]=0

##   Generate   weights   for   distance   based   on   median   travel   time

#           Get           list           of           CT           Identifiers
ctuids=unique(pp$CTUID)


#               Skewnormal               distribution               parameters
```

```
w = 8 # scale
a = -4 # drop rate on right side of distribtuion
o = w/2.4444444444# offset for center of distribution


outv = vector("numeric")
outk=vector("numeric")
ssetout=vector("numeric")
outw=vector("numeric")
for(i in 1:length(ctuids)) {
  #print(i)
  subs = which(x$CTUID == pp$CTUID[i])
  sset = x[subs,]
  e = pp$MEDCOM[i] + o # location+2
  k = sset$Total_Minutes # minutes to each CT from current CT
  # Get weghts
  weights = skew(k, e, w, a)
  wz = (weights - min(weights,na.rm = T)) / (max(weights,na.rm=T) - min(weigh
ts,na.rm=T))
  outw=c(outw,wz/sum(wz))
  wz = wz / sum(wz) * pp$POPCOM[i]
  outv = c(outv, wz)
  outk=c(outk,k)
  ssetout=rbind(ssetout,sset)
}

ssetout$pop=outv
ssetout$k=outk
ssetout$w=outw


# Get distance weihts to commercial center(s) where the distance is weighted a
ccording to the average
# distance of each CT from the center of the city using zonal statistics.
# tr = read.csv("e:/att/dist.csv",colClasses = c("character","numeric"))
# names(tr)=c("CTUID","DIST")
# tr$DIST=scalex(tr$DIST)


# Scale function 0 to 1
scalex = function(x){
  px= ((x - min(x)) / (max(x) - min(x)))
  return(px/sum(px))
}

# Multiply each CTs distance weights to other CTs by the commercial weights i
n the other CTs to redistribute the commerical population to other cts accord
ing to their median travel time and relative intensity of commercial activity
.  This uses the median travel weights to ajust the intensity of commercial a
ctivity weights towards the median travel time and also weights all commercia
```

```
l intensity weights zero that are zero in the median travel skewnormal.
mgd = data.frame()
for (i in 1:length(ctuids)) {
  m = merge(ssetout[which(ssetout$CTUID == ctuids[i]),], pp, by.x = "CTUIDTO"
, by.y = "CTUID")
  #q = merge(m,tr,by.x = "CTUIDTO", by.y = "CTUID")
  # # # make weight for curretn ctuid proportional to the total
  wframe = data.frame(rep(data.frame(m[,7]),length(names(m[, 11:23]))))
  wframe[i,]=m[i,11:23]
  mframe=as.data.frame(lapply(wframe, scalex))

  ## Multiply by distance weigths plus other weights
  #wts=q$DIST*(mframe*m[,11:23])

  ## Multiply by disance weights only
  wts=mframe*m[,11:23]

  wts = as.data.frame(lapply(wts, scalex), col.names = paste(names(m[,11:23])
, "p", sep =
  ""))
  wts = data.frame(mapply("*", wts, pp[pp$CTUID == m$CTUID[i], 18:30]))
  m = cbind(m, wts)
  mgd = rbind(mgd, m)
  #print(i)
}

#tr = read.csv("e:/att/dist.csv",colClasses = c("character","numeric"))
#names(tr)=c("CTUID","DIST")
#tr$DIST=scalex(tr$DIST)

