Intraplate seismicity in Canada: a graph theoretic approach to data analysis and interpretation

K. Vasudevan¹, D. W. Eaton¹, and J. Davidsen²

¹Department of Geoscience, University of Calgary, Calgary, Alberta T2N 1N4, Canada
²Department of Physics and Astronomy, University of Calgary, Calgary, Alberta T2N 1N4, Canada

Received: 2 April 2010 – Revised: 13 July 2010 – Accepted: 17 August 2010 – Published: 4 October 2010

Abstract. Intraplate seismicity occurs in central and northern Canada, but the underlying origin and dynamics remain poorly understood. Here, we apply a graph theoretic approach to characterize the statistical structure of spatiotemporal clustering exhibited by intraplate seismicity, a direct consequence of the underlying nonlinear dynamics. Using a recently proposed definition of “recurrences” based on record breaking processes (Davidsen et al., 2006, 2008), we have constructed directed graphs using catalogue data for three selected regions (Region 1: 45°–48°N/74°–80°W; Region 2: 51°–55°N/77°–83°W; and Region 3: 56°–70°N/65°–95°W), with attributes drawn from the location, origin time and the magnitude of the events. Based on comparisons with a null model derived from Poisson distribution or Monte Carlo shuffling of the catalogue data, our results provide strong evidence in support of spatiotemporal correlations of seismicity in all three regions considered. Similar evidence for spatiotemporal clustering has been documented using seismicity catalogues for southern California, suggesting possible similarities in underlying earthquake dynamics of both regions despite huge differences in the variability of seismic activity.

1 Introduction

While earthquakes near tectonic plate boundaries fit well within the plate tectonics paradigm, the causes and dynamics of intraplate seismicity within continental interiors remain enigmatic. This class of seismicity has been extensively studied in recent decades (Gregersen, 1979; Adams and Basham, 1990; Johnston and Kanter, 1990; Denith and Featherstone, 2003; Sandiford and Egholm, 2008; Stein and Liu, 2009) and continues to draw the interest of seismologists and seismic hazard analysts, including earthquakes in the continental interior of the US (Coppersmith et al., 1987; Dewey, 1988; Johnston and Kanter, 1990; Zoback, 1992; Li et al., 2009; Stein and Liu, 2009), several regions of Canada (Basham et al., 1977; Quinlan, 1984; Adams, 1989a; Bent, 1996; Eaton et al., 2005; Darbyshire et al., 2007; Wang et al., 2008; Woodgold, 2010), Australia (Bowman et al., 1990; Denith and Featherstone, 2003; Sandiford and Egholm, 2008), southeast Brazil (Assumpção et al., 2004), and northern Norway (Gregersen, 1979; Atakan et al., 1994; Bungum et al., 2010; Gregersen and Voss, 2009). One particularly intriguing aspect is the high spatial and temporal variability of seismic activity in continental interiors compared with plate boundaries (Stein and Liu, 2009). Glacial Isostatic Adjustment (GIA) (Quinlan, 1984; Adams, 1989b; Atakan et al., 1994; Wu and Hasegawa, 1996; Wu, 1999; Wu and Johnston, 2000; Ma et al., 2008; Wang et al., 2008; Wu et al., 2010), hotspot tracks (Ma and Eaton, 2006), localized high strain-rates because of thermal fluctuations in the lower crust (Sandiford and Egholm, 2008), and reactivation of ancient terrane boundaries (Hasegawa, 1991; Denith and Featherstone, 2003) have been invoked as possible causative factors for earthquakes in these regions.

The focus of this study is on intraplate seismicity in parts of eastern and northern Canada, where all of these factors are potentially significant. Patterns of seismicity derived from earthquake catalogues and faults generated by rock deformation occur within a broad range of coupled spatiotemporal scales, requiring an understanding of the collective behaviour of earthquakes and faults (Ben-Zion, 2008). There has been a concerted effort to utilize empirical laws in order to establish simple stochastic models that can be used for seismic hazard assessment and potentially for earthquake prediction (Guo and Ogata, 1997; Turcotte, 1997; Schoenberg, 2003; Li et al., 2009; Vere-Jones, 2009). One way to gain insight into the underlying dynamics of earthquakes is...
to study the spatiotemporal patterns of seismicity, wherein each earthquake is treated as a point event in space and time. Such an approach may shed light on the fundamental physics since these patterns are emergent processes of the underlying many-body non-linear system. This approach has been successfully applied in many cases, leading to the discovery of new key features of seismicity (Tata, 1969; Rundle et al., 2003; Corral, 2004; Davidsen and Goltz, 2004; Shcherbakov et al., 2004; Davidsen and Paczuski, 2005; Corral, 2006a, b; Hainzl et al., 2006; Stein and Mazzotti, 2007; Davidsen et al., 2008; Marsan and Lengliné, 2008).

