Title: Using spatial Markovian chain for the statistical analysis of seismic occurrences in the Azores Region

Abstract

The objective of this paper is to study the spatial sequence of the epicentral locations of seismic occurrences (Space variable) in the Azores region. In a previous investigation based on geological considerations and statistical criteria that addresses the existing historical and instrumental information, the epicentral locations were assigned to seven seismic zones.

This paper focuses on the analysis of occurrences in the seven seismic zones using Markovian chains.

The probability of occurrence of an earthquake in one of the seven adopted zones is estimated, revealing great differences among the seismic zones.

Additionally, the one-step transitions of this variable (seismic zone for the next occurrence in time) are explored and show an evident dependence between consecutive earthquake locations. Assuming that the process is stationary, N-step transitions are also discussed. Based on the developed Markovian model, this study also simulates the sequence of epicentral zones where occurrences can take place, which is an important component of modelling the entire process of seismic occurrences in the Azores region.

1 Introduction

The aim of this study is to estimate the probability of occurrence of an earthquake in each of the seven distinct seismic zones in the Azores region (as explained later) based on the existing historical and instrumental information and also on geological data with special emphasis on spatial location "memory". This goal also includes evaluation of the influence of the location of an earthquake on the location of the next earthquake, i.e., the first order (or one-step) dependence of the Space variable. We also want to quantify the influence of the location of an earthquake not only on the next earthquake but after N seismic events. N-step transitions are also analyzed, and Markovian chains are applied for this purpose.

This study also simulates how seismicity "moves" from zone to zone, which is an important component of the entire process of seismic occurrences in the Azores region and a topic of ongoing study.

Traditionally, seismic phenomena have been described using Poisson models, but these models do not have "memory", a fact that contradicts the reality. There are many ways to analyze the seismic phenomena, using physical or mathematical models.

After an earthquake, it is common to ask where the next earthquake will occur. This study attempts to give a statistically sound answer to this question.

The Azores Archipelago is renowned due its intense seismicity, a characteristic that can be explained by its location at the triple junction of the Mid-Atlantic Rift, where the Eurasian, Nubian, and American Plates meet. The consequent seismic activity in this region is the result of the existence of active complex geological structures such as rifts, trenches, volcanoes, banks and faults.

This intense seismic activity has been studied by several authors, including Bezzeghoud et al. (2008), Borges et al. (2008), Carvalho et al. (2001) and Nunes et al. (2000). Other authors such as Kagan et al. (2010) and Reiter (1991) also studied this particular region.

The Azores Archipelago consists of nine islands distributed among three different groups: the islands of Flores and Corvo, which constitute the Western Group; the islands of Terceira, Graciosa, São Jorge, Faial and Pico, which are components of the Central Group; and the islands of São Miguel and Santa Maria that form the Eastern Group.

Modelling of natural phenomena, especially seismic processes, is a complex task that has been attempted by several authors in different areas of the world. As examples, Nanjo et al. (2011) presented selected models for Japan that are still undergoing testing and development. Other authors focused on tectonic aspects, such as Burford et al. (2000), who built a 3-D subsurface model, and still others have focused on prediction of damages caused by earthquakes (Sopra and Patrizi 1987) to better understand consequences to buildings.

Other models appeal to "memory", such as Markovian models, but these models have not experienced great development. Kirimidjien and Anagnos (1984) first attempted to implement such models in the framework of California, but these studies were restricted to the same seismic region, and the results obtained were never transposed to practice.

Memory is present in several natural phenomena and is an important aspect in simulation of natural processes that are not independent. For example, Russo and Soares (2014) used conditional space simulations in the context of urban air pollution forecasting.

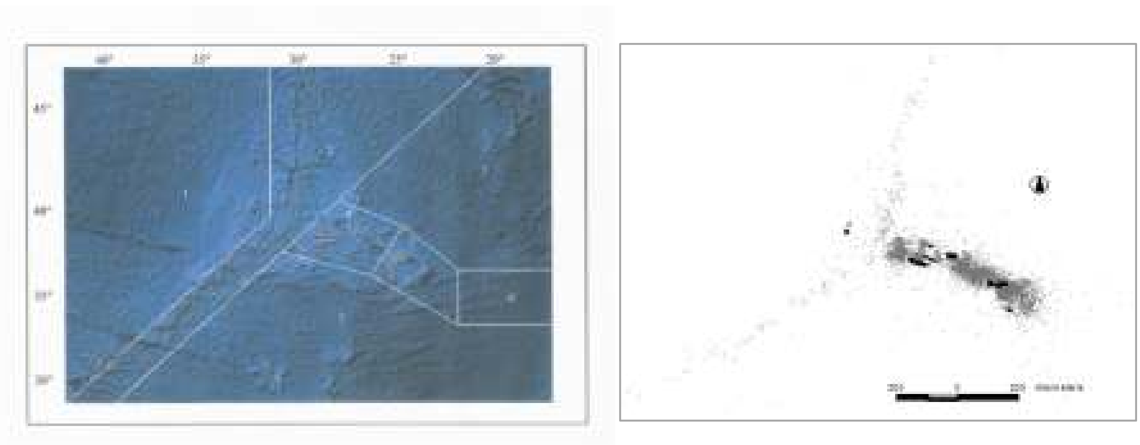
Recently, Cavers and Vasudevan (2015) presented a space-time Markovian chain to represent a global model for earthquakes sequences. These researchers recognized that the sequence of epicenters has "memory", that is the location of an epicenter may constrain the location of the next event and this fact is of fundamental importance.

We aim to estimate the probability of earthquake occurrence in a target seismic zone using knowledge of the seismic zone of the previous event. The process can be repeated, thus allowing the simulation of the sequence of seismic locations.

Cavers and Vasudevan (2015) used a different methodology for macro-earthquake zonation (i.e., active continent, trenches, etc.,) and applied weights for the state-to-state transition probabilities.

Each earthquake occurrence can be characterized by the three variables of Time, Size and Space. The variable Time (Dt) is defined by the time intervals between consecutive events, the variable Size (S) is the Richter magnitude (M_L) associated with an earthquake, and the Space variable (Sp) represents the zone in which the epicenter of the earthquake is located.

