Documentation for the
2009 Canadian Composite Seismicity Catalogue
(CCSC09)

Final Report

By

Miguel Macias-Carrasco, Azadeh Fereydouni, Katsuichiro Goda, Gail Atkinson

September 2010

CONTENTS

1. Introduction

This document describes the Canadian Composite Seismicity Catalogue (CCSC09) developed under the NSERC Strategic Network Grant "Reducing Urban Seismic Risk". This report describes the procedures used to analyze the data, the input and output files, and each one of the variables contained in the composite catalogue as well as the references to the diverse sources of information.

The primary data for the development of the CCSC09 catalogue was the SHEEF08 (seismic hazard earthquake epicentre file, 2008) catalogue, which is an updated version of the SHEEF catalogue used by Adams and Halchuk (2003) as the basis for the fourth generation seismic hazard maps of Canada. The SHEEF08 catalogue covers all the Canadian territory, including earthquakes at a minimum distance of 200 km for the southwestern and northwestern boundaries and at 300 km for the southeastern boundary. Data from a total of 26217 earthquakes, ranging in magnitude (different types) from 2.5 to 9.0 that occurred between 1627 and 2008 constitute the catalogue (West and East regions); details on the variables extracted from this database are given in Chapter 3 (Input / output files). The secondary sources of information, used in a selective way to enhance the data already contained in the SHEEF08 catalogue, were as follows: (1) The Ma and Atkinson (2006) catalogue for eastern Canada (MASE) was developed for eastern North America and describes seismicity that occurred from 1980 to 2006 in the northern Ontario and the western Quebec seismic zones; the reported magnitudes are Nuttli magnitudes (M_N). (2) The Petersen et al. (2008) catalogues for western North America (PETW) and for eastern North America (PETE) were developed for the 2008 update of the U.S. national seismic hazard maps. The PETE catalogue lists about 3350 earthquakes, most of them reported as m_{bLg}, from 1700 through 2006, with magnitudes $\geq$ 3.0. The PETW catalogue includes data from about 3260 earthquakes that occurred between 1850 and 2006 with magnitudes $\geq$ 4.0, where M_W and M_L dominate the different reported magnitudes. (3) The catalogue obtained from the Advanced National Seismic System at http://earthquakes.usgs.gov/research/monitoring/anss (ANSS) was used to extend the period from U.S. databases through 2008; thus the downloaded catalogue included information for the period 2007-2008, between the limits: 35° to 60° in latitude, -135° to -60° in longitude, 0 km to 99 km in depth, and from 2.5 to 9.0 in magnitude (different types). (4) The catalogue developed by the Canadian Pacific Geosciences Centre at Sidney, BC (PGCW) and the Canadian earthquake epicentre catalogue developed by the Geological Survey of Canada at Ottawa (GSCE) contain data through 1991, and mainly were used to enhance the SHEEF08 database. (5) The moment tensor solution catalogue published by Ristau (2004) for western Canada (RIST) was used to provide high-quality moment magnitude (M_W) values for a suite of western earthquakes that occurred in the period 1976-2004.

The variables that are generated from the input files and organized in the final output files are, for each event:

- Date/time of occurrence composed of five elementary variables (1-5): year, month, day, hour, and minute.
- Location (6-7): latitude and longitude.
- Reported magnitudes (8-15): Mb, M_N, M_L, M_S, Mc, M_W, MM [*], and MZ [*] (the last two magnitude types are used for unknown magnitude types).
- Preferred magnitude (16-17) [*]: value and type.
- Depth (18-19) [*]: value and designation.
- Flags for (20-21) [**]: data source (catalogue) and seismic source zone
- Conversion factor (22) [**]: used for M_W calculations.
- Assigned moment magnitude (23) [**]: M_W.

[*] – Explained in Chapter 3 (Input / output files)

[**] – Explained in Section 4 of Chapter 2 (Procedures)

The main program designed to process the input information is written in MATLAB language, named **ccsc09.m** and generates two final composite catalogues, one for West Canada (data located at longitudes < -110º) and the other one for East Canada (data located at longitudes ≥ -110º). These two output files can be tested against the information available from Earthquakes Canada at http://earthquakescanada.nrcan.gc.ca/stndon/NEDB-BNDS/bull-eng.php, which provides seismicity catalogues for periods since 1985 to date for the whole territory. A second computer program (**plotccsc09.m**) was developed to visualize this information, which only requires the output files (composite catalogues for both regions) from the main program and the catalogue files provided by Earthquakes Canada.

2. Procedures

The primary source of information is the SHEEF08 catalogue. This data set is cleaned of duplicates, modified and / or enhanced with magnitude values, and new information is added through four steps as described below, before the output catalogues are generated. The fundamental idea is to preserve information of the primary data and to add, in the subsequent steps, reported magnitude values on common events from secondary catalogues and / or add the complete set of variables values for non-common events. Where available, moment magnitudes from the Ristau catalogue are used in preference to those in the other catalogues. Figure 1 shows a general representation of the procedure and mainly gives the order in which information was processed.

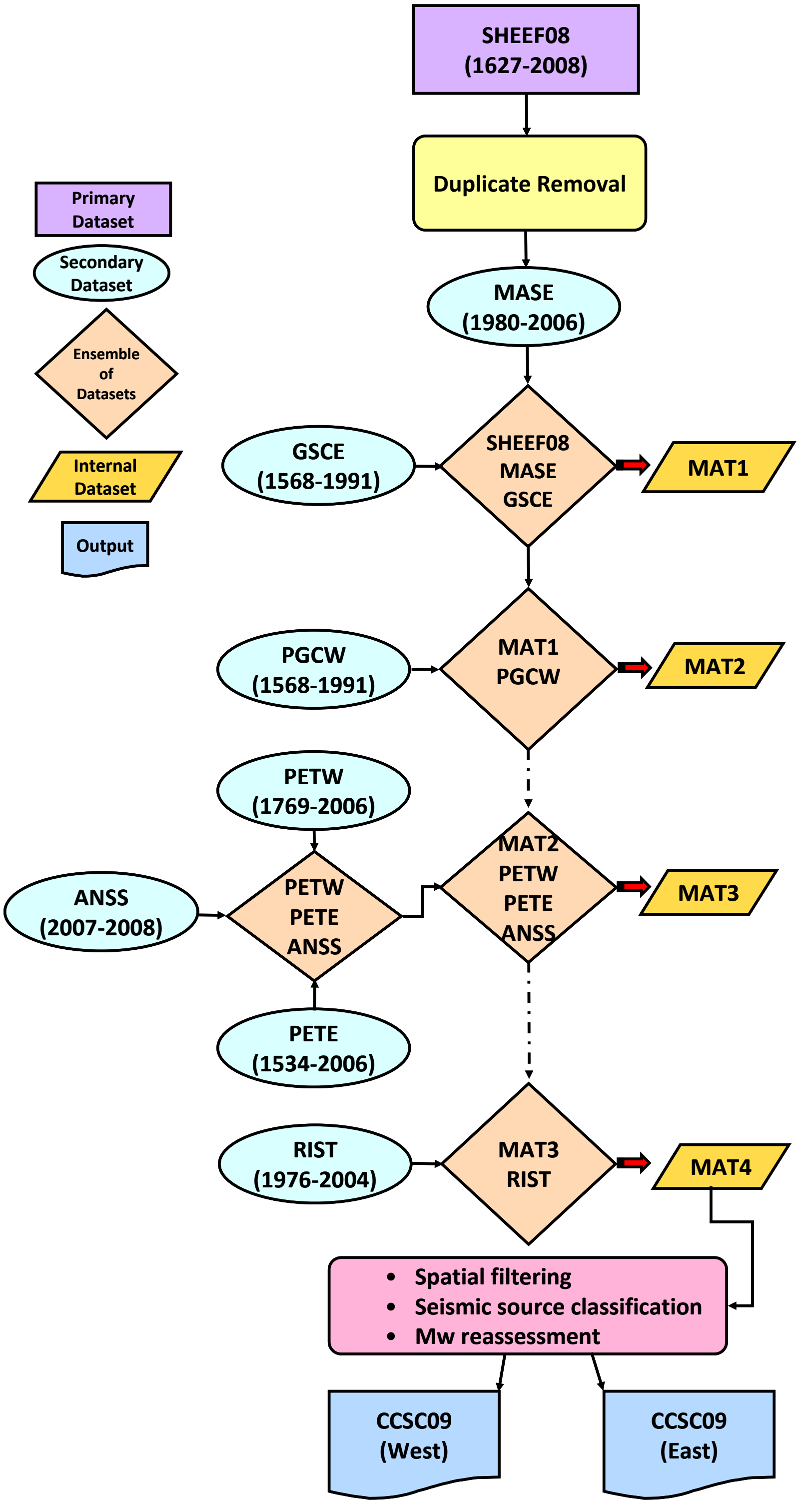


Figure 1 – Schematic representation of the process used to develop the Canadian Composite Seismicity Catalogue (CCSC09). MAT1, MAT2, MAT3, and MAT4 are the variables in the **ccsc09.m** program that contain the output data ensemble from Sections 1 to 4, respectively.

Section 1

The SHEEF catalogue is primarily based on the earthquake epicenter and magnitude information in the Geological Survey of Canada (GSC)'s Canadian Earthquake Epicenter File (CEEF), which has been maintained and augmented for about a century (Halchuk, 2009). Through the development of the SHEEF catalogue, the primary source has been augmented and extended by inclusion of events from various earthquake databases. The procedure of updating and integration of the solutions have been resulted in multiple solutions for a single event (S. Halchuk, personal communication, 2010).