The framework of spatiotemporal point processes is especially well suited for complex network analysis (Albert and Barabási, 2002; Newman, 2003; Fagiolo, 2007; Fortunato, 2010), which has recently been applied to earthquake data (Baiesi and Paczuski, 2004, 2005; Davidsen et al., 2008; Jiménez et al., 2008; Marsan and Lengliné, 2008; Zaliapin et al., 2008). This approach has proved helpful in characterizing the spatiotemporal clustering of earthquakes in southern California and to establish non-trivial features by direct comparison with stochastic null models (Davidsen et al., 2008; Peixoto and Davidsen, 2008; Peixoto et al., 2010). Most importantly, the latter approach has led to a new and independent estimate of the rupture length and its scaling with magnitude.

Here, we apply the approach of Davidsen (2008) to seismicity data from three selected regions in eastern and northern Canada (Figs. 1 and 2). These three regions were chosen to sample intraplate seismicity within tectonic settings where pre-existing terrane boundaries, hotspot tracks and GIA have been variably cited as potential causes of intraplate seismicity. Region 1 contains the West Québec Seismic Zone (Forsyth, 1981; Mereu et al., 1986; Ma and Atkinson, 2006; Ma et al., 2008), a moderately active area of intraplate seismicity, thought to be controlled by crustal structures associated with both a failed Precambrian rift (the Ottawa-Bonnechere graben) and/or a Mesozoic hotspot track (Great Meteor hotspot) (Ma and Eaton, 2006; Ma et al., 2008). Region 2 contains diffuse, shallow intraplate seismicity within topographically elevated areas of an Archean craton (Superior Province), as well as localized activity along one or more hotspot tracks (Forsyth, 1981; Ma and Eaton, 2006). Region 3 includes an area of Paleozoic fault structures (Boothia-Bell uplift) and is arguably most strongly influenced by GIA, which is induced by unloading of the continental Laurentide Ice sheet (Fig. 1a). Although GIA perturbs the crustal stresses in all three regions, inter-regional difference may arise as a result of time-dependent relaxation of the mantle (Peltier, 1996; Wu, 1998). Due to the location of Region 3 near the centre of the former ice sheet, glacial unloading occurred most recently there.

The outline of the paper is as follows. After briefly reviewing intraplate seismicity, we present the concept of spatiotemporal recurrence that incorporates the theory of records, showing how directed, recurrence graphs are constructed. Next, we present the properties of these graphs for the regions considered. In particular, we consider the distance-interval and time-interval (or waiting-time) probability distributions, including consistency checks such as the probability distributions derived from magnitude-limited subsets of the seismicity catalog. To link causal characteristics of the dynamics to the network of recurrences noted in probability distributions, we use the statistical properties of a null model, mathematically established by Davidsen et al. (2008), where the events in space and time are random, uncorrelated, and causally unrelated. We present two variants of the Monte Carlo simulations with 100 realizations to emulate the null model. We compare these results with the results from the original and segments of the data. In so doing, we examine the causal structure present in the complex network of intraplate earthquakes. Finally, we interpret our results in the context of intraplate seismicity and compare them with the results obtained by Davidsen et al. (2008) for plate-boundary seismicity in southern California.

2 Intraplate seismicity of Canada

Intraplate seismicity is characteristic of all continents, but it is generally weaker and less frequent than seismicity observed near active plate boundaries. One class of intraplate seismicity is concentrated along passive continental margins (Sandford and Egholm, 2008), whereas another comprises events that are scattered heterogeneously throughout the stable continental interior (Basham et al., 1977; Dewey, 1988; Adams, 1989a; Adams and Basham, 1990; Johnston and Kanter, 1990; Hasegawa, 1991; Denith and Featherstone, 2003). Our present work focuses on the latter class of intraplate earthquakes, which are less well understood. A global review (Dewey, 1988) concluded that 71% of the seismicity in stable continental regions is associated with embedded passive margins or continental rift zones. From palaeoseismic studies and crustal deformation rates, which are usually much slower (less than 1 mm/year) in continental interiors than at many plate boundaries (3–16 cm/year), the average recurrence times for large earthquakes in continental areas is inferred to be of the order of thousands of years (Lee et al., 2007).