Rodrigues and Oliveira (2013) defined 7 distinct seismic zones in the Azores Region based on statistical and geological information and verified that the magnitude and time between consecutive earthquakes differ significantly among these seismic zones. Figure 1 shows the seven defined zones together with the epicenter locations and a zoom for the Azores Archipelago with the main Geological features.



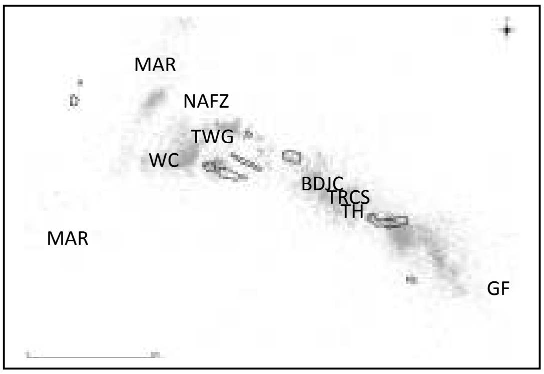


Fig. 1 (a) Schematic representation of the 7 defined seismic zones proposed by Rodrigues and Oliveira (2013), (b) Epicentral locations, (c) Zoom of Azores Archipelago - Mid-Atlantic Ridge (MAR); West of Capelinhos (WC); North Azores Fracture Zone (NAFZ); Bank D. João de Castro (BDJC); Trench Hirondelle (TH); Trench West of Graciosa (TWG); Terceira Rift Central Sector (TRCS); Gloria Fault (GF).

This 7-zone definition is adopted in this work, and the Space variable is represented as $Sp \in \{1, 2, 3, 4, 5, 6, 7\}$.

2 Data

For the Azores region, the available data are collected from two different sources. The catalogue of Nunes et al. (2004) is the source of data for the period 1915-1998, and for the period 1999-2011, the data are directly obtained from the site of Instituto Português do Mar e da Atmosfera (I.P.M.A. - 2011).

The first period covers the area encompassed by longitude 11.50º W – 42.86º W and latitude 10.80º N – 47.54º N. A total of 9214 records are available, of which 5456 contain information on magnitude according to the Richter scale (M_L).

The catalogue of the second period covers an area delimited by longitude 21.31º W – 35.42º W and latitude 34.3º N – 45.57º N containing 9608 events, all of which contain magnitude information (M_L).

A total of 18822 seismic records are available with information on Time and Space, 15064 of which contain information on magnitude according to the Richter scale.

The data were analyzed as a whole, including foreshocks and aftershocks. Fig. 2 shows a Gutenberg–Richter plot, which indicates that the dataset is not complete. Many small-magnitude events occur in the sea, far from the seismic network, and thus are not recorded. According to the Gutenberg–Richter law, a linear trend should exist between Log N and m:

$$\text{Log } N(m) = a - b \times m, \qquad (1)$$

where N is the number of events of magnitude greater than m, and a and b are constants fitted to the data.

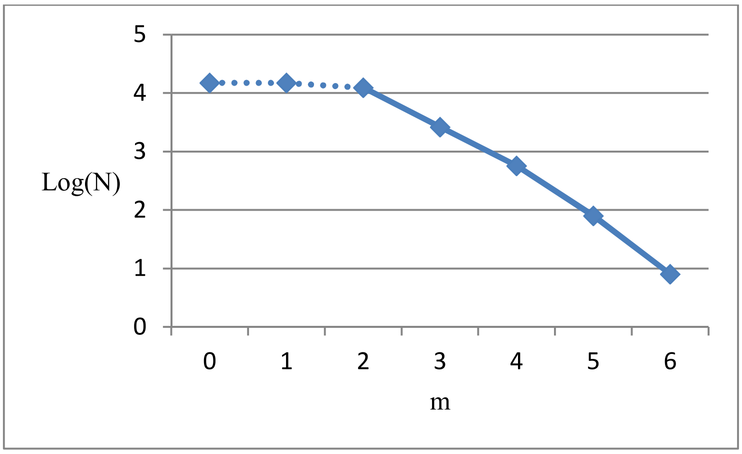


Fig. 2 Gutenberg-Richter plot for the Azores Region based on the entire data-base.

Removing earthquakes smaller than magnitude 2, a least squares approximation leads to

$$\text{Log } N(m) = 5.77611 - 0.79\, m, \qquad (2)$$

with a correlation coefficient $R = -0.996$, which indicates a significant linear correlation and that the catalog is complete for earthquakes with magnitude larger than 2.

Data are not constrained to "completeness of data for the low magnitude values" or to "filtration of aftershocks", that is all seismic events are considered.

3 Methodology

The current study is performed in the following stages:

1. The Space variable is characterized among the seven defined seismic zones, and the main statistics of Sp are determined.

2. The "memory" effect is analyzed. The influence of the seismic zone of an earthquake in the location (seismic zone) on the next seismic event is quantified by estimating the conditioned probability of occurrence of an event in each one of the seven seismic zones and knowing the location of the epicenter (seismic zone) of the last occurrence.

$Sp_{t+1} = j$ signifies that at epoch t+1, an event took place in zone j, and $Sp_t = i$ represents that at epoch t, an event took place in zone i. $P\ [\ Sp_{t+1} = i \mid Sp_t = j\ ]$ is the probability a seismic occur in zone i at epoch t+1 given than an earthquake have occurred in zone j at epoch t.

This step includes the definition of seven statistical conditioned distributions:

$$P\ [\ Sp_{t+1} = i \mid Sp_t = 1\ ],$$

$$P\ [\ Sp_{t+1} = i \mid Sp_t = 2\ ],$$

$$P\ [\ Sp_{t+1} = i \mid Sp_t = 3\ ],$$

$$P\ [\ Sp_{t+1} = i \mid Sp_t = 4\ ],$$

$$P\ [\ Sp_{t+1} = i \mid Sp_t = 5\ ],$$

$$P\ [\ Sp_{t+1} = i \mid Sp_t = 6\ ],$$

$$P\ [\ Sp_{t+1} = i \mid Sp_t = 7\ ],\ \ i \in \{1, 2, 3, 4, 5, 6, 7\}.$$

These formulations represent the "one-step" transition probabilities. The "N-step" transitions are also referred.