The SHEEF catalogue provides the main source of information for the CCSC09, so it is critical to clean the catalogue from replicated information to the level that is practically achievable. Considerable work was done to discriminate and eliminate the duplicate solutions in the SHEEF data. The SHEEF catalogue is available online through the GSC's Open File Report 6208 (Halchuk, 2009). The report is accompanied by the full solution lines of the database and the lists of aftershocks removed from the main database for hazard calculation purposes. It also includes the list of events identified as duplicates by GSC in late 2009. Our initial investigation uncovered that some duplicates still exist in the full solutions, and more comprehensive investigation was required to clean the SHEEF data of replicated information. So before we began to re-check the data, we combined the following files accompanying the Open File 6208 into one single catalogue, SHEEF2008, simplified the format, and retained just the essentials.

- SHEEF –list of solution lines (used in the fourth generation hazard models).
- byammartin_rem.txt – list of aftershocks removed form the Byam-Martin, NU region.
- miramichi_rem.txt – list of aftershocks removed from the Miramichi, NB region.
- Nahanni_rem.txt – list of aftershocks removed from the Nahanni, NT region.
- duplicates_to_1991.ceef.txt - list of 26 events identified as duplicates and removed from the SHEEF catalogue in late 2009.
- duplicates_1992_2008.ceef.txt – list of 104 events identified as duplicates in late 2009.

- SHEEF199192to2008.txt – a SHEEF compatible list of additional, more recent solutions from the end of the original SHEEF to the end of 2008, (not used in the fourth generation hazard models).

A semi-automatic procedure was applied to identify the duplicate events in SHEEF2008. A set of possible candidate pairs (potential duplicates) was extracted based on the occurrence of the events in the same time up to the minute. Each pair was examined individually and identified as duplicates, unless there were legitimate reasons to consider it as two separate events, such as the distance between events, or their occurrence within an aftershock sequence. Since the accuracy of earthquake location varies significantly over time and space in the SHEEF data (Halchuk, 2009), it was impossible to set certain criteria for determining the reasonable distance between two separate earthquakes, so it was decided subjectively through the investigation.

In SHEEF2008, a total of 228 “possible” duplicate pairs was identified. Of these, 227 pairs classified as “real” duplicates. The number of real duplicates increases noticeably after 1990/91. During that period of time the information from the GSC’s Sidney office in the west has been routinely incorporated into the database in Ottawa office. Since the two offices have been using different packages to locate earthquakes, the integration of the solutions has resulted in several duplicates in SHEEF data after 1990/91.

The cleaned SHEEF08 and MASE catalogues are then merged as follows. The month and day values from SHEEF08 data are set equal to one if these data are missing (equal zero). There were only 7 cases in the SHEEF08 data set, where the month and day were missing. Then, SHEEF08 and MASE data sets are merged by identifying common events, and replacing just the location (coordinates) and depth information (value and designation) in the SHEEF08 data set by the corresponding information from MASE data set on these events. The procedure used in this section, as well as in the other sections, to identify common or repeated events is based on time differences: a) the date and time (up to the minute) of occurrence of each event is converted to days and fraction of days (from an specific reference time, in this case [year, month, day, hour, minute] = 0), b) for a specific dataset (repeated event case) or for an ensemble of two data sets (common event case), the events are sorted in time (increasing order), c) time differences are calculated between contiguous rows of the time column of the data set, and d) an index is identified where these differences are equal to zero, and are related to repeated or common events. For events in the MASE data that are not reported in the SHEEF08 catalogue, all the information (date, location, magnitude, depth, and flags) is added to the SHEEF08-MASE ensemble. Note that the MASE catalogue was developed for eastern North America, focused in the northern Ontario and the western Quebec seismic zones and the reported magnitudes are Nuttli magnitudes (M_N). The output data set at this stage is then updated with reported magnitudes from the GSCE catalogue only for common events; the preferred magnitude in the SHEEF08-MASE data is not altered. This step adds magnitudes not contained in the SHEEF08. The output variables at this stage in the computer program are referred as mat1, tpo1, mty1, dty1; mat1 is a matrix that contains all the variables, the variables giving the magnitude type for the preferred magnitude and depth designation, are contained in mty1 and dty1, respectively; tpo1 is the time variable in days and fraction of days for events in mat1 and is used for time comparisons. Note that an explanation for the different depth designations found in each one of the input files is given in Chapter 3.

Section 2

In this section, the PGCW catalogue is read, filtered in space (longitude ≤ -110°, latitude ≤ 60°), and all the necessary variables are generated. It is emphasized that because this catalogue contains, for many events, information about different magnitude types (not only for the preferred magnitude), it is a remarkable source of information on reported magnitudes for West Canada(note: the GSCE catalogue is for East Canada). At this point of the process, the preferred magnitude information in Section 1 is preserved and only new reported magnitude information for common events between PGCW and mat1 is added. The non-common event information is also added but for the whole set of variables. The output from this section is the ensemble SHEEF08-MASE-GSCE-PGCW, and the variables are referred as mat2, tpo2, mty2, and dty2 in the computer program.

Section 3

The PETW, PETE, and ANSS catalogues are read, and repeated events were removed based on common time of occurrence. These catalogues provide events given in the U.S. databases. The PETW and PETE data cover up to 2006, and the ANSS catalogue is defined for the period 2007-2008; the data were added to the PETE-PETW composite. The reported magnitude values in mat2 are enhanced with the corresponding information in the ensemble PETW-PETE-ANSS for common events. For non-common events, information from this data ensemble is added to the total ensemble. Again, as in previous sections, the preferred magnitude information in mat2 is preserved. The output from this section is the ensemble SHEEF08-MASE-GSCE-PGCW-PETW-PETE-ANSS and the variables with this information are referred as mat3, tpo3, mty3, and dty3 in the computer program.

Section 4

In this section, the RIST catalogue is used to add M_W information to the output data mat3. The RIST catalogue for this work was integrated with the earthquake date / time, location and M_W data contained in Tables A1 and A2 in Ristau (2004). Because of the good quality of this information, reported and preferred magnitude data in mat3 were replaced by the corresponding (common events) data in the RIST catalogue. The final ensemble is composed by the SHEEF08-MASE-GSCE-PGCW-PETW-PETE-ANSS-RIST catalogues and is captured in the internal variables mat4, tpo4, mty4, and dty4. Then, the mat4 data are filtered in space ($-110° \leq$ longitude $\leq -45°$, $35° \leq$ latitude $\leq 80°$) for eastern region and ($-160° \leq$ longitude $< -110°$, $43.5° \leq$ latitude $\leq 75°$) for western region, to assure convenient spatial coverage of the composite catalogues for further analyses.

Each one of the events in the final ensemble is referred to its original database by using the source catalogue flag **cf**; this variable takes the values: 0, 1, 2, 3, 4, 5, or 6 for SHEEF08, MASE, PETW, PETE, ANSS, PGCW, or RIST, respectively.

Seismic source classification

Because moment magnitude conversion from different magnitude types may depend on the location and seismotectonic setting, all the events in mat4 were classified according to specific seismic source zones. The reference source zones, used in this work for M_W-M (different types) comparison purposes, are those contained in the **R** model of seismicity published by Adams and Halchuk (2003) for the whole territory. It is assumed that the seismogenic features of each zone may be reflected in magnitude estimations and consequently in magnitude residuals, i.e., M_W-M (different types) between zones. This classification is captured in the seismic source zone flag **zf**, included in the output file; the relation between the source flag and the code (abbreviation) given by Adams and Halchuk (2003) for each zone is explained in Table 1, and Figure 2 shows the spatial distribution of each seismic source zone. Note that for the following code and for the CCSC09 catalogue, East Canada is defined by (-45° ≥ longitude ≥ -110°; 35° ≤ latitude ≤ 80°), and West Canada is (-110° > longitude ≥ -160°; 43.5° ≤ latitude ≤ 75°).

Table 1 – Flag numbers and codes for the seismic source zones used in the calculation of M_W values for events in the CCSC09 catalogue. Source zones are as in Adams and Halchuk (2003).

Region	Zone flag	Code	Description
West Canada	**2**	**ALC**	ALASKA COASTAL
	3	**ALI**	ALASKA INLAND
	4	**BFT**	BEAUFORT SEA
	5	**BRO**	BROOKS PENINSULA
	7	**CASR**	CASCADE MOUNTAINS (shallow)
	8	**CST**	COASTAL
	9	**DENR**	DENALI
	10	**EXP**	EXPLORER PLATE BENDING
	11	**FHL**	FLATHEAD LAKE
	13	**GOA**	GULF OF ALASKA
	14	**GSP**	GEORGIA STRAIT/PUGET SOUND(deep)
	15	**HECR**	HECATE STRAIT
	16	**JDFF**	JUAN DE FUCA PLATE BENDING OFFSHORE
	17	**JDFN**	JUAN DE FUCA PLATE BENDING ONSHORE(deep)
	19	**MMB**	MACKENZIE MOUNTAINS
	20	**NBC**	NORTHERN BC
	21	**NOFR**	NOOTKA FAULT
	22	**NYK**	NORTHERN YUKON
	23	**OFS**	OFFSHORE
	25	**QCFR**	QUEEN CHARLOTTE FAULT
	26	**RMN**	RICHARDSON MTNS-NORTH
	27	**RMS**	RICHARDSON MTNS-SOUTH
	28	**ROC**	ROCKY MOUNTAIN F and T BELT
	29	**SBC**	SOUTHERN BC
	34	**SOY**	SOUTHERN YUKON
	32	**YAK**	YAKATAT COLLISION
	101	**WNE**	NORTH-EAST of WEST CANADA
	102	**WE**	EAST of WEST CANADA
	103	**WSE**	SOUTH-EAST of WEST CANADA
	104	**WS1**	SOUTH-1 of WEST CANADA
	105	**WS2**	SOUTH-2 of WEST CANADA
	106	**WW**	WEST of WEST CANADA
	107	**WNW**	NORTH-WEST of WEST CANADA
	1	**Unknown**	PROBABLY THE EVENTS ON THE BOUNDARIES OF ZONES