Figure 1 shows a tectonic map as well as a simplified seismic hazard map of Canada. The stable continental interior of Canada is predominantly comprised of the Canadian Shield, where exposed rocks are of Precambrian age and have experienced little tectonic activity for the past 600 million years or more. In terms of seismic hazard (Fig. 1b), Canada is divided into two principal seismogenic regions, western and eastern Canada. In western Canada, seismicity is linked to plate-boundary tectonic forces. Relatively high levels of seismic activity along the west coast are attributed to the active convergent plate boundaries, as reflected by both subduction-related earthquakes (Adams, 1990; Clague, 1997; Clowes and Hyndman, 2002) and transform-fault-related strike-slip
Fig. 1. (a) A simplified tectonic map of Canada. Tectonic age is defined as the time since the most recent episode of tectonic deformation. Tectonic ages in Baffin Island, not shown in the figure, are comprised of both Archean and Paleoproterozoic ages. The dark purple line marks roughly the Laurentide ice margin at glacial maximum. (b) 2005 seismic hazard map of Canada (http://earthquakescanada.nrcan.gc.ca/hazard-alea/) showing the three regions considered for study. SHmax (maximum horizontal stress) orientations for early postglacial times and the current orientations are also shown.
earthquakes (Bird et al., 1996; Bufe, 2005). The seismicity of northern and eastern Canada is classified as a single seismogenic region because of the common intraplate tectonic setting. Some of the seismicity appears to be localized near Paleozoic fault structures, such as the Boothia Uplift–Bell Arch zone in northern Hudson Bay and Paleozoic rift systems along the St. Lawrence and Ottawa rivers (Adams, 1989b; Adams and Basham, 1990; Hasegawa, 1991). The northern band of seismicity in western Quebec may be linked to a hotspot track (Ma and Eaton, 2006).

The Laurentide ice sheet at the last glacial maximum, which occurred about 21 000 years ago (Denton and Hughes, 1981), covered most of northern and eastern Canada (see Fig. 1a). In the case of Regions 1 and 2, the retreat of the ice sheet occurred about 11 500 years before present, whereas for Region 3 it occurred no earlier than approximately 8400 years before present. This difference in time since glacial unloading may be significant, as crustal stresses induced by GIA as well as seismicity rates are thought to depend on time after deglaciation (Wu, 1999; Wu and Johnston, 2000).

The spatiotemporal clustering of intraplate earthquakes in these areas has previously received little attention. Here, we follow a recently proposed approach to characterize the spatiotemporal clustering of earthquakes based on modern tools of complex network theory.

3 Directed recurrence networks of earthquakes

In the area of statistical seismology, earthquakes are often treated as marked point processes in space and time, which is a convenient way to analyze the spatiotemporal patterns of seismicity (Turcotte, 1997; Ogata, 1998; Schoenberg, 2003; Vere-Jones, 2009). The observed patterns serve as a benchmark test for models that aim to reproduce natural seismicity. Here, we quantify the dynamics of such spatiotemporal point processes in terms of the properties of complex networks or graphs.

The key idea is to represent seismicity as a directed graph by linking each earthquake to its “recurrences”. We apply a recently proposed definition of recurrence (Davidsen et al., 2008), which is different from the notion used in recurrence time statistics (Saichev and Sornette, 2007) widely considered in the literature. We illustrate our definition of recurrences of a given event and the mapping of the data to graph in Fig. 3 and an associated Table 1. Seismic events taken from a seismicity catalogue are treated as nodes, \( a_i \), in this graph. In Fig. 3, the nodes are labeled sequentially with subscripts from 1 to 8 according to their occurrence in time (i.e., from first to last). The spatial distance between two events, \( a_i \) and \( a_j \), is denoted by \( l_{ij} \). Edges, \( e_{ij} \), taken into consideration in the directed graph here are for the recurrences only, based on the definition given in (Davidsen et al., 2008): event \( a_j \) is a recurrence of any previous event \( a_i \) if it is closer to \( a_i \) than every other event \( a_k \) that occurred between the two, i.e. for
an aftershock sequence (ETAS) models (Ogata, 1998). Indeed, ETAS models are able to capture many observed features of these sequences. More importantly, however, no physical difference between earthquakes has been established that would support the traditional terminology of mainshocks, aftershocks and foreshocks as separated classes of events (Davidson et al., 2008). Hence, we treat all of the observed earthquakes in the catalogue alike.

In order to investigate the influence of catalog incompleteness and detection thresholds on the properties of the recurrence networks, we analyze several networks for each earthquake catalog obtained by only considering events above a given magnitude threshold, $m$. In addition, we also study networks generated by considering only segments of the catalogs such as the recent half (or first half) and the early half (or last half). Here, the main purpose is to understand whether or not such a partition, where station monitoring of earthquakes could differ, results in recurrence networks with differing statistical properties.

To establish if the observed properties of the recurrence networks indeed reflect non-trivial spatiotemporal clustering or spatiotemporal correlations between earthquakes, we also consider the recurrence networks of shuffled catalogs obtained through Monte Carlo simulations. Our Monte-Carlo simulations fell into two categories, namely, shuffling the spatial locations of the events in the catalogue (MC1) and shuffling the spatial locations and magnitudes of the events, prior to working with the newly generated directed recurrence graph (MC2). In both cases, each shuffled catalog is a realization of a random process with no spatiotemporal correlations but the same (independent) spatial and temporal correlations as the original catalog. Hence, differences between the properties of the network of recurrences for the original catalog and those for the shuffled catalogs indicate non-trivial spatiotemporal clustering. A typical Monte Carlo simulation entailed 100 realizations allowing us to estimate average properties.