3. Assuming that the sequence of epicentral locations is stationary in time, a simulation model for the Space sequence is built and generated according to the corresponding $P [Sp_{t+1} = j \mid Sp_t = i], i, j \in \{1, 2, 3, 4, 5, 6, 7\}$.

To evaluate the quality of the generated Space sequence, samples of generated sequences are compared with the Space data using goodness-of-fit statistical tests.

This methodology is a component of an ongoing study for modelling the seismic process of occurrences in the Azores region, and the main steps are represented in Fig. 3.

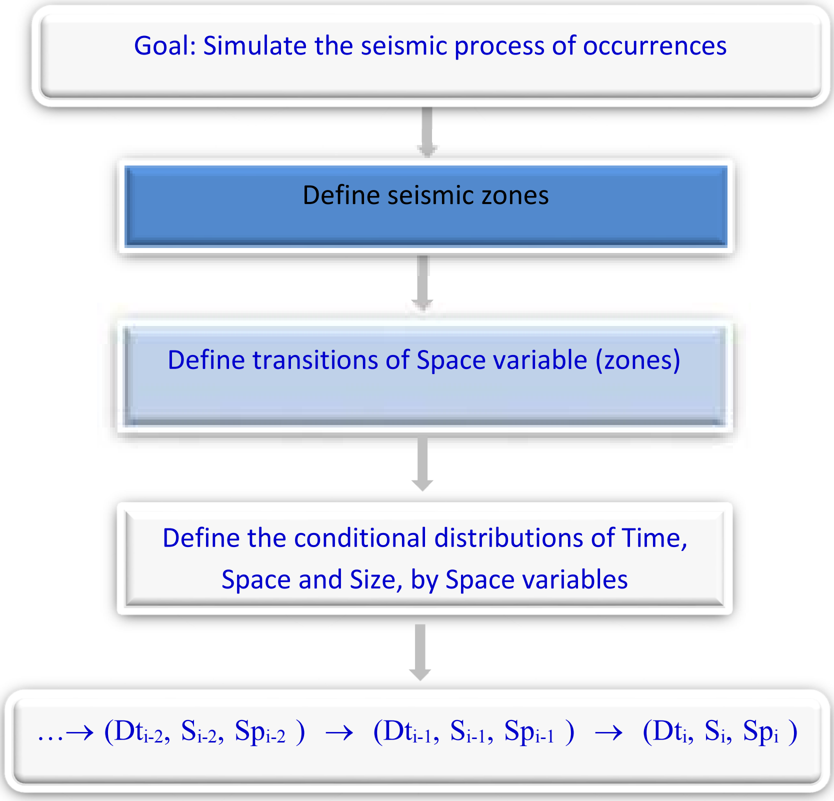


Fig. 3 Schematic representation of the global process of occurrences. The dark blue box indicates previous achievements, the light blue box denotes the subject of this paper, and the white boxes portray work in progress.

The statistical software R® (see, e.g., Dalgaard 2008 or Venables et al. 2011) and the Turbo Pascal® language are used to implement the procedure described above.

4 Space data analysis

4.1 Space variable analysis

Based on the data, the main statistics of Sp are computed and presented in Table 1, and the corresponding histogram of probability density function is shown in Fig. 4.

Table 1. Statistics of Sp

	Statistics of Sp
Mean	4.38
Standard deviation	1.02
Skewness	-1.35
Kurtosis	4.55
Minimum	1
quantile	
0.1	2
0.2	4
0.3	4
0.4	4
0.5	5
0.6	5
0.7	5
0.8	5
0.9	5
1 (max.)	7
Total number of records	18 822

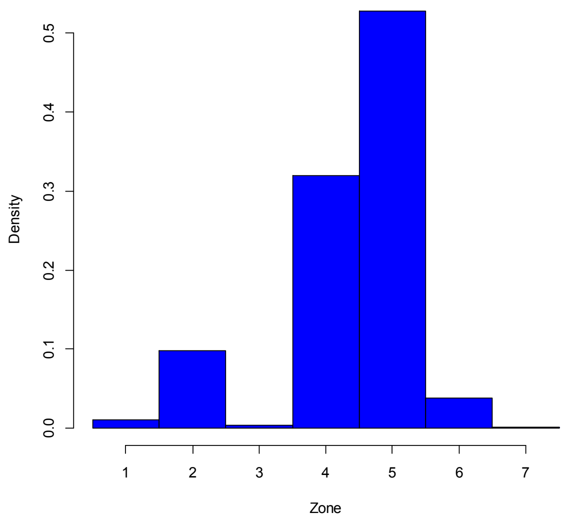


Fig. 4 Histogram of probability density function of Sp.

As shown, the frequency of seismic events in each one of the seven adopted zones is quite different, highlighting the great dispersion of Sp among the seven zones.

Additionally, zones 1, 3 and 7 include small numbers of events compared with the other seismic zones, and for this reason, they are considered to be background zones of seismicity.

Zones 4 and 5 have the highest number of occurrences, followed by zone 2.

Zone 2 is a maritime zone corresponding to the Mid-Atlantic Ridge and its transform faults to the north. This zone also includes the North Azores Fracture Zone.

Zone 5 consists of the Eastern Group of the Archipelago, the Hirondelle Trench, the D. João de Castro Bank and the two islands of São Miguel and Santa Maria.

Zone 4 encompasses the Central Group of the Archipelago west of Capelinhos and the Terceira Rift central sector with five islands.

Zone 6 is a maritime zone and includes the Gloria Fault to the East. The number of seismic events is moderate, but this zone has the highest magnitude of all zones: M_L8.2. The number of seismic events is not high but the magnitudes can be high. (For details on these morphological structures, see Madeira et al., 2015).

4.2 One-step transitions of space variable analysis

It is known that the zone of an epicenter might constrain the zone of the next earthquake (see for example, Cavers and Vasudevan 2015).

The objective is to analyze the statistical distribution of the Space variable (Sp) and the one-step transitions of this variable. If an earthquake takes place in a seismic zone, the aim is to answer the question: "which is likely to be the next seismic zone?"