East Canada	**149**	**ACM**	ARTIC CONTINENTAL MARGIN
	52	**ADR**	NORTHERN ADIRONDACKS
	53	**AOBR**	ATLANTIC OFFSHORE BACKGROUND
	54	**BFI**	BAFFIN ISLAND
	55	**BOU**	BOOTHIA UNGAVA
	57	**CMF**	COASTAL MAINE FUNDY
	58	**COC**	COCHRANE
	59	**DIB**	DEVON ISLAND BACKGROUND
	60	**ECM**	EASTERN CONTINENTAL MARGIN
	61	**GAT**	GATINEAU
	63	**GLD**	GREENLAND
	64	**IRB**	IAPETAN RIFT BACKGROUND
	65	**IRM**	IAPETAN RIFT MARGIN
	66	**JMS**	JAMES BAY
	67	**LAB**	SOUTHERN LABRADOR
	69	**LBR**	LABRADOR RIDGE
	70	**NAI**	NORTHERN APPALACHIANS INTERIOR
	71	**OBGR**	ONTARIO BACKGROUND R
	72	**SGL**	SOUTHERN GREAT LAKES
	149	**SVDR**	SVERDRUP BASIN R
	73	**WLB**	WILLISTON BASIN
	50	**Unknown**	
	51	**Unknown**	IN SOUTHERN BORDER OF CANADA
	0	**Unknown**	IN SOUTHERN BOUNDARY OF CATALOGUE IN THE U.S.

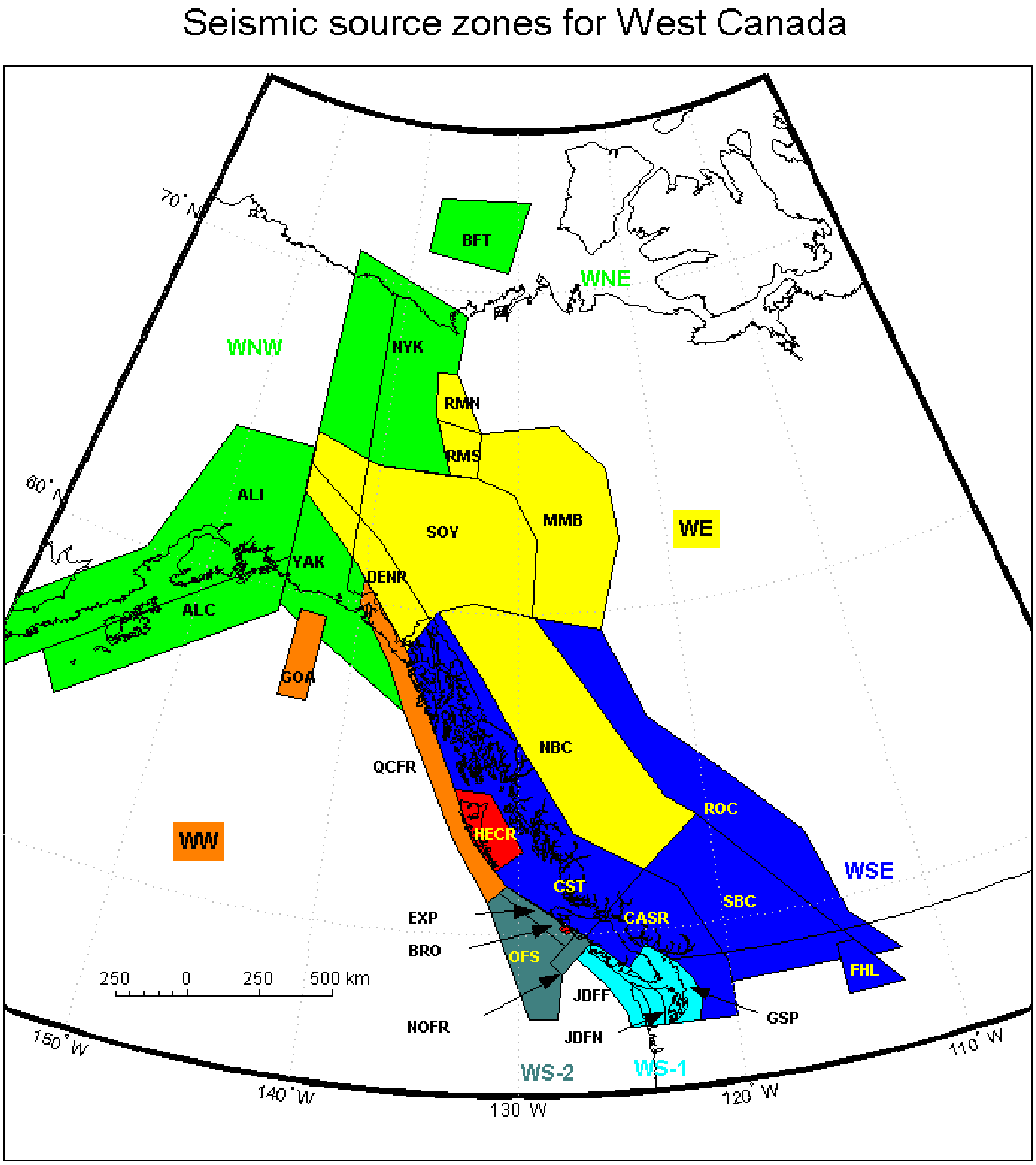


Figure 2 – West Canada seismic source zones (Adams and Halchuk, 2003) used to classify events included in the CCSC09 catalogue. Codes for zones are as in Table 1; colours identify sub-regions (used to assign M_W in the catalogue) that are described in Table 2. Source zones, WNE, WE, WSE, WS-1, WS-2, WW, and WNW, are complementary zones (not defined in Adams and Halchuk (2003)) used to cover the whole catalogue and are not completely coloured in this figure; their actual limits are given in the next section.

M_W for West Canada

In order to investigate M_W values for the composite catalogue, events with M_W-M (different types) pairs of reported magnitudes were grouped according to seismic zone and M type, plotted against pre-existing models of M_W-M (different types) for the regions, and analyzed to estimate factors to convert from: Mb, M_N, M_L, M_S, or Mc to M_W. The correction factors were calculated as the mean value of the residuals M_W-M for each combination of seismic zone and M type, under the assumption that gradient between M_W and any of the M types equals 1.0. This assumption was confirmed by plots for which the number of M_W-M pairs allowed delineation of trends. Figure A-1 (Appendix A) contains some examples of M_W-M data pairs, plotted against pre-existing models (described later), for the whole West region and for each M type. Under this assumption, M_W is obtained simply by adding the correction factor listed in the output catalogue to the corresponding reported preferred magnitude.

In the first stage of this analysis, plots and mean residuals of M_W-M data pairs were done for each one of the magnitude - source zone combinations (5 M types × 29 zones). Based on these preliminarily results, in about 50% of the cases it was possible to observe trends and confirm previous M_W conversion models for the region (mainly for the M_L case). However, most of these results were not robust enough to determine reliable correction factors for all the M types and sources zones, but were nonetheless useful to observe spatial trends of these residuals, group specific source zones, and form sub-regions for a similar analysis. Table 2 describes these sub-regions, their related seismic source zones and the possible event types that are more common in each sub-region. Note that the event type column is just given as a reference and is not used for M_W assignment in the composite catalogue. See Atkinson (2005) for a more detailed description of event types in the region.

Because some earthquakes in the composite catalogue may be located in zones not defined in the R model given by Adams and Halchuk (2003), seven complementary source zones were defined in this work to cover the whole catalogue, and were added to specific sub-regions according to their location and tectonic setting. Each one of the complementary zones was defined as the remaining area, out of the already defined source zones in Adams and Halchuk (2003), which is localized into a specific rectangle area. The rectangle areas are given, for each case, by the limits:

- WNE: [65° ≤ latitude ≤ 75°] and [-140° ≤ longitude < -110°]
- WE: [54° ≤ latitude < 65°] and [-130° ≤ longitude < -110°]
- WSE: [43.5° ≤ latitude ≤ 54°] and [-121.5° ≤ longitude < -110°]
- WS1: [43.5° ≤ latitude ≤ 47°] and [-124.9° ≤ longitude < -121.5°]
- WS2: [43.5° ≤ latitude < 50°] and [-130° ≤ longitude < -124.9°]
- WW: [43.5° ≤ latitude < 60°] and [-160° ≤ longitude < -130°]
- WNW: [60° ≤ latitude ≤ 75°] and [-160° ≤ longitude < -140°]

Table 2 – West Canada sub-regions, the corresponding seismic sources zones for M_W-M residual analysis and the event types that are expected in each sub-region. Colour in sub-region column identifies the corresponding sub-region in Figure 2.