Our seismicity catalogues were downloaded from the Earthquakes Canada online database (earthquakescanada. nrcan.gc.ca) and cover the time period from 1990–2009. The online database is generated primarily from events that are automatically picked and located from the continuous data stream. As new stations become available, they are added to the “pick list” of network stations. Thus, when new stations are added the quality of the catalogue changes due to the inherent reduction in location uncertainty and reduction in the magnitude detection threshold. In Table 2, we summarize the partitioned catalogues according to the magnitude-threshold and two segments, the early half and the recent half. The ratio of the recent half to the early half signifies the role of the new stations in detecting a large number of small magnitude threshold events, as one goes to Region 3 from Regions 1 and 2.

all $a_k$ with $i < k < j$. Thus, event $a_j$ is automatically a recurrence of event, $a_{j-1}$, for $2 < j < n$, and it is actually the first recurrence of $a_{j-1}$. Since the events are time-ordered, we represent the edges between nodes with a pointed arrow, as is done in a directed graph and shown in Fig. 3. In Table 1, the included and excluded distance-intervals, $l_{ij}$, correspond to the included and excluded edges, $e_{ij}$, of the recurrence network, respectively. Based on the definition of recurrences, the included distance-intervals, $l_{ij}$, for a given event $i$ decrease with increasing index $j$. Thus, the distance-intervals form a record-breaking process with respect to the shortest distance.

It is important to point out that in our analysis of different earthquake catalogs we do not remove aftershocks or apply any other declustering algorithm (Zaliapin et al., 2008). This is based on a hypothesis that earthquakes can be classified within a hierarchical structure in space and time such that each earthquake can simultaneously be an aftershock, mainshock and foreshock. This premise forms the basis of epidemic type aftershock sequence (ETAS) models (Ogata, 1998), in contrast to the traditional classification of an earthquake as a mainshock, aftershock or foreshock. ETAS models have been used to investigate sequences of earthquakes, following a large earthquake, which have been historically identified as aftershocks due to an increased seismic activity as quantified by the Omori law (Turcotte, 1997). Indeed, ETAS models are able to capture many observed features of these sequences. More importantly, however, no physical difference between earthquakes has been established that would support the traditional terminology of mainshocks, aftershocks and foreshocks as separated classes of events (Davidson et al., 2008). Hence, we treat all of the observed earthquakes in the catalogue alike.

In order to investigate the influence of catalog incompleteness and detection thresholds on the properties of the recurrence networks, we analyze several networks for each earthquake catalog obtained by only considering events above a given magnitude threshold, $m$. In addition, we also study networks generated by considering only segments of the catalogs such as the recent half (or first half) and the early half (or last half). Here, the main purpose is to understand whether or not such a partition, where station monitoring of earthquakes could differ, results in recurrence networks with differing statistical properties.

To establish if the observed properties of the recurrence networks indeed reflect non-trivial spatiotemporal clustering or spatiotemporal correlations between earthquakes, we also consider the recurrence networks of shuffled catalogs obtained through Monte Carlo simulations. Our Monte-Carlo simulations fell into two categories, namely, shuffling the spatial locations of the events in the catalogue (MC1) and shuffling the spatial locations and magnitudes of the events, prior to working with the newly generated directed recurrence graph (MC2). In both cases, each shuffled catalog is a realization of a random process with no spatiotemporal correlations but the same (independent) spatial and temporal correlations as the original catalog. Hence, differences between the properties of the network of recurrences for the original catalog and those for the shuffled catalogs indicate non-trivial spatiotemporal clustering. A typical Monte Carlo simulation entailed 100 realizations allowing us to estimate average properties.

Our seismicity catalogues were downloaded from the Earthquakes Canada online database (earthquakescanada. nrcan.gc.ca) and cover the time period from 1990–2009. The online database is generated primarily from events that are automatically picked and located from the continuous data stream. As new stations become available, they are added to the “pick list” of network stations. Thus, when new stations are added the quality of the catalogue changes due to the inherent reduction in location uncertainty and reduction in the magnitude detection threshold. In Table 2, we summarize the partitioned catalogues according to the magnitude-threshold and two segments, the early half and the recent half. The ratio of the recent half to the early half signifies the role of the new stations in detecting a large number of small magnitude threshold events, as one goes to Region 3 from Regions 1 and 2.
Table 2. Number of events according to the magnitude. Ratio of time-interval of the recent half (First Half) to that of the early half (Last Half) for three regions studied: a) Region 1: 0.92 b) Region 2: 0.47; c) Region 3: 0.17. We used the event catalogue from July 1990 to July 2009, available at Earthquakes Canada, Ottawa, Canada (source: http://earthquakescanada.nrcan.gc.ca).