Markovian chains are suitable for modelling this phenomenon (see, for example Ravindran et al. 1987).

The possible values of Sp are considered "states", and each change of state is a transition. The points in time at which the system is observed (in this case, the seismic events) are the epochs.

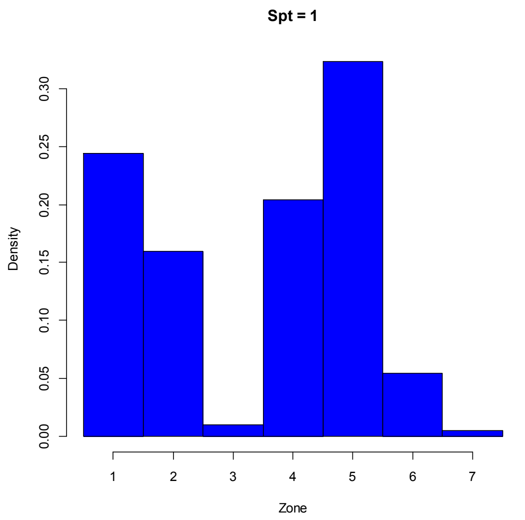
~~$Sp_{t+1} = j$ signifies that at epoch t+1, an event took place in zone j, and $Sp_t = i$ represents that at epoch t, an event took place in zone i.~~

Let p_{ij} be the conditional probability $P\,[\,Sp_{t+1} = j \mid Sp_t = i\,]$, i.e., the one-step transition probability, $i, j \in \{1, 2, 3, 4, 5, 6, 7\}$.

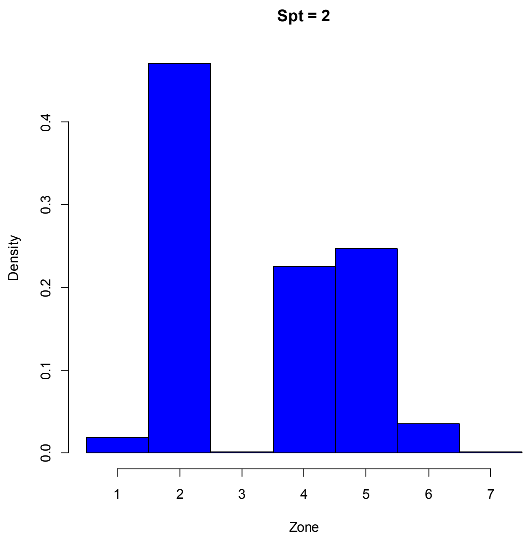
Based on the data, the one-step transition matrix T was computed for estimation of the seven empirical conditioned distribution functions $Sp_{t+1}|Sp_t = i$, $i \in \{1, 2, 3, 4, 5, 6, 7\}$. Figures 5(a) to (g) display the histograms of corresponding probability density functions.

$$T = \begin{bmatrix} 0.244 & 0.159 & 0.01 & 0.204 & 0.323 & 0.055 & 0.005 \\ 0.019 & 0.472 & 0.001 & 0.225 & 0.247 & 0.035 & 0.001 \\ 0.062 & 0.046 & 0.015 & 0.138 & 0.709 & 0.016 & 0.015 \\ 0.006 & 0.072 & 0.002 & 0.633 & 0.251 & 0.034 & 0.002 \\ 0.007 & 0.045 & 0.005 & 0.154 & 0.753 & 0.035 & 0.001 \\ 0.011 & 0.094 & 0 & 0.273 & 0.508 & 0.114 & 0 \\ 0 & 0.12 & 0.04 & 0.32 & 0.48 & 0.04 & 0 \end{bmatrix} \qquad (3)$$

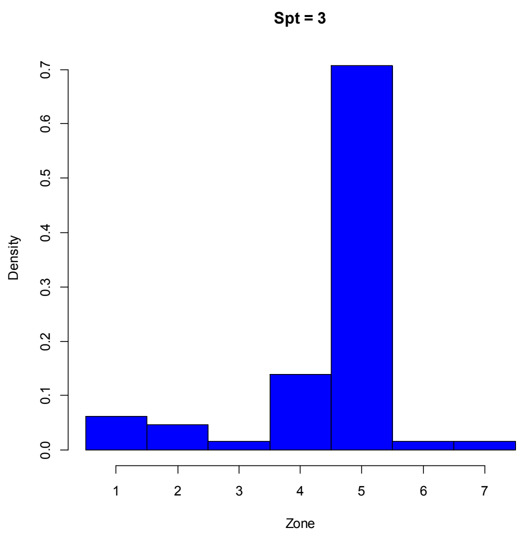
For example, Fig. 5(a) shows the probability of an earthquake occurring in each one of the seven seismic zones knowing that the last earthquake took place in zone 1.


Spt = 1
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
1
2
3
4
5
6
7
Zone

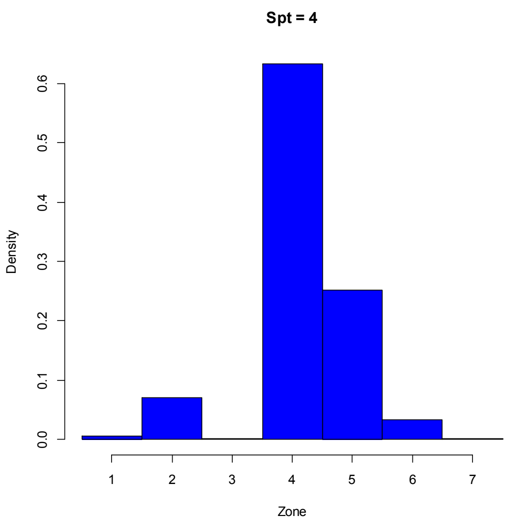
(a)