Sub-region	Flag numbers for seismic source zones	Expected event type	Codes for seismic source zones
1 **red**	**5, 15**	**Transition, shallow crustal**	**BRO, HECR**
2 **heavy green**	**10, 21, 23, 105**	**Offshore, interface, transition**	**EXP, NOFR, OFS, WS2**
3 **orange**	**13, 25, 106**	**Offshore, transition**	**GOA, QCFR, WW**
4 **blue**	**7, 8, 11, 28, 29, 103**	**Shallow crustal, interface, continental**	**CASR, CST, FHL, ROC, SBC, WSE**
5 **yellow**	**9, 19, 20, 26, 27, 34, 102**	**Continental**	**DENR, MMB, NBC, RMN, RMS, SOY, WE**
6 **green**	**2, 3, 4, 22, 32,101, 107**	**Interface, inslab, continental**	**ALC, ALI, BFT, NYK, YAK, WNE, WNW**
7 **cyan**	**14, 16, 17, 104**	**Interface, inslab, transition**	**GSP, JDFF, JDFN, WS1**

Table 3 gives, for Mb, M_N, and M_L, the correction factors and the number of M_W-M data pairs used for each correction factor calculation. As the number of pairs for Mb is low, the M_W-Mb data pairs are combined in sub-regions # 2 and 3; these sub-regions cover the offshore events and thus the combination of the data is reasonable. The data pairs in sub-regions # 1, 4, 5, 6, and 7 are also combined together to increase the number of data pairs used to determine the correction factor. For M_N, the data in all the regions are combined to derive the M_W-M_N conversion factor based on a reliable number of data pairs. Since M_N is being used for crustal events only, the inclusion of offshore sub-regions is irrelevant (i.e., M_N values are not reported for offshore events). Because M_W-M_S plots do not follow a model with the slope = 1.0, these data pairs were fitted to a model that is shown in Appendix A (Figure A-2), and was used for M_S-M_W conversion. For the Mc case, due to the very low number of the M_W-Mc data pairs, the model proposed by Dewberry and Crosson (1995) [DC95] for the region was applied. Figures A-3, A-4 and A-5 in Appendix A compare data, from different sub-regions, versus M_W conversion models developed for western North America region by Ristau et al. (2003) [RRC03] and Cassidy et al. (2005) [CRR05] for M_L, Braunmiller and Nabelek (2002) [BN02] for Mb, and Dewberry and Crosson (1995) [DC95] for Mc. When a specific conversion model is applied (M_S and Mc cases), the correction factor is the difference between the calculated M_W and the corresponding reported magnitude. For Mb, M_N, and M_L, the correction factors in Table 3 are simply added to the reported magnitude.

Table 3 – West Canada M_W additive conversion factors for Mb, M_N, and M_L magnitude types and for seismic source sub-regions as defined in Table 2. Assigned M_W in the composite catalogue is then given by: M_W = M [Mb, M_N, or M_L] + conversion factor.

Sub-region	Mb		M_N		M_L	
	factor	# pairs	factor	# pairs	factor	# pairs
1	-0.06 ± 0.31	3	0.05 ± 0.29	0	0.57 ± 0.15	74
2	0.37 ± 0.32	49	0.05 ± 0.29	0	0.63 ± 0.15	1268
3	0.37 ± 0.32	12	0.05 ± 0.29	0	0.55 ± 0.21	251
4	-0.06 ± 0.31	6	0.05 ± 0.29	3	0.12 ± 0.37	36
5	-0.06 ± 0.31	3	0.05 ± 0.29	6	-0.18 ± 0.30	80
6	-0.06 ± 0.31	2	0.05 ± 0.29	1	0.03 ± 0.35	33
7	-0.06 ± 0.31	4	0.05 ± 0.29	0	0.60 ± 0.48	16

M_W for East Canada

Due to the lower number of M_W-M (different types) pairs per seismic source zone in eastern Canada, the assignment of M_W to the earthquakes in the eastern section of the composite catalogue followed a different approach, compared to the western region. For the entire East region, three different relationships between M_W and M (different types) were applied to assign M_W values. For M_N type, the Sonley and Atkinson (2005) model [SA05] was used not only because it was derived from this type of data but also because of the robustness of this model. For Mb, given the observed trends between M_W-Mb data pairs, the SA05 model was also assumed to apply (i.e., Mb = M_N). For M_L, a model was derived from the relationship reported by Kim (1998) [WK98] between M_L-M_N for Eastern North America (ENA). This author developed a M_L scale for ENA and explored the differences between M_L and m_{bLg}; his database is representative of eastern Canadian seismicity. The available data confirmed the appropriateness of the models described above, except for the M_S case. Therefore, a specific model was determined from the data for ENA (Figure A-2) and applied for conversion. For Mc, because it was not possible to observe trends from M_W-Mc data pairs, Mc ≈ M_L was assumed so that the model for M_L can be applied to Mc data. The equations used for M_W assignment for East Canada are as follows:

- $M_W = 1.03 \times M_N - 0.61$; Sonley and Atkinson (2005) model [SA05]
- $M_W = 1.03 \times Mb - 0.61$; Sonley and Atkinson (2005) model [SA05]
- $M_W = 1.03 \times M_L - 0.46$; derived from Kim (1998) model [WK98]
- $M_W = 0.67 \times M_S + 1.8$; model for M_S after linear fit to M_W-M_S data pairs
- $M_W = 1.03 \times Mc - 0.61$; Sonley and Atkinson (2005) model [SA05]

Note that the conversion factors given in the composite catalogue for the East region are the differences between the calculated M_W values (after model application) and the reported magnitudes for each particular case. As the different approach is used for the M_W assignment for the East region, the seismic source zone classification (Table 1) and consequently the map shown in Figure 3 were not relevant for this process but are included for consistency and reference.

Section 4, and consequently the whole process, ends with the M_W assignment for each event, and the generation and recovery of the two output files (for eastern and western Canada) that integrate the CCSC09 catalogue: **ccsc09west.txt** and **ccsc09east.txt**. The input files required for the program **ccsc09.m** are described in some extent in the next chapter as well as the output files.

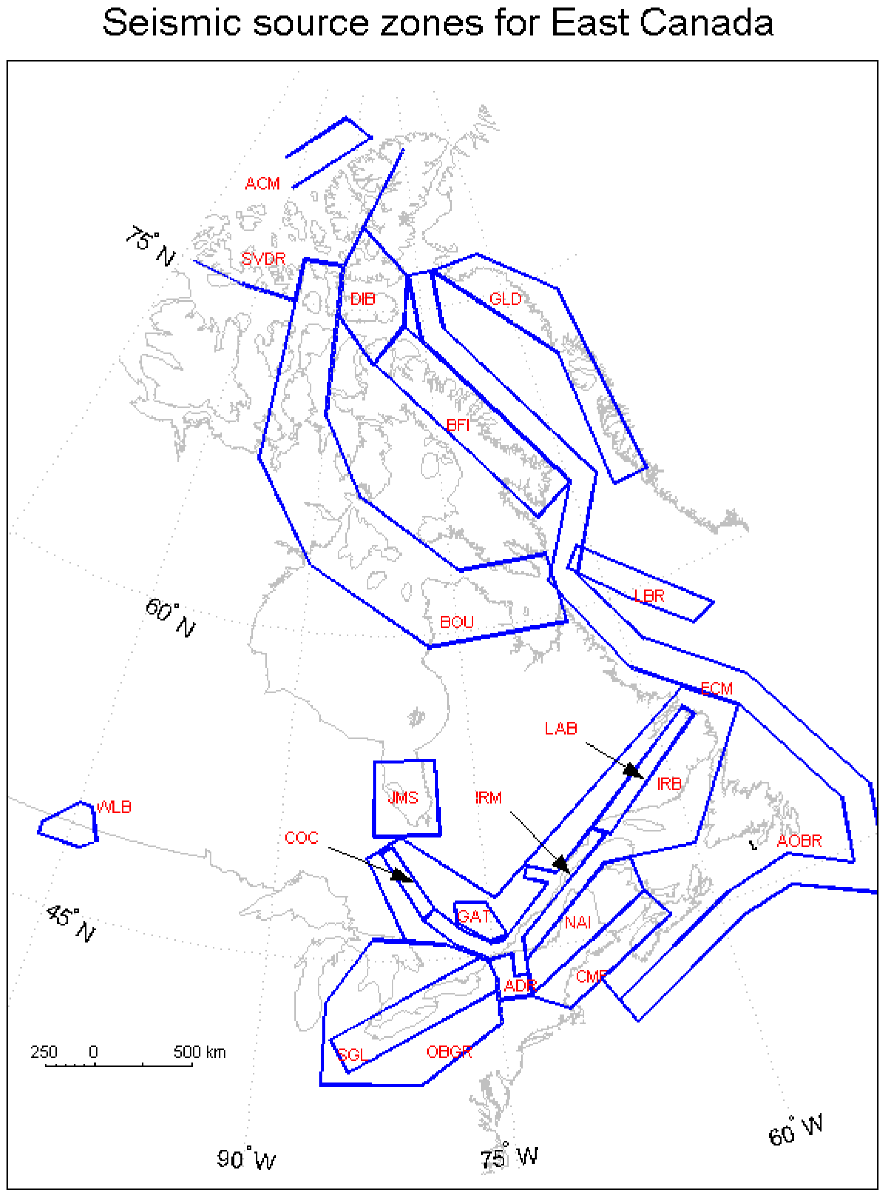


Figure 3 – Seismic source zones for East Canada (Adams and Halchuk, 2003). Codes for source zones (red fonts) are as described in Table 1. These source zones have no related conversion factors for M_W assignment.