# show a plot of the medain travel time weights, Original
par(mfrow=c(3,1))
idx=112 # A census tract
subs=which(mgd$CTUID==ctuids[idx])
sset = mgd[subs,]
plot(sset$Total_Minutes,sset$w, xlab="Mean travel time for CT", ylab="Time we
ight")
abline(v=pp[pp$CTUID==ctuids[idx],4])
plot(sset$Total_Minutes,sset$COM1_distr,xlab="Mean travel time for CT",ylab="
COM1 relative weights",col=2)
tot = sset$w*sset$COM1_distr
tot = ((tot-min(tot))/(max(tot)-min(tot)))
tot =tot/sum(tot)
plot(sset$Total_Minutes,sset$COM1_distrp,xlab="Mean travel time for CT",ylab=
"Final time weights")
```

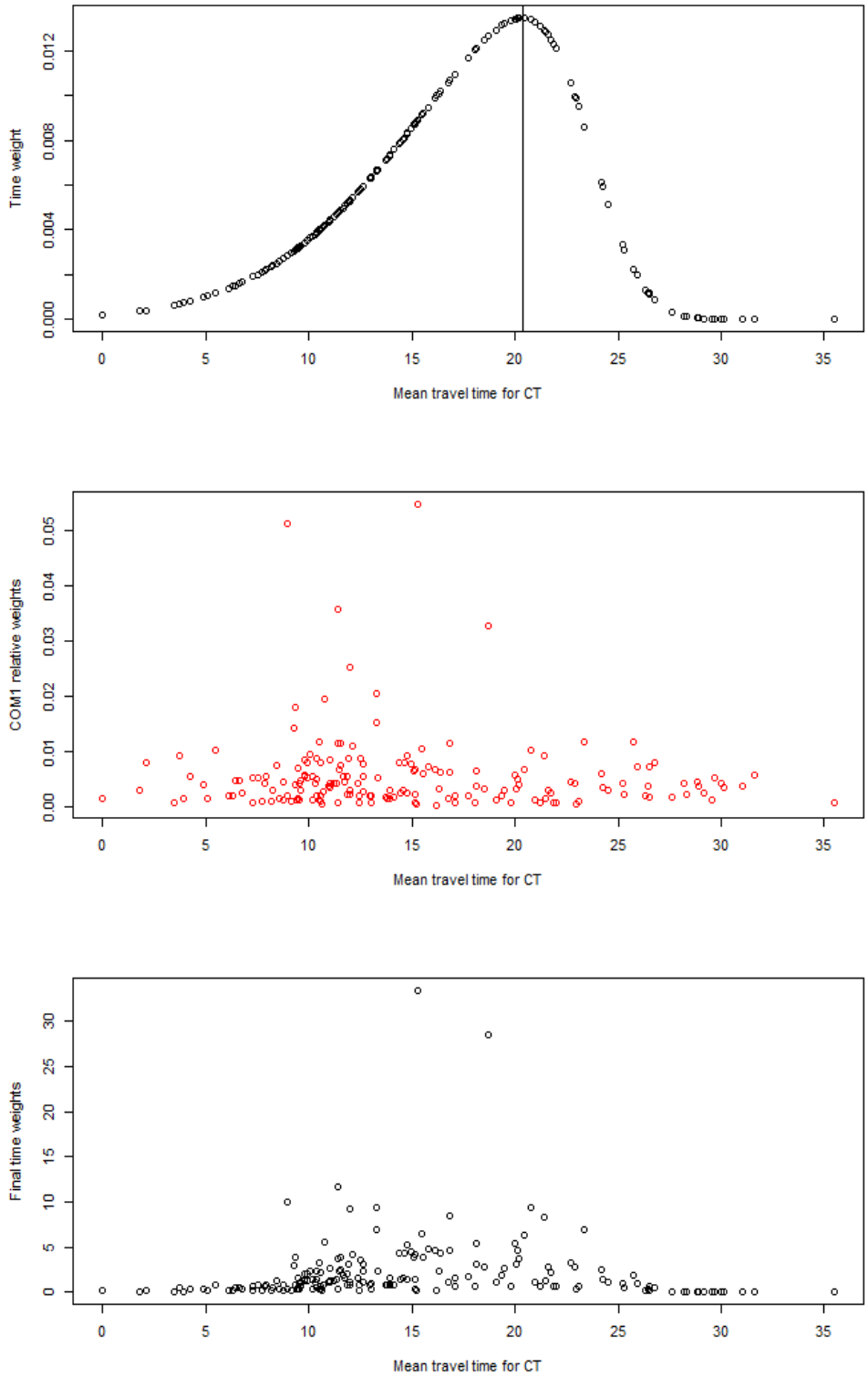


Figure E-1: Time weights, commercial weights and total time weight plots.

```
svec=vector("numeric")
pres=vector("numeric")
opop=vector("numeric")

nm                              =                              names(mgd)[39:51]
for(k                   in                   1:length(nm))                    {
  for            (i            in            1:length(pp$CTUID))             {
  svec = c(svec, sum(mgd[[nm[k]]][which(mgd$CTUIDTO == pp$CTUID[i])], na.rm =
  T))
  #pres=c(pres,pp$POP[pp$CTUID==ctuids[i]]-pp$POPCOM[pp$CTUID==ctuids[i]])
  #opop=c(opop,pp$POP[pp$CTUID==ctuids[i]])

  }
  pp[[nm[k]]]=svec

  svec=vector("numeric")
}

pp$inct=pp$POP-rowSums(pp[,18:30])
pp$tpop=pp$inct+rowSums(pp[,33:45])

write.csv(pp,"e:/att/final/PETERB.csv")

flist                                                =list.files("e:/att/final")
dfout=data.frame()
for(i                           in                           1:length(flist)){
  x            =            read.csv(paste("e:/att/final/",flist[i],sep=""))
  dfout=rbind(dfout,x)

}
write.csv(dfout,"e:/att/final/finalmodel.csv")
```

```
x=read.csv("e:/att/torontodists.csv",colClasses                              =
c(rep("character",6),rep("numeric",2)))
x=x[,-c(1,3,4,5,6,8)]

#       Long       tailed       parameters       for       Log-Normal       curve
C                                      =                                        1
s                                                                            =.5
m                              =                              log(x$Total_Minutes)
routesf=C*exp(-((log(x$Total_Minutes)-m)^2)/(2*(s^2)))
x$weights=routesf


#                                   show                                  weights
ctuids=unique(x$CTUID)
subs=which(x$CTUID==ctuids[55])
sset                                   =                                x[subs,]
plot(sset$Total_Minutes,sset$weights,type="l")
```

```
# Plot weights for one ct to see distribution
subs=which(ssetout$CTUID==ctuids[1])
sset = ssetout[subs,]
plot(sset$Total_Minutes, sset$pop)
abline(v=pp[pp$CTUID
```