Spt = 2
Density
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
6
7
Zone

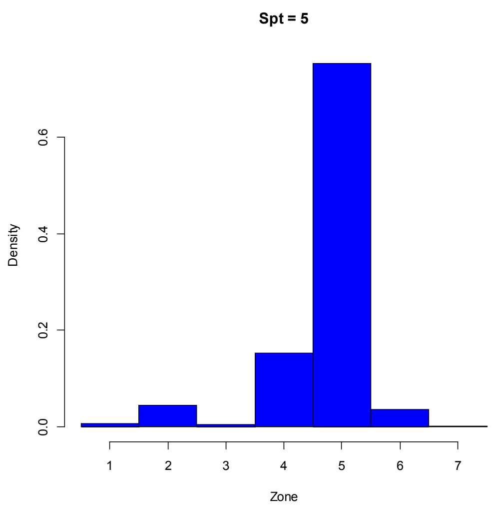
(b)


Spt = 3
Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
1
2
3
4
5
6
7
Zone

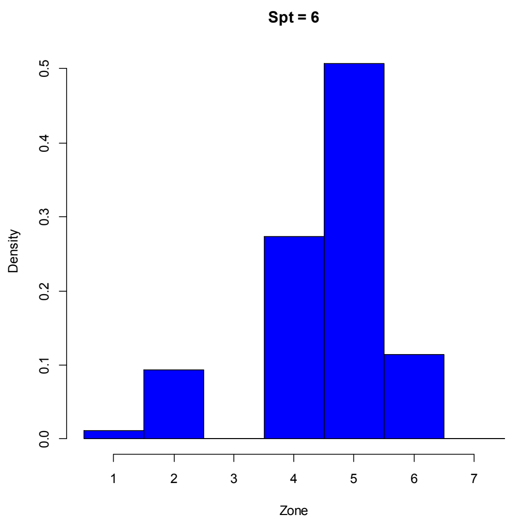
(c)

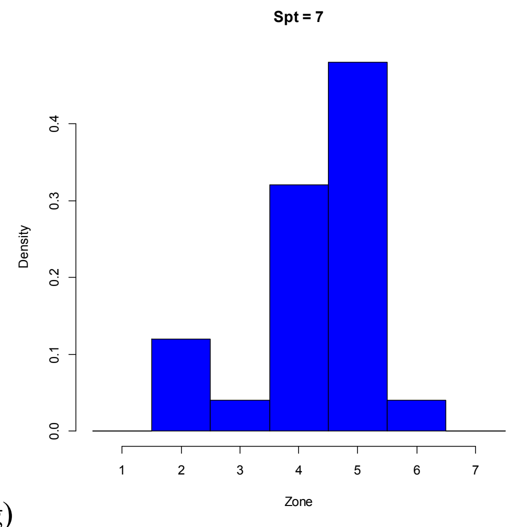

Spt = 4
Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1
2
3
4
5
6
7
Zone

(d)


Spt = 5
Density
0.0
0.2
0.4
0.6
1
2
3
4
5
6
7
Zone

(e)


Spt = 6
Density
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
6
7
Zone



(f)g)

Fig. 5 Histogram of probability density function of: (a) $Sp_{t+1}|Sp_t=1$, (b) $Sp_{t+1}|Sp_t=2$, (c) $Sp_{t+1}|Sp_t=3$, (d) $Sp_{t+1}|Sp_t=4$, (e) $Sp_{t+1}|Sp_t=5$, (f) $Sp_{t+1}|Sp_t=6$, (g) $Sp_{t+1}|Sp_t=7$.

The seven conditioned distributions are highly different. The most similar to the Sp distribution function is the distribution of $Sp_{t+1}|Sp_t=5$, which is not surprising given that zone 5 contains the greatest number of seismic occurrences.

If a seismic event occurs in zone 1, the most likely target zone is zone 5, which means that the next earthquake will not be an aftershock. Zone 1 is a background zone with few seismic events. Zone 1 comprises the Western Group of the Azores Archipelago and is situated NW of the Mid-Atlantic Ridge. The islands of Flores and Corvo are in this zone.

Notice that in eq. (3), in the first line of matrix T, the largest element is t_{15}, that corresponds to a transition from zone 1 to zone 5, but the other elements of the first line are not zero, that is, with small likelihood another transition may occurs. For example, an aftershock may occur with a probability of 0.244 (t_{13}).

If a seismic event occurs in zone 2, the most likely target zone is the same zone, which means that in zone 2, a great probability exists that the next earthquake will be an aftershock. This zone corresponds to the Mid-Atlantic Ridge, an active zone. It seems reasonable that an event in this zone may be followed by another near it. In fact observing the second line of matrix T, $t_{22} = 0.472$ is the largest value of the second line. A transition to the others active zones, as zone 4 and zone 5, has a likelihood of 0.255 and 0.247. On the other hand, an earthquake in zone 2 followed by an event in zone 7, a background zone, has only 0.001 of likelihood.

If a seismic event occurs in zone 3, the most likely target zone is zone 5, which means that the next earthquake will probably not be an aftershock. Zone 3 is a background zone with a small number of earthquakes. As expected, the most probably next event will be in another zone.

If a seismic event occurs in zone 4, the most likely target zone is the same zone. Again, in zone 4, a great possibility exists that the next earthquake will be an aftershock. This is an active zone. Zone 4 encompasses the Central Group of the Archipelago, islands of Faial, Pico, São Jorge, Terceira and Graciosa. Also comprises west of Capelinhos and the Terceira Rift central sector. It features very high seismicity. As certain, the next earthquake will be an aftershock and the seismicity "moves" to another zones with low probability, except to the neighbor zone 5.

If a seismic event occurs in zone 5, the most probable target zone is the same zone. Again, in zone 5, a great possibility exists that the next earthquake will be an aftershock, $t_{55} = 0.753$ is the largest value of T matrix.

Zone 5 comprises the Eastern Group of the Archipelago, containing two islands: São Miguel and Santa Maria, the Hirondelle Trench and the D. João de Castro Bank.

These geological structures explain the fact that this zone has the highest seismicity of all seven zones. It is expected that the seismicity "moves" to another zones with low probability.

If a seismic event occurs in zone 6, the most likely target zone is zone 5, which signifies that the next earthquake will probably not be an aftershock.

Zone 6 is a maritime zone and includes the Gloria Fault. In this zone the seismicity is moderate, but this zone has the highest magnitude of all zones: 8.2. It is characterized by a moderate number of earthquakes, which can be of relatively high magnitude. Gloria Fault justifies the high magnitudes but seismicity is moderate, the most expected next zone is not the same, but zone 5.