Adoption of actual M_W values

When M_W is converted from other instrumental magnitudes, the associated uncertainty is generally high; both the uncertainty in the original magnitude as well as that of the conversion equations must be taken into account. Thus, the moment magnitudes obtained by other means in previous studies have been adopted in the composite catalogue. The preference has been given to the values determined by detailed earthquake source studies and moment tensor conversion or those converted from instrumental moments. The use of MMI data was the second choice, as felt data provide good indication of the overall event size, and are likely superior to instrumentally converted values, particularly in early events for which the stations coverage was poor and the magnitude determination methods were poorly documented. For the purpose of current study, two recent databases were considered: Atkinson and Boore (2006) [AB06] and Bent (2009) [B09], with preference given to AB06 except in cases where the B09 study provided new or additional information. Appendix B shows the adopted values and their sources. These values have replaced converted values in the M_W column of the catalogue. For these events, the conversion factors are simply zeros. As the moment magnitude is the most preferred magnitude in the seismological analysis, the preferred magnitude is also changed to M_W in these events.

3. Input / output files

Input files

Table 4 (below) gives comprehensive information about the necessary input files to run the program **ccsc09.m**. In the first part of the table, rows are for input files and columns for the different variables explored in each case (input file). The input file name column gives information on the file name extension, the catalogue name used in this work (related with a specific input file), and the section (procedure) where the input file is explored. All the *.txt, *.dat, and *.1991 files, are the original catalogues obtained from different sources, meanwhile the *.xls files were generated from their original txt versions or from published catalogues (e.g., RIST catalogue).

It is emphasized that for some catalogues, more than one magnitude type is reported for a specific event and, from these possible two or more values (reported magnitudes), one value is selected as the preferred magnitude. In this regard, the SHEEF08 catalogue only provides preferred magnitude (value and type), however, in Table 4, all the reported magnitude columns have marks for this file because this is the primary source of information and the preferred magnitude type can be any of the types considered in the reported magnitude columns. For East Canada, the SHEEF08 catalogue is mainly composed of Nuttli magnitude (M_N) data, meanwhile for West Canada, M_L type is predominant over other types (M_W, M_S, or Mc) (Halchuk, 2009). In this catalogue, there are also magnitudes reported as M type (older events with undefined type), OT type (usually based on intensity data), or the magnitude type is left blank (unknown type for which the MZ code was assigned in the computer program). For the SHEEF08 catalogue, the depth designation (indication on how the depth is determined or assigned) can be: F or G (value determined by a geophysicist; both codes are the same but were used at different time periods), N (assigned hypocenter and time), H (assigned hypocenter but calculated origin time), or V (depth determined from regional depth phases) (S. Halchuk, personal communication, 2010). For cases where the depth designation is left blank, the code Z is assigned in the composite catalogue.

For the GSCE and PGCW catalogues from the GSC offices, the previous codes and magnitude types are valid but the main feature in these files is that they may contain different magnitude types for the same event, allowing enhancement of magnitude information in the composite catalogue.

For the PETE catalogue (central and eastern U.S.), the preferred magnitude is always assumed to be $m_{bLg,}$ which, in fact, is the original magnitude for most of the events (C. Mueller, personal communication, 2009). It also provides for some events different original magnitude types. Because there is no information for depth designation, the code P was assigned to all the events from this catalogue (in the program and in the composite catalogue) as the depth designation, with no specific meaning but to indicate they are from the PETE catalogue. For the PETW catalogue (western U.S.), even though about 95% of the reported magnitudes may be M_W, Mb, M_S or M_L, the preferred magnitude is assumed to be M_W (C. Mueller, personal communication, 2009). Again, the depth designation is set to P. The third U.S. source of information is the catalogue ANSS, which may contain different magnitude types for the same event. For the depth designation, the code A was assigned to events from this catalogue , with no specific meaning but to indicate they are from this catalogue. Finally, the MASE catalogue, as mentioned in the preceding section, is based on Nuttli magnitudes (M_N), describes the seismicity occurred from 1980 to 2006 in eastern Canada; the main feature of this database is that it provides high quality estimation of earthquake depth (Ma and Atkinson, 2006) and epicentre coordinates. For this catalogue, the depth designation in the composite catalogue is set to M.

For all the input catalogues, events for which there was no information (blank space) for magnitude type or an unknown type was found, the codes MM or MZ were used, and these cases were grouped under these column names in the computer program. The way they were grouped and processed for the final composite catalogue is explained in the following.

In Section 4, there are two more file couples (not included in Table 4): 1) **maprw.dat** and **maplegendrw.dat,** and 2) **mapre.dat** and **maplegendre.dat,** raster files that represent the seismic source zones map and which are used to classify the catalogue earthquakes and assign the corresponding seismic source zone flag. These are indexed matrices, with index values equal to the code number (zone flag) assigned to each seismic source zone, as in Table 1.

Table 4 – Input files used to develop the CCSC09 catalogue. The ● mark in cells represents the primary data (SHEEF08 catalogue) for each particular variable. The ✕ mark in cells means that the data from a particular variable (column) are available from the corresponding input file (row) but were not used to update the data for the composite catalogue; the ✳ mark in cells means availability of data for a particular variable – file combination and that these data are potential candidates to update the composite catalogue data set. Blank cells means that there are no data for the corresponding type of variable – file combination. Last column gives an indication of the depth designation recovered from each file and, at the bottom of the table, a brief description on these values (depth designation) as well as notes on reported magnitudes are added.

File name **[Format; catalogue; section]**	**Date**			**Time**		**Location**		**Reported magnitudes**								**M prf**		**Depth**		**Dty**
	year	**month**	**day**	**hour**	**minute**	**latitude**	**longitu**	**MB**	**MN**	**ML**	**MS**	**MC**	**MW**	**MM**	**MZ**	**value**	**type**	**value**	**type**	
sheef2008 [*.dat; SHEEF08; 1]	●	●	●	●	●	●	●	●	●	●	●	●	●	● 1*	● 2*	●	●	●	●	**G, F, H, N, Z**
WQSZ_80_06 [*.txt; MASE; 1]	✳	✳	✳	✳	✳	✳	✳									✕		✳	**3***	**M**
NONT_80_06 [*.txt; MASE; 1]	✳	✳	✳	✳	✳	✳	✳									✕		✳	**3***	**M**
ceef1991 [*.dat; GSCE; 1]	✕	✕	✕	✕	✕	✕	✕	✳	✳	✳	✳					✕	✕	✕	✕	✕
pgc_ceef [*.1991; PGCW; 2]	✳	✳	✳	✳	✳	✳	✳	✳	✳	✳	✳	✳ **4***	✳			✳	✳	✳	✳	**G, Z**
emb_cat3_export [*.txt; PETE; 3]	✳	✳	✳	✳	✳	✳	✳	✳	✳ **5***	✳	✳	✳	✳	✳ **5***	✳ **5***	✳	✳	✳	**6***	**P**
wmm_cat3_export [*.txt; PETW; 3]	✳	✳	✳	✳	✳	✳	✳	✳	✳ **5***	✳	✳	✳	✳	✳ **5***	✳ **5***	✳	✳	✳	**6***	**P**
ceus3 [*.xls; PETE; 3]	✳	✳	✳	✳	✳	✳	✳	✳	✳ **5***	✳	✳	✳	✳	✳ **5***	✳ **5***	✳	✳	✳	**6***	**P**
wus3 [*.xls; PETW; 3]	✳	✳	✳	✳	✳	✳	✳	✳	✳ **5***	✳	✳	✳	✳	✳ **5***	✳ **5***	✳	✳	✳	**6***	**P**

ANSS78 [*.txt; ANSS; 3]	✳	✳	✳	✳	✳	✳	✳	✳	✳	✳ **7***	✳	✳	✳	✳	✳ **7***	✳	✳	✳	**6***	**A**
ristau [*.xls; RIST; 4]	✳	✳	✳	✳	✳	✳	✳						✳			✳	✳			
ristauxx [*.xls; RIST; 4]	✳	✳	✳	✳	✳	✳	✳						✳			✳	✳			
Depth designation:	**G** and **F** (both have the same meaning) = value determined by a geophysicist; **Z** = unknown (blank space); **M** = determined by Ma and Atkinson (2006); **P** or **A** = artificial designation for Petersen or ANSS catalogues; **H** = assigned hypocenter but calculated origin time; **N** = assigned hypocenter and time.																			
Notes for referenced cells	1* - **MM** is formed by M; 2* - **MZ** is formed by blank space or OT in sheef08 file, refers to unknown magnitude type; 3* - artificial designation for values calculated by Ma and Atkinson (2006); 4* - MC column from pgc_ceef data is formed by MC or OT (other) values indistinctly; 5* - **MN**, **MM**, **MZ** columns from Petersen data are formed by: [**MN**&**LG**&**MD**& **blanks**], [column of **zeros**], [**MI**&**UK**&**GR**& **blanks**] respectively; 6* - artificial designation for depth values in Petersen or ANSS catalogues; 7* - **ML**, **MZ** columns from ANSS data are formed by: [**ML**&**MD**], [**UK**] respectively. **MD**, **MI**, **GR**, **UK**, refers to values calculated from duration, intensity, Gutenberg-Richter relationship, and unknown magnitude type.																			

Output files

The output files **ccsc09west.txt** and **ccsc09east.txt** for West and East Canada, respectively, integrate the output information from the process described above. Two header files (**headerw.txt** for West and **headere.txt** for East) are used to describe all the output variables contained in each composite catalogue. This header text precedes the data organized according to the following format.

Table 5 – Format used in CCSC09 catalogue files (according to MATLAB code). The code column gives the actual letters (abbreviation) used in the header of the output files to refer the corresponding variable.