<table>
<thead>
<tr>
<th>Earthquake Region</th>
<th>Magnitude Threshold (≥)</th>
<th>Number of events</th>
<th>Early half of the data (Last Half or Lhalf)</th>
<th>Recent half of the data (First Half or Fhalf)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ottawa-Bonnechere Graben (Region 1)</td>
<td>2.0</td>
<td>830</td>
<td>415</td>
<td>415</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>347</td>
<td>183</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>122</td>
<td>64</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>43</td>
<td>26</td>
<td>17</td>
</tr>
<tr>
<td>Northern Ontario (Region 2)</td>
<td>2.0</td>
<td>347</td>
<td>174</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>130</td>
<td>86</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>31</td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>9</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Boothia Uplift – Bell Arch (Region 3)</td>
<td>2.0</td>
<td>931</td>
<td>466</td>
<td>465</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>605</td>
<td>369</td>
<td>236</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>268</td>
<td>182</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>72</td>
<td>44</td>
<td>28</td>
</tr>
</tbody>
</table>

4 Results: properties of the network of recurrences for intraplate seismicity

4.1 In-degree and out-degree histograms

The properties of the network of recurrences – which is simply a binary directed graph – can be analyzed by several means. For example, some of its topological features can be captured by the in-degree and the out-degree associated with each node. The in-degree, \( k_i^{\text{in}} \), is simply the number of events pointing to an event, \( a_i \), from events in the past, while the out-degree, \( k_i^{\text{out}} \), is the number of subsequent events that \( a_i \) points to. In the context of recurrence networks, the out-degree is the number of recurrences of a given event and the in-degree is the number of times a given event is the recurrence of the preceding events. The in-degree and out-degree histograms for the three regions considered here are shown in Fig. 4a, b, and c. We compare them with Poisson distributions with the same respective average, which is the expected shape for both the in-degree and the out-degree distribution of a process with trivial spatiotemporal clustering in the limit of large event numbers (Davidsen et al., 2008). Such a process corresponds to the case of a catalogue with randomly occurring (in space and time), independent events, which only has trivial spatio-temporal correlations in the sense that they directly follow from the independent spatial and temporal properties of the event catalog. Note that for a directed graph, the average in-degree has to be the same as the average out-degree but the in-degree or the out-degree distributions themselves can be very different. In fact, this is the case for the directed graphs of all three regions. In addition, both the in-degree and out-degree distribution differ slightly from the expected Poisson behavior in each region. Most strikingly, the frequency with which number of nodes with small out-degree and large out-degree occur exceeds the frequency predicted by the Poisson distributions for all three regions. These findings indicate the presence of non-trivial spatiotemporal clustering.

4.2 Clustering

The topology of directed graphs can also be quantified in terms of its clustering. The clustering coefficient of a network quantifies how well connected the neighbors of a node are among themselves (Fagiolo, 2007). In the case of recurrences, it refers to the likelihood that recurrences of the same event are also recurrences of each other. In this context, it is important to note that there is a clear distinction between the notion of network clustering (Fagiolo, 2007) and communities/cliques (Fortunato, 2010). In graph theory, “communities” arise in the context of graph partitioning (see e.g. Fortunato, 2010) and correspond to more or less well-separated subgraphs.

While the presence of distinctive areas with vastly different seismic activity in the three earthquake regions studied (Fig. 2a, b and c) already suggests that the respective recurrence networks should possess a relatively high clustering coefficient, this needs to be quantified by a comparison with the random uncorrelated case. As Davidsen et al. (2008) noted, however, there is not one unique definition of the clustering coefficient. Here, we choose to follow the approach of Fagiolo (2007), which is particularly well-suited for directed graphs. The approach is based on the notion of triads or triangles. In the triad representation of a complex network (Fig. 5), all possible directed triangles formed by the nodes, \( i, j, h \), lead to four patterns, namely “cycle”, “middleman”, “heart”, and “star”.
Fig. 4. In-degree and out-degree histograms for binary directed recurrence networks along with the Poisson probability distribution. (a) Region 1; (b) Region 2; (c) Region 3. The average degrees computed are 5.95, 4.61, and 5.47 for Region 1, Region 2, and Region 3, respectively for magnitude threshold of 2.

Fig. 5. Possible patterns and graphs, adapted from Fagiolo (1997). The “cycle” configurations do not contribute to the clustering coefficient.

“in”, and “out” (see Table 1 in Fagiolo, 2007). A separate clustering coefficient can then be assigned to each of the patterns. Here, the clustering coefficient is defined as the ratio between the number of triangles of that pattern actually formed by $i$ and the total number of triangles of that pattern that $i$ can possibly form. Since no recurring event of any source event in a triad can go back to the source, “cycle” patterns never occur for recurrence networks. We restrict ourselves to the contribution of the “middleman”, “in” and “out” patterns to the clustering coefficient at each node.

Figure 6a, b and c depict the distribution of the clustering coefficients for the three regions. For comparison, we have included the MC1 results. In all three cases, the average clustering coefficient of the original recurrence networks is at least one order of magnitude larger than that for MC1. This is a clear sign that one cannot simply separate spatiotemporal correlations between earthquakes into spatial and temporal correlations without losing significant information about the underlying microscopic dynamics.