Zone 7 is a maritime zone and is the furthest south of all seismic zones. It has the lowest seismicity.

If a seismic event occurs in zone 7, the most likely target zone is zone 5, which signifies that the next earthquake will probably not be an aftershock. Zone 7 is a background zone with a small number of earthquakes.

Fig. 6 presents these data in the form of “arrows” and loops” or a scheme of the most likely one-step transitions for each seismic zone.

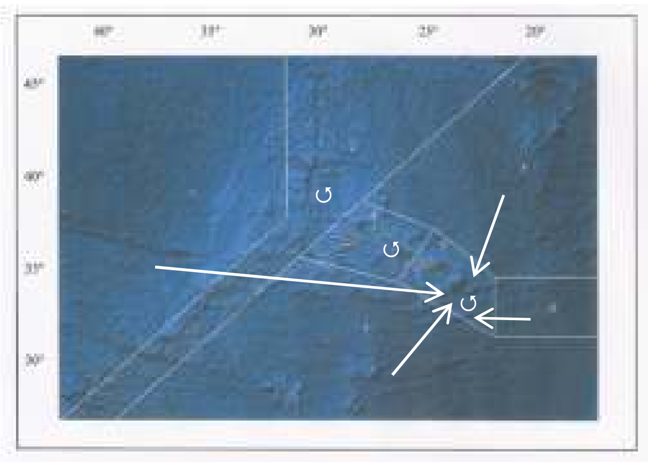


Fig. 6 Scheme of the most likely one-step transitions for each seismic zone.

It can be noted that in zones with great number of occurrences, active zones (zones 2, 4 and 5), the most probable case is that the next seismic event (event) will take place in the same zone.

In zones with few earthquakes, the most probable case is that the next seismic event will occur in a different zone because zones 1, 2, 6, and 7are located far from the islands, and if the aftershocks were weak, they might not be recorded.

We can notice that a characteristic of Azores region seismicity is that in the actives zones (zones 2, 4, 5 and 6) a seismic event will probably followed by an event in the same seismic zone, in background zones (zones 1, 3 and 7) a seismic event will probably followed by an event in another seismic zone.

On the other hand background zones are located far from the islands, and if the aftershocks were weak, they might not be recorded.

These results show that the sequence of events is not independent of the preceding earthquake.

We note that the current results represent only one-step transitions, and we must quantify the N-step transitions.

We assume that the transition probabilities do not change with the passage of time. In this situation, if we use Markovian chains to model the sequence of epicenters, the Markovian chain will be stationary.

For a stationary Markovian chain, the matrix T creates the one-step transition probabilities for any time and is sufficient to describe the entire process, i.e., the entire sequence of epicentral seismic zones.

For example, if a seismic event has occurred in zone 2 at epoch t, i.e., $Sp_t = 2$, and we desire to estimate the probability of $Sp_{t+2} = 5$, this is a two-step transition, $p_{ij}^{(2)}$. Using matrix T, $p_{52}^{(2)}$ can be computed.

$$p_{52}^{(2)} = p_{51}.\ p_{12} + p_{52}.\ p_{22} + p_{53}.\ p_{32} + p_{54}.\ p_{42} + p_{55}.\ p_{52} + p_{56}.\ p_{62} + p_{57}.\ p_{72} \quad (4)$$

By generalizing, it is possible to obtain the N-step transition matrix (see, for example, Ravindran and Dolberg, 1987). Given that the Markovian chain is stationary:

$$TM^{(n)} = TM^{n}. \quad (5)$$

This procedure allows reproduction of the seismic epicentral sequence.

5 Space generation

Space variables can be generated, i.e., a sample of values of the random variable can be produced and compared with the data sample. First, the aim is to only test whether the proportion of the number of generated events in each zone matches the corresponding proportion in the data. Subsequently, the sequence of the generated sample must be checked.

5.1 Generation of Space variable

The number of earthquakes in each seismic zone was computed based on the data, allowing estimation of the probability of a seismic event occurring in each one of the seismic zones.

The inverse–transform method and the statistical distribution function of Space were used to generate pseudo-random values of Space (see, for example, Pidd (1994) or Rubinstein and Melamed (1998)).

We use Spg as the random variable that represents the generated values of Space.

Table 2 shows the number of seismic events for each zone in the data and in a generated sample of size 19 000. Fig. 7 presents the histogram of probability density function of Spg.

Table 2 Number of generated seismic events for each seismic zone in the data and in a generated sample.

Zone	1	2	3	4	5	6	7	Totals
Data	201	1847	65	6009	9948	727	25	18822
Generate Sample	204	1892	60	6148	9953	730	13	19000

The generated sample and the data samples should exhibit the same features. A goodness-of-fit test was used to check whether the generated random values of Space fit the corresponding data values.

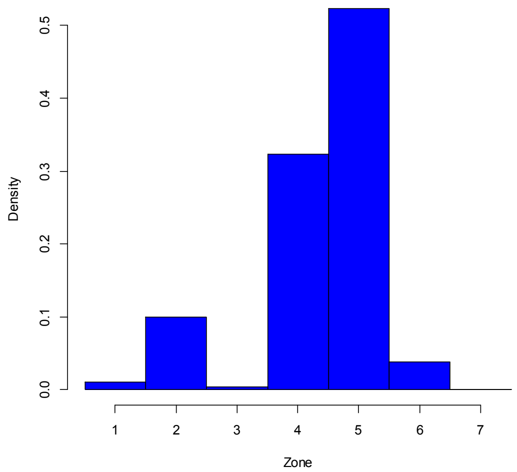


Fig. 7 Histogram of probability density function of Spg.

The Chi-square test for two independent samples (see, e.g., Siegel and Castellan, 1988) was used to verify whether the samples formed by Sp and Spg can be considered to come from the same population. Table 3 summarizes the results obtained.

Table 3. Summary of results obtained in the Chi-square two-sample test (Space variable).