Variable		Code	Columns	Format
Year		**yr**	1-4	%4d
Month		**mo**	6-7	%2d
Day		**dy**	9-10	%2d
Hour		**hr**	12-13	%2d
Minute		**mi**	15-16	%2d
Latitude		**lat**	19-25	%7.3f
Longitude		**lon**	28-36	%9.3f
Reported magnitudes		**MB**	40-43	%4.2f
		MN	46-49	%4.2f
		ML	52-55	%4.2f
		MS	58-61	%4.2f
		MC	64-67	%4.2f
		MD	70-73	%4.2f
		MW	76-79	%4.2f
		MM	82-85	%4.2f
		MZ	88-91	%4.2f
Preferred magnitude	Value	**Mpf**	94-97	%4.2f
	Type	**Tm**	99-100	%s
Depth value		**Dep**	103-108	%6.2f
Depth designation		**dd**	110	%s
Source catalogue flag		**cf**	113	%1d
Seismic source flag		**zf**	116-118	%3d
M_W correction factor		**mf**	121-125	%5.2f
Moment magnitude		**M_W**	129-132	%5.2f

Notes on MATLAB format: The format argument is a string containing C language conversion specifications. It starts with the % character (required). Digits between the % character and the one-letter conversion character, specify field width (digit to the left of decimal point), which is the total number of characters to be printed, and precision which is the number of digits to the right of decimal point (if required). Conversion characters for these cases are: d = decimal notation (signed), f = fixed-point notation, and s = string of characters.

Some notes on specific variables from output files

For the preferred magnitude of any event, from any input catalogue, with no information (blank space) on magnitude type, or with unknown magnitude type codes, the variable MM or MZ was generated and used in the **ccsc09.m** program. These variables are integrated in different ways according to the input file under analysis.

Table 6 – Magnitude types of the preferred magnitude for the listed catalogues (first column), used to integrate the MM or MZ internal variables. These magnitude types represent unknown types or lack of information.

Input catalogue	Input M types for MM variable integration	Input M types for MZ variable integration
SHEEF08	M	OT and blank space
PETE	Non applicable	UK, MI
PETW	Non applicable	UK, MI, GR, blank space
ANSS	M	UK
Notes: M = older events with undefined type; OT = other, mainly based on intensity data; UK = unknown type; MI = based on intensity data; GR = based on the Gutenberg-Richter model; and blank space = no magnitude type information.		

Even though the variables MM and MZ are used internally, their values are included as reported magnitudes in the composite catalogue and the M type code, instead of being MM or MZ, is replaced in one of the following ways: 1) by using, if available, recovered information on reported magnitude (value and type) for the event under analysis during the exploration process of different databases, or 2) by assuming that the unknown magnitude types are M_L for East and West regions before 1960. After 1960, the assumption is that they are M_L in the West and M_N in the East for Canadians earthquakes. For U.S. information, the assumption given by Petersen et al. (2008) (as explained above) is considered.

Minimum magnitude in output files

It is important to note that no minimum magnitude limit is introduced in the final catalogue. Since it was originally intended to retain all the available seismic information in the composite catalogue, all the events in the source databases are include in the output files without any minimum magnitude consideration. It is recommended that the user should apply a magnitude completeness analysis to assess the threshold magnitude spatially and temporally. Earthquakes in the CCSC09 catalogue with magnitude $M_W >= 2.5$ are shown in Appendix C.

4. 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 4459, 155 pp.

Atkinson, G.M. (2005). Ground motions for earthquakes in South-western British Columbia and North-western Washington: crustal, in-slab, and offshore events, Bull. Seism. Soc. Am. **95**, 1027-1044.

Atkinson, G.M., and D. M. Boore, D.M. (2006). Earthquake ground motion prediction equations for eastern North America, Bull. Seism. Soc. Am. **96**, 2181-2205.

Bent, A.L. (2009). A moment magnitude catalog for the 150 largest eastern Canadian earthquakes, Geological Survey of Canada, Open File 6080, 23 pp.

Braunmiller, J., and Nabelek J. (2002). Seismotectonics of the Explorer region, J. Geophys. Res. **107**, ETG 1-1-ETG 1-25.

Cassidy, J.F., Rogers, G.C., and Ristau, J. (2005). Seismicity in the vicinity of the SNORCLE corridors of the Northern Canadian Cordillera, Can. J. Earth Sci. **42**, 1137-1148.

Dewberry, S., and Crosson, R. (1995). Source scaling and moment estimation for the Pacific Northwest seismograph network using S-Coda amplitudes, Bull. Seism. Soc. Am. **85**, 1309-1326.

Halchuk, S. (2009). Seismic Hazard Earthquake Epicenter File (SHEEF) used in the fourth generation seismic hazard maps of Canada, Geological Survey of Canada, Open File 6208, 15 pp.

Kim, W.Y. (1998). The M_L scale in eastern North America, Bull. Seism. Soc. Am. **88**, 935-951.

Ma, S., and Atkinson, G.M. (2006). Focal depths for small to moderate earthquakes ($m_N \geq 2.8$) in western Quebec, southern Ontario, and northern New York, Bull. Seism. Soc. Am. **96**, 609-623.

Petersen, M., Frankel, A., Harmsen, S., Mueller, C., Haller, K., Wheeler, R., Wesson, R., Zeng, Y., Boyd, O., Perkins, D., Luco, N., Field, E., Wills, C., and Rukstales, K.

(2008). Documentation for the 2008 Update of the United States National Seismic Hazard Maps: U.S. Geological Survey Open-File Report 2008-1128, 61 p.

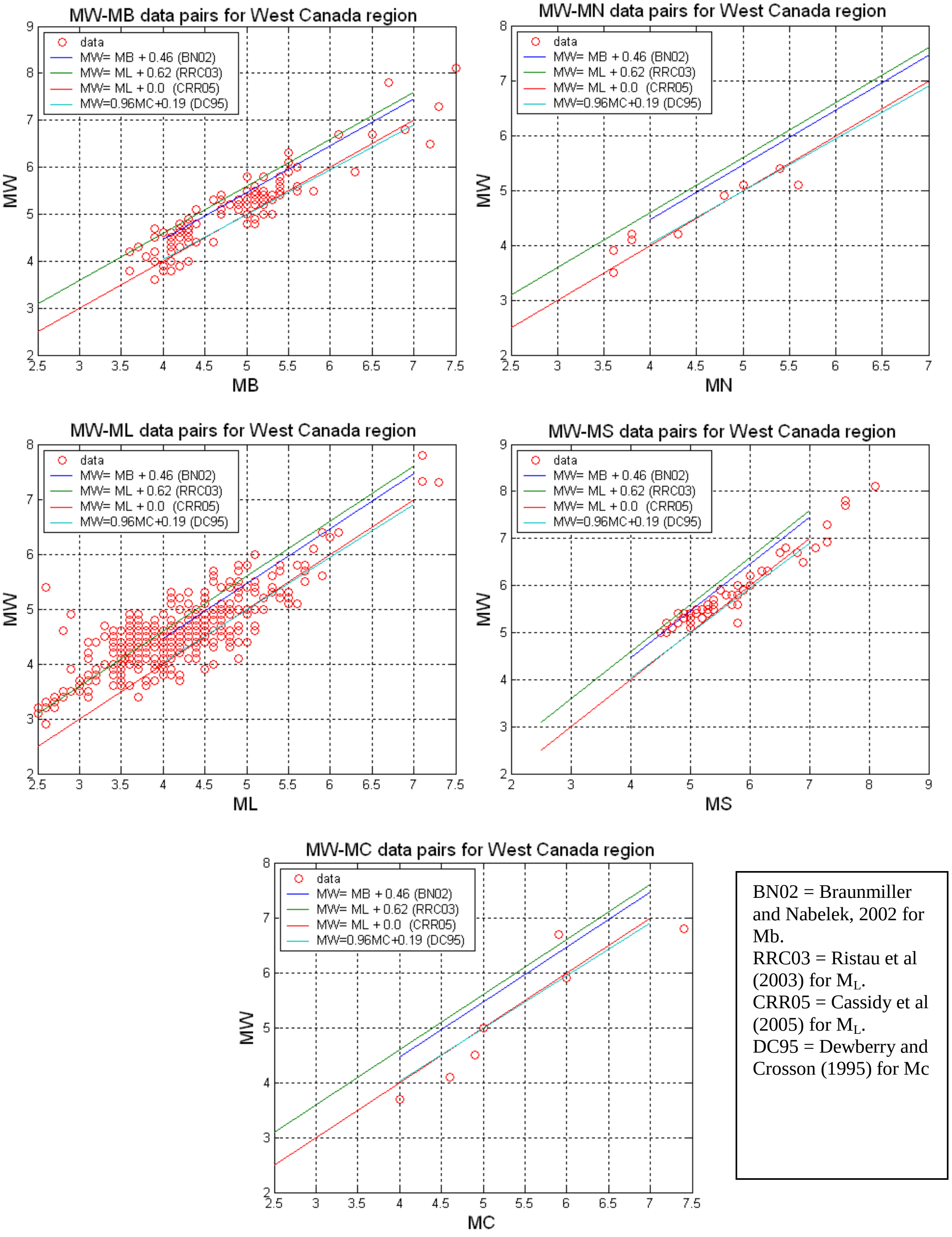
Ristau, J., Rogers, G.C., and Cassidy, J.F. (2003). Moment magnitude-local magnitude calibration for earthquakes off Canada's west coast, Bull. Seism. Soc. Am. **93**, 2296-2300.