4.3 Temporal separation of recurrences

In addition to the topological properties of the recurrence networks, we also examine the distribution of the time-intervals (or waiting-times) between events and all their recurrences. For example, given an event $j$ and its three recurrences $k$, $m$ and $n$, we consider $t_{jk}$, $t_{jm}$, $t_{jn}$ as the valid time-intervals (or waiting-times). The analysis gives rise to a suite
Fig. 6. Clustering coefficients of binary directed recurrence network of events using Fagiolo’s approach (1997). The clustering coefficient used here is the sum of “in”, “middleman”, and “out” contributions of the triad discussed in the text. (a) Region 1; (b) Region 2; (c) Region 3.

of probability density functions, $p^m(t)$, for different time-intervals (or waiting-times), $t$, for the three regions (Figs. 7a, b, c, and 8a, b, c), where $m$ denotes the applied magnitude threshold.

Fig. 7. Distribution of time-intervals between events and their recurrences for the full data sets, the recent half (Fhalf) and the early half (Lhalf) of the data sets for magnitude threshold of 2. For comparison, we also show the MC1 results. (a) Region 1; (b) Region 2; (c) Region 3.
In the case of Region 1 (Figs. 7a and 8a) and Region 2 (Figs. 7b and 8b), the overall behavior of $p_m(t)$ is very similar for all considered (sub-)catalogues that include the original catalogs and the MC1 catalogs. For time-intervals less than one hour, there is a rather wide scatter of data. While the scatter is due to poor statistics, the overall decay for increasing $t$ over this range of time-intervals is related to temporarily-increased seismic activity as characterized by the Omori law (Ogata, 1998). For $1 < t < 100$, $p_m(t)$ is roughly constant. For intermediate time-interval values, $200 < t < 10000$, the $p_m(t)$’s show a scaling behaviour. The least-squares fit to the data in this range of time-interval values lead to power-law exponents of 1.05 and 1.10 for Regions 1 and 2, respectively. For longest time-intervals, $t > 10000$, there is an observational cut-off since the duration of the catalogue is finite. It is important to note that the functional characteristics of $p_m(t)$ are similar for the different catalogs, for segments of the catalogs, and also for the MC1 catalogs. Small differences in $p_m(t)$ with varying magnitude-thresholds for $1 < t < 100$ in Region 1 (Fig. 8a) and Region 2 (Fig. 8b) can be attributed to poor statistics. In particular, the number of earthquakes in Region 2 above magnitude 3 is simply insufficient to allow one to make any significant statements — see Table 2.

While for $t > 500$ the distributions of the time-intervals for Region 3 (Figs. 7c and 8c) show a behavior similar to those for Regions 1 and 2 (Figs. 7a, b, and 8a, b) for all catalogs considered, $p_m(t)$ has clearly a different behavior for smaller time-intervals for the full catalog. In particular, it seems that $p_m(t)$ follows another power law for $t < 500$. Note, however, that there are huge differences in $p_m(t)$ between the first and second half of the catalog for Region 3 (Fig. 7c). In particular, the distribution of the time-intervals for the second half of the catalog is rather similar to the distributions found for Regions 1 and 2. This suggests that there are some temporal variations in Region 3 that might be related to variations in the detection network, as discussed in Sect. 5.

The results presented in Figs. 7 and 8 seem to indicate that there are no differences between the original catalogs and the MC1 catalogs. If this were indeed true, it would imply that the distribution of time-intervals is fully determined by the spatial and the temporal correlations between earthquakes and, thus, insensitive to spatiotemporal correlations. To test this hypothesis, we consider MC2 catalogs, which allow us to study the distribution of time-intervals for different magnitude-thresholds $m$. The corresponding $p_m(t)$’s are shown in Fig. 9 for the different regions. It is evident that the distributions vary with $m$, in sharp contrast to the original catalogs shown in Fig. 8. Hence, the observation that $p_m(t)$ is independent of $m$ within the statistical uncertainties

---

1 The fits were obtained by standard linear regression on the logarithm of the data. The data were only fitted over “linear” segments which were chosen based on visual inspection. The latter leads to systematic uncertainties such that it is difficult to give error bars or confidence intervals. A rough estimation based on a graphical analysis gives error bars of about ±0.1 for the exponents.
Fig. 9. Distribution of time-intervals between events and their recurrences for different threshold magnitudes obtained for the MC2 catalogs. (a) Region 1; (b) Region 2; (c) Region 3.

is the same for the original catalogs and the MC2 catalogs. As shown in Davidsen et al. (2008), this exponent is generally expected and related to our definition of recurrences as record-breaking events.