Chi-square two-sample test	
Test statistic	5.31
Critical value (α = 0.05)	12.59
Conclusion	Do not reject H0

Therefore, it can be considered that Sp and Spg have the same distribution.

5.2 Generation of conditioned Space variable

Given an initial value of Spg_t, it is possible to generate the next value Spg_{t+1} according to the respective conditioned distribution functions.

If the procedure is repeated, it generates a sample of pseudo-random values. Figure 8 displays the flowchart for Space generation.

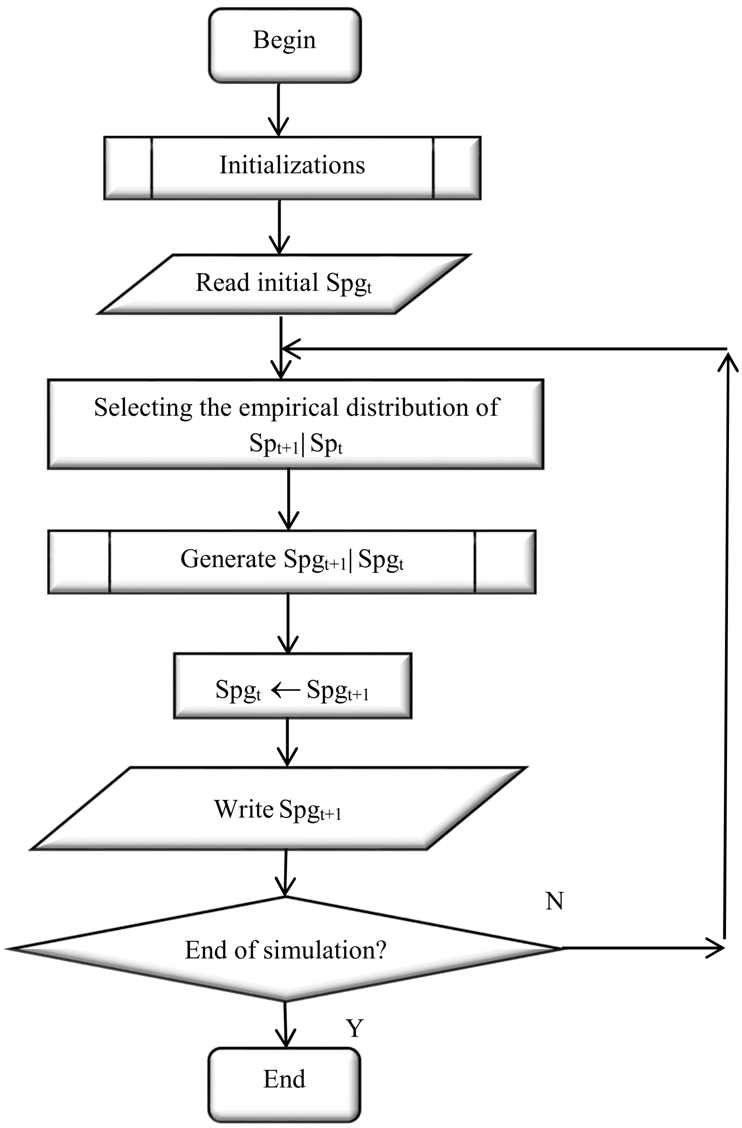


Fig. 8 Flowchart of Space generation.

To check the first-order transitions of Space, a sample of data and the generated one-step transitions were compared.

Based on the generated sample (of size 19000), the number of one-step transitions of the variable Space was also generated, and the results are presented in the transition frequency matrix TG:

$$TG = \begin{bmatrix} 72 & 27 & 3 & 42 & 79 & 9 & 0 \\ 41 & 999 & 3 & 432 & 455 & 54 & 2 \\ 5 & 5 & 2 & 11 & 38 & 3 & 3 \\ 35 & 433 & 10 & 3882 & 1515 & 207 & 6 \\ 73 & 444 & 50 & 1519 & 7466 & 357 & 8 \\ 6 & 73 & 0 & 197 & 354 & 63 & 0 \\ 0 & 4 & 0 & 6 & 9 & 0 & 0 \end{bmatrix} \qquad (6)$$

The generated and data conditioned distribution functions must be compared, which means comparison of 7 conditional distributions. The variables

$$Sp_{t+1} \mid Sp_t = i \text{ and } Spg_{t+1} \mid Spg_t = i, \qquad i = 1, 2, \ldots, 7 \qquad (7)$$

were compared using the Chi-square test for two independent samples (see, e.g., Siegel and Castellan 1988). Table 4 condenses the results.

Table 4. Summary of results obtained in the Chi-square two-sample test (conditioned distribution functions of Space).

Samples compared	Test statistic	C. V. (α = 0.05)	Conclusion
$Sp_{t+1} \mid Sp_t = 1$ $Spg_{t+1} \mid Spg_t = 1$	5.31	9.49	NR
$Sp_{t+1} \mid Sp_t = 2$ $Spg_{t+1} \mid Spg_t = 2$	5.56	9.49	NR
$Sp_{t+1} \mid Sp_t = 3$ $Spg_{t+1} \mid Sp_t = 3$	0.88	5.99	NR
$Sp_{t+1} \mid Sp_t = 4$ $Spg_{t+1} \mid Spg_t = 4$	1.41	12.59	NR
$Sp_{t+1} \mid Sp_t = 5$ $Spg_{t+1} \mid Spg_t = 5$	1.37	12.59	NR
$Sp_{t+1} \mid Sp_t = 6$ $Spg_{t+1} = j \mid Spg_t = 6$	2,71	9.49	NR
$Sp_{t+1} \mid Sp_t = 7$ $Spg_{t+1} \mid Spg_t = 7$	0.20	5.99	NR

C.V. = Critical value; NR= Do not reject $H0_i$; R = Reject $H0_{i;}$

According to the statistical tests, it can be assumed that one-step transitions in the generated sample fit the one-step transitions in the data.

Assuming that the sequence of epicenters is stationary in time (which makes sense, in our opinion), it is possible to properly generate the entire sequence using the one-step transition matrix.