Ristau, J. (2004). Seismotectonics of western Canada from regional moment tensor analysis, Ph. D. dissertation. School of Earth and Ocean Sciences, University of Victoria, B. C.

Ristau, J., Rogers, G.C., and Cassidy, J.F. (2005). Moment Magnitude-Local Magnitude Calibration for Earthquakes in Western Canada, Bull. Seism. Soc. Am. **95**, 1994-2000.

Sonley, E., and Atkinson, G.M. (2005). Empirical relationship between moment magnitude and Nuttli magnitude for small-magnitude earthquakes in South-eastern Canada, Seis. Res. Lett. **76**, 752-755.

Appendix A


MW-MB data pairs for West Canada region

MW-MN data pairs for West Canada region

MW-ML data pairs for West Canada region

MW-MS data pairs for West Canada region

MW-MC data pairs for West Canada region

BN02 = Braunmiller and Nabelek, 2002 for Mb.
RRC03 = Ristau et al (2003) for M_L.
CRR05 = Cassidy et al (2005) for M_L.
DC95 = Dewberry and Crosson (1995) for Mc

Figure A-1. M_W-M (different types) data pairs for West Canada (whole region). Colour lines are M_W calibration models developed for the region. Codes for models are given in the text box.

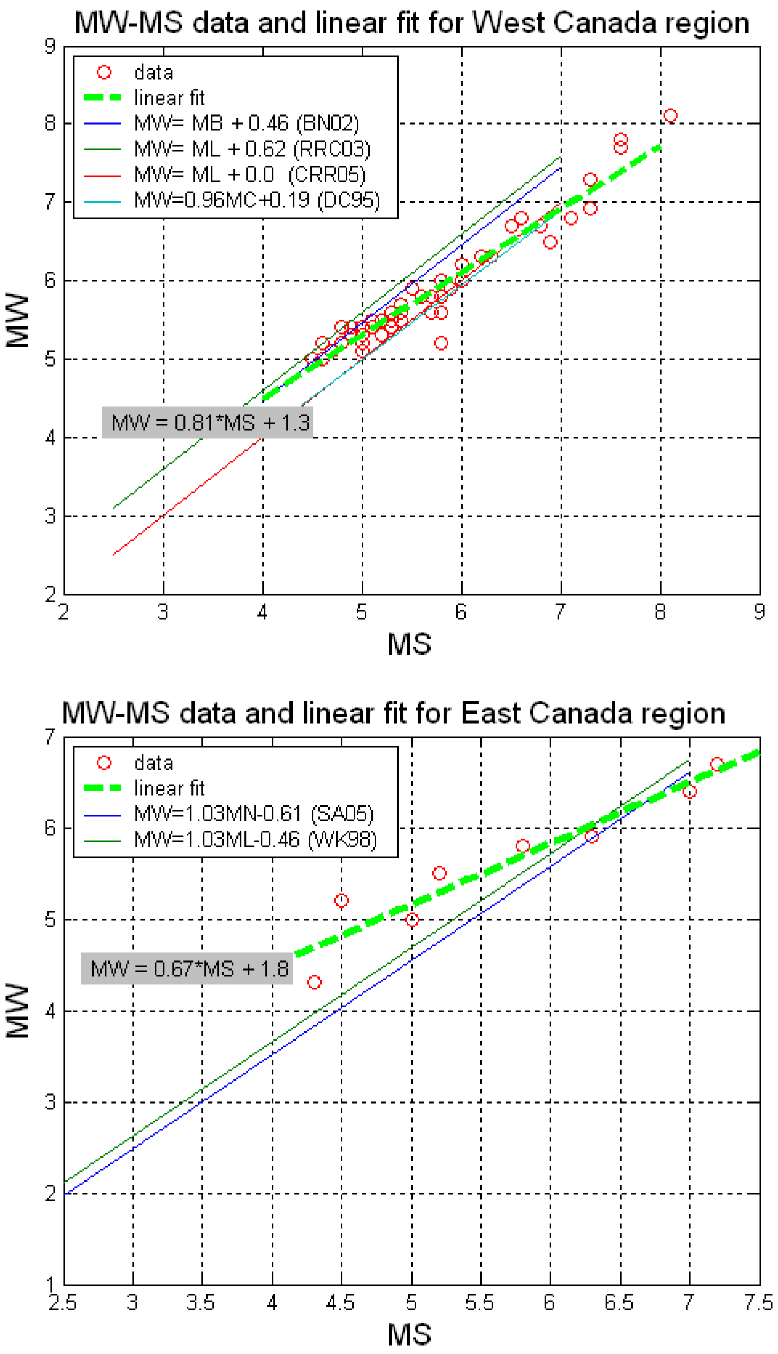


Figure A-2. M_W-M_S data pairs for West (top panel) and for East (bottom panel) regions. Continuous lines in the bottom panel shows the Sonley and Atkinson (2005) model [SA05] for M_N and the Kim (1998) model [WK98] for M_L. Top panel models are as in Figure A-1. For both graphs, the green dashed line is the linear fit to the data and the resulting model is in the grey box for each region.

Mw-ML conversion for West Canada region

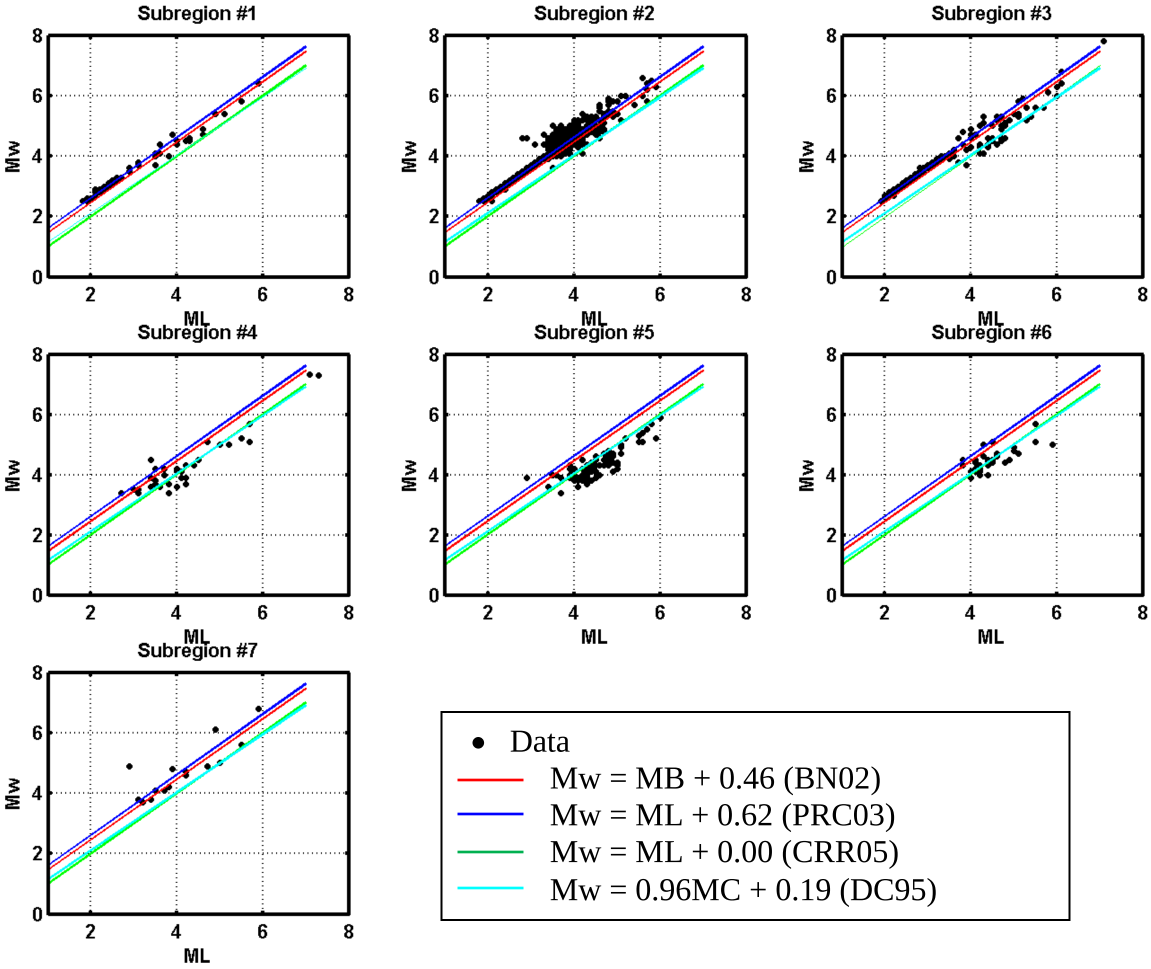


Figure A-3. M_W-ML data pairs for western Canada. Continuous lines are models as given in Figure A-1. Data from different source zones were grouped (sub-regions) and plotted in each graph. Source zones codes and their corresponding sub-regions are given in Table 2.

Mw-MB conversion for West Canada region

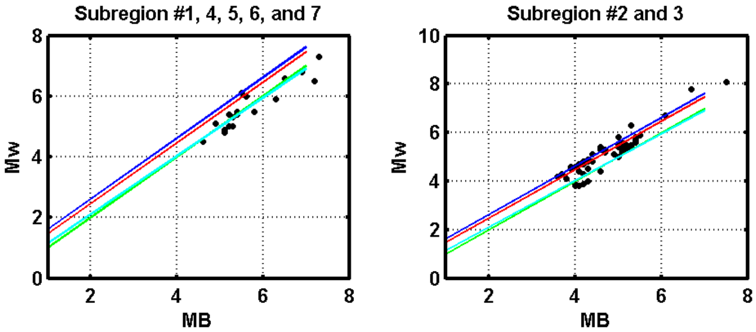


Figure A-4. M_W-M_B data pairs for western Canada. Continuous lines are models as given in Figure A-1. Data from different subregions were grouped and plotted in each graph. Source zones codes and their corresponding sub-regions are given in Table 2.