### 4.4 Spatial distances of recurrences

In addition to the distribution of the time-intervals, we also consider the spatial distances, $l$, between events and all their recurrences (see Fig. 3) for the different catalogs. This allows us to estimate the corresponding probability density functions, $p^m(l)$, which are shown in Figs. 10 and 11 for the three regions of intraplate seismicity considered. For all regions, $p^m(l)$ has a power-law regime for distance larger than the location errors (about 10 km) up to a transition point beyond which we observe cutoff behavior related to the finite spatial size of the observation regions. The transition point between the two different regimes is about 80 km for Region 1 and about 200 km for Regions 2 and 3. While the decay is approximately exponential in the cut-off regime for Regions 1 and 2, the behavior is more complicated in the cut-off regime of Region 3 and could even be divided into two sub-regimes. Note, however, that the decay seems to approach an exponential behavior for catalogs with larger threshold magnitudes (Fig. 11b). This suggests that results for Region 3 partially suffer from catalog incompleteness for lower magnitudes, as is shown in Fig. 11 and further discussed in the Sect. 5. Yet, some of the differences can likely be attributed to the strongly inhomogeneous spatial distribution of seismic activity in that region (Fig. 2). The larger scatter in the cut-off regime for Regions 2 and 3 compared to Region 1 (Fig. 10) is due to fewer events – see Table 1.

For Regions 1 and 2, the exponent in the power-law regime lies in the vicinity of 1 (1.1 for Region 1 and 1.0 for Region 2 as estimated by the least-squares method). In contrast to that, the power-law exponent for Region 3 is about 1.40. Yet, the exponent tends towards a smaller value, 1.29, for $m = 3$ (Fig. 11b) indicating again that the differences between Regions 1 and 2 on one side and Region 3 on the other side might be due to catalog incompleteness (see Fig. 12 and the Sect. 5 for details). It is important to realize that an exponent 1 is generally expected for a spatiotemporal point process without any spatiotemporal correlations (Davidsen et al., 2008). This is confirmed by the observation that there are no significant differences between the distributions of the spatial distances for the original data and those for the MC1 catalogs (Fig. 10). This implies that the observed behavior of $p^m(l)$ is not indicative of spatiotemporal correlations between events.

The dependence of $p^m(l)$ on the threshold magnitude $m$ investigated in Fig. 11 for Regions 1 and 3. Since there are insufficient statistics for Region 2 for magnitude equal to or greater than 2.5, we do not include a figure here. For Regions 1 and 3, we find that the $p^m(l)$ curves vary only weakly with $m$. The only noticeable effects are small
changes for distances less than about 10–20 km, which are roughly comparable to the uncertainty in the event locations. With increasing \( m \), these changes lead to the emergence of a constant regime. This effect is particularly pronounced in Region 1.

5 Detection network and catalog completeness

A marked difference in the shape of probability distributions, \( p(l) \) and \( p(t) \), between Region 3 and the other two regions raised a question if there were differences in the data collection conditions over the historical time period of seismicity in these regions (see Table 1). Figure 12 shows calculated magnitude detection threshold for two scenarios, before and after installation of seismograph stations for the Canadian POLARIS initiative (Eaton et al., 2005). Installation of the POLARIS stations commenced in 2002 within various regions of Canada. These stations were used, as they became available, for the construction of the earthquake catalogue used for this study (Wetmiller et al., 1988). The magnitude detection threshold maps were calculated for a frequency of 10 Hz using the attenuation model and earthquake source model of Atkinson and Boore (1995). The earthquake source model was modified to include earthquakes
of moment magnitude $M_w < 4$ using the single corner frequency formula of Brune (1970) with a stress drop of 55 bars. This value of stress drop was selected to ensure continuity for the calculated values of ground motion for $M_w = 4$. The threshold value for background noise was selected based on observed 10 Hz ambient noise at representative POLARIS stations. The detection cut-off was applied using the rule that the ground motion must exceed the noise cut-off amplitude value for at least 4 stations. This value was chosen since it represents the effective minimum number of stations needed for an event to appear in the earthquake catalog, and so gives a useful guide for magnitude completeness.

The maps show that, prior to installation of POLARIS stations (i.e., during the first half of the time window considered here), most of Region 3 had a high ($M_w > 3.5$) magnitude detection threshold, whereas Regions 1 and 2 had detection thresholds of $M_w = 2.5$ or less. After installation of POLARIS stations (i.e., during a large part of the second half of the time window considered here), all three regions have a relatively uniform and consistent magnitude detection threshold of $M_w = 2.5$ or less. This analysis represents the most extreme difference between these two time periods, as the POLARIS stations were installed over a number of years and were not present throughout all of the second half of the time window. Nevertheless, this improvement in network performance in Region 3 probably may explain the temporal differences in seismicity patterns documented in the present study.

6 Discussion

In this study, we treat intraplate seismicity as a spatiotemporal point process and capture its dynamics using binary directed graphs. Davidsen et al. (2008) used this approach to identify non-trivial features of spatiotemporal clustering of earthquakes in southern California, leading to a new and independent estimate of the rupture length and its scaling with magnitude. Their results also furnished evidence in support of a shadowing effect associated with smaller earthquakes (Rubin, 2002; Fischer and Horálek, 2005; Hainzl and Marsan, 2008). The topological structure of the directed graphs obtained here for intraplate seismicity is similar in some regards to the structure found for southern California. In both cases, there are significant deviations from a Poissonian out-degree distribution and clustering is much stronger than expected in the absence of spatiotemporal correlations between events. Moreover, the distribution of time-intervals between events and their recurrences is, in both cases, independent of the applied magnitude threshold, in sharp contrast to the shuffled catalogs. Since, by construction, the shuffled catalogs only preserve spatial and temporal correlations separately but are void of any spatiotemporal correlations between events, this is a clear indication of the presence of spatiotemporal correlations in seismicity. While also the functional forms of the distribution of time-intervals is similar for our intraplate Regions 1 and 2 and southern California for intermediate to large time scales, the differences for time-intervals less than about 1 h are probably related to the large differences in seismic activity.

Compared to seismicity catalogues for southern California, the seismicity catalogues for eastern and northern Canada suffer from several shortcomings. For example, one of the main findings of Davidsen et al. (2008) was to provide a new and independent estimation of the rupture length and its scaling with magnitude. This result was enabled by the remarkably small location uncertainties (50–100 m) for the southern California catalog. In contrast, due to relatively sparse seismograph distributions, the location errors for the catalogues of the three regions studied here are typically in the range of 10–20 km, which is significantly greater than the rupture length expected for the small events observed in the catalogue. Due to the much lower rates of seismicity, our catalogues also contain far fewer events (by a factor of about 20) than in southern California.

The important question of the potential influence of GIA-induced stresses on seismicity remains unanswered. Compared to the other two regions, we observe tantalizing differences in Region 3 with respect to the structure of the probability distributions of recurrences, $p(t)$ and $p(l)$. These differences invite speculation that they arise from GIA, since deglaciation occurred more recently in Region 3 than in Regions 1 and 2. Given changes in the magnitude detection threshold due to installation of new seismograph stations for the POLARIS projects, further study is needed to test this model.
7 Conclusions

In this paper we have applied analytical tools of spatiotemporal recurrence statistics to regions of intraplate seismicity in eastern and northern Canada, using earthquake catalogue data for the time-interval 1990–2009. Through comparisons with equivalent analyses using either a model of random events or Monte Carlo shuffling of the catalogue data, our results provide strong evidence in support of non-trivial spatiotemporal correlations of seismicity in all three regions considered. This evidence includes:

For all three regions, the in-degree and out-degree parameters significantly exceed Poissonian distributions near the tails of the distributions (i.e., far away from the mean).

Using clustering coefficient as defined in Fagiolo (2007), the average clustering coefficient of the original recurrence networks is at least an order of magnitude greater than the location-shuffled catalogue, MC1.

The temporal probability distribution \( p^m(t) \) is independent of \( m \) in the original catalogues, but not in catalogues where both location and magnitude are shuffled (MC2).

Similar evidence for spatiotemporal clustering has been documented using seismicity catalogues for southern California, suggesting similarities in underlying earthquake dynamics in both regions (Canada and California) despite differences in tectonic setting (stable continental interior versus tectonically active plate boundary).

In Regions 1 and 2, the temporal probability distribution, \( p(t) \), has a structure that is roughly constant for the interval \( 1 < t < 100 \) h and exhibits normal power-law behavior with exponent \( \sim 1 \) for \( 200 < t < 10^4 \) h. When the entire catalogue is considered, Region 3 does not show the roughly constant behavior for \( 1 < t < 100 \) h, but this behavior begins to emerge when only the second half of our time window (2000–2009) is considered. In terms of the spatial probability distribution, \( p(l) \), Regions 1 and 2 exhibit normal power law behavior with an exponent close to unity for correlation distances greater than the location uncertainty (\( \sim 10 \) km) and less than a cut-off imposed by spatial dimensions of the study area. Region 3, on the other hand, exhibits power law behavior with an exponent of \( \sim 1.4 \) in this range. The more northern location of Region 3, where removal of the lithospheric load due to Laurentide glaciation occurred much later than in Regions 1 and 2, invites speculation that these apparent differences in graph structure may reflect underlying differences in patterns of seismicity related to glacial isostatic adjustment (GIA). However, differences between Region 3 and the other two areas may be due to seismograph network changes arising from installation of POLARIS stations in this area commencing in 2002.

Acknowledgements. The authors would like to acknowledge the Natural Sciences and Engineering Research Council of Canada for financial support. They would like to thank Earthquakes Canada for the use of the earthquake data used in this study (http://earthquakescanada.nrcan.gc.ca/stndon/NEDB-BNDS/bull-eng.php). The authors express gratitude to C. Woodgold of Earthquakes Canada for assistance with the earthquake catalogue and Phil Hammer of the University of British Columbia for the tectonics map of Canada. They would like to thank one of the referees for helpful comments.

Edited by: I. Zaliapin
Reviewed by: M. Baeisi and another anonymous referee

References


www.nonlin-processes-geophys.net/17/513/2010/


Marsan, D. and Lenglé, O.: Extending earthquakes’