Mw-MN conversion for West Canada region

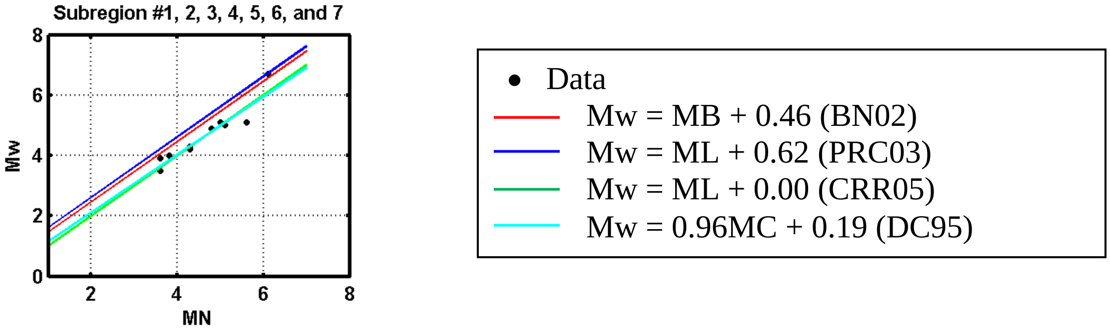


Figure A-5. M_W-M_N data pairs for western Canada. Continuous lines are models as given in Figure A-1. Data from different subregions were grouped and plotted in one graph. Source zones codes and their corresponding sub-regions are given in Table 2.

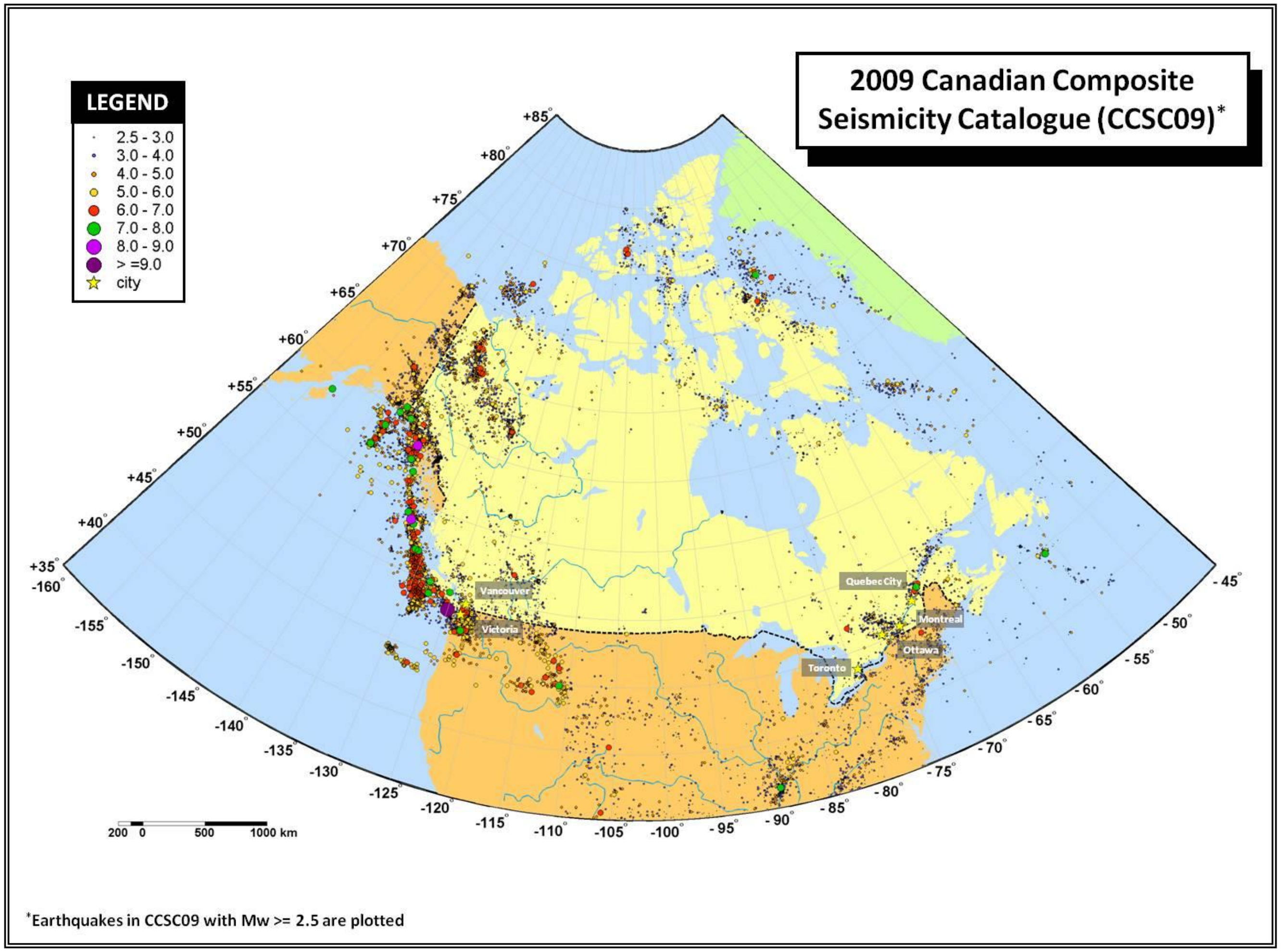
Appendix B

Table B-1. The database of adopted M_W values from Atkinson and Boore (2006) [AB06] or Bent (2009) [B09].

Event	Year	Month	Day	M_W	Source
1	1663	2	5	7	B09
2	1791	12	6	5.5	B09
3	1816	9	9	4.8	B09
4	1831	5	8	4.8	B09
5	1860	10	17	6.1	B09
6	1870	10	20	6.6	B09
7	1884	9	19	4.8	B09
8	1893	11	27	5.1	B09
9	1895	9	1	4.1	B09
10	1897	3	23	4.6	B09
11	1905	7	27	4.2	B09
12	1906	5	10	3.7	B09
13	1909	5	26	5.3	B09
14	1909	7	19	4.3	B09
15	1912	1	2	4.7	B09
16	1914	2	10	5.1	B09
17	1917	9	3	4.3	B09
18	1925	3	1	6.4	AB06
19	1926	11	5	3.8	B09
20	1927	6	1	4.6	B09
21	1929	8	12	4.9	AB06
22	1929	11	18	7.3	AB06
23	1931	9	20	4.6	B09
24	1933	11	20	7.4	B09
25	1934	7	30	4.3	B09
26	1934	11	12	4	B09
27	1935	11	1	6.2	AB06
28	1939	10	19	5.3	AB06
29	1940	12	20	5.5	AB06
30	1940	12	24	5.6	B09
31	1944	9	5	5.8	AB06
32	1946	7	23	4.1	B09
33	1952	6	20	4.1	B09
34	1961	12	31	4.2	B09
35	1963	9	4	6.1	B09
36	1968	11	9	5.4	AB06
37	1971	10	2	4.7	B09
38	1971	12	7	5.6	B09
39	1972	1	21	4.6	B09
40	1972	11	19	5.9	B09
41	1972	11	21	6	B09
42	1972	12	27	6.3	B09
43	1972	12	28	5.9	B09
44	1975	10	6	5.1	B09

Event	**Year**	**Month**	**Day**	**M_W**	**Source**
45	1979	6	27	5	B09
46	1979	8	19	4.8	B09
47	1980	7	27	5.1	AB06
48	1982	1	9	5.5	AB06
49	1982	1	9	4.6	AB06
50	1982	1	11	5.2	AB06
51	1982	1	19	4.3	AB06
52	1982	3	31	4.2	AB06
53	1982	6	16	4.2	AB06
54	1983	10	7	5	AB06
55	1985	10	5	6.7	AB06
56	1985	12	23	6.8	AB06
57	1985	12	25	5.2	AB06
58	1986	1	31	4.8	AB06
59	1986	7	12	4.5	AB06
60	1987	6	10	5	AB06
61	1987	12	13	5.4	B09
62	1988	3	25	6.3	AB06
63	1988	11	23	4.3	AB06
64	1988	11	25	5.8	AB06
65	1989	3	16	5	AB06
66	1989	12	25	5.9	AB06
67	1990	10	19	4.7	AB06
68	1992	1	4	5.5	B09
69	1997	11	6	4.5	AB06
70	1997	12	6	5	B09
71	1998	9	25	4.5	AB06
72	1999	3	16	4.5	AB06
73	2000	1	1	4.7	AB06
74	2001	8	14	5.2	B09
75	2002	4	20	5	AB06
76	2002	6	18	4.6	B09
77	2004	8	26	4.3	B09
78	2005	3	6	5	AB06
79	2008	4	18	5.2	B09

Appendix C



2009 Canadian Composite Seismicity Catalogue (CCSC09)*

*Earthquakes in CCSC09 with Mw >= 2.5 are plotted

Figure C-1. Earthquakes in CCSC09 with moment magnitude equal or greater than 2.5 are plotted.