As an example, the matrixes of the second and sixth order are presented below. Note that in the 6^{th} order matrix, the columns are nearly equal. As anticipated, the transition matrices are expected to have equal columns when the order increases, and each line tends to be equal to the Space distribution function.

$$T^{(2)} = T^2 = \begin{bmatrix} 0.07 & 0.149 & 0.005 & 0.282 & 0.500 & 0.044 & 0.002 \\ 0.017 & 0.257 & 0.002 & 0.301 & 0.384 & 0.038 & 0.001 \\ 0.023 & 0.078 & 0.005 & 0.231 & 0.626 & 0.037 & 0.002 \\ 0.009 & 0.095 & 0.003 & 0.467 & 0.387 & 0.037 & 0.002 \\ 0.009 & 0.071 & 0.004 & 0.236 & 0.641 & 0.038 & 0.001 \\ 0.011 & 0.099 & 0.003 & 0.306 & 0.536 & 0.044 & 0.001 \\ 0.010 & 0.107 & 0.004 & 0.320 & 0.520 & 0.037 & 0.002 \end{bmatrix} \qquad (8)$$

This result means that the seismic process of occurrences loses “memory” with time. After several events, the probability of an occurrence among the seismic zones is nearly independent of the location of a far event.

$$T^{(6)} = T^6 = \begin{bmatrix} 0.011 & 0.102 & 0.004 & 0.320 & 0.525 & 0.038 & 0.001 \\ 0.011 & 0.105 & 0.004 & 0.323 & 0.519 & 0.038 & 0.001 \\ 0.011 & 0.098 & 0.004 & 0.314 & 0.535 & 0.037 & 0.001 \\ 0.011 & 0.100 & 0.004 & 0.327 & 0.519 & 0.038 & 0.001 \\ 0.011 & 0.097 & 0.004 & 0.314 & 0.536 & 0.038 & 0.001 \\ 0.011 & 0.099 & 0.004 & 0.318 & 0.529 & 0.038 & 0.001 \\ 0.011 & 0.099 & 0.004 & 0.319 & 0.528 & 0.038 & 0.001 \end{bmatrix} \qquad (9)$$

6 Conclusions

In this work, the sequence of Space locations of the epicenters (seismic zones) in the Azores region was analyzed. The seismic zones were defined in a previous study (Rodrigues and Oliveira 2013), and based on the existing historical and instrumental information, this paper focuses on the analysis of the Space variable using Markovian chains. The one-step transitions of this variable were explored and showed an evident dependence on consecutive earthquake locations.

It can be noted that in zones with greater number of occurrences, (zones 2, 4 and 5) the most probable event is that the next seismic event will take place in the same zone. The next event should be an aftershock. In zones with few earthquakes, the most probable event is that the next seismic event will occur in a different zone. This observation can be explained because zones 1, 3, 6, and 7 are located far from the islands, and if the aftershocks were weak, they might not be recorded.

These results confirm that the spatial sequence of events is not independent of the preceding earthquakes.

These results make sense assuming that the transition probabilities do not change with time, i.e., the Markovian chain is stationary. With this assumption, it is possible to obtain the N-step transitions of Space, which allows estimation of the probability of occurrence of a seismic event in a specific zone based on the “recent seismological past”.

Using simulation techniques, the sequence of epicentral zones was reproduced. The current simulation is a component of ongoing work aimed at modelling the process of seismic occurrences in the Azores region. This project will include a visual computer simulation of the entire process of seismic occurrences in the Azores region.

The Markov chain developed here can be easily exported to the analyses of time events of other natural phenomena such as temperature, precipitation, etc.

References

Bezzeghoud, M., Borges, J.F., Caldeira, B., Buforn, E., Udias, A.: Seismic Activity in the Azores region in the context of the Western part of the Eurasia-Nubia Plate boundary. Proc. International Seminar on Seismic Risk, Azores, Portugal, paper 7 (CDROM), 1-11, 2008.

Borges, J.F., Bezzeghoud, M., Caldeira, B., Buforn, E.: Recent seismic activity in the Azores Region. Proc. International Seminar on Seismic Risk, Azores, Portugal: paper 8 (CDROM), 1-5, 2008.

Burford, D.J., Ger, L., Blake, E.H., Wit, M.J., Doucoré, C.M., Hart, R.: A seismic modelling environment as a research and teaching tool for 3-D subsurface modelling. Journal of Astronomy and Geophysics, 2: 69-77, 2000.

Carvalho, A., Sousa, M.L., Oliveira, C.S., Campos-Costa, A., Nunes, J.C., Forjaz,V.H.: Seismic hazard for Central Group of the Azores Islands. Bull. Geofis. Teor. Appl., 42: 89-105, 2001.

Cavers, M., Vasudevan, K.: Spacio-Temporal Complex Markovian Chain (SCMC) Model Using Directed Graphics: Earthquakes sequencing. Pure. Appl. Geophys., 172: 225-241, 2015.

Dalgaard, P.: Introductory Statistics with R. © Springer-Verlag, 2008.

I.P.M.A.: Instituto Português do Mar e da Atmosfera. http://www.ipma.pt/pt/sismologia/actividade/Accessed in December 2011

Kagan, Y.Y., Bird, P., Jackson, D.D.: Earthquake Patterns in Diverse Tectonic Zones of the Globe. Pure Appl. Geophys., 167: 721-741, 2010.

Kirimidjien, A.S., Anagnos, T.: Stochastic Slip-Predictable Model for Earthquake Occurrences. Bulletin of Seismological Society of America, 74(2): 739-755, 1984.

Madeira, J., Brum da Silveira, A., Hipólito, A., Carmo, R.: Active tectonics in the central and eastern Azores islands along the Eurasia-Nubia boundary: a review. *In* Volcanic Geology of São Miguel Island (Azores Archipelago), edition JL Gaspar et al. The Geological Society Memoir no.44, Chapter 3, 15-32, 2015.

Nanjo, K.Z., Tsuruoka, H., Hirata, N., Jordan, T.H.: Overview of the first earthquake forecast testing experiment in Japan. Earth Planets Space, 63, 159-169, 2011.

Nunes, J.C., Forjaz, V.H., Oliveira,: C.S. Zonas de geração sísmica para o estudo da casualidade do Grupo Central do Arquipélago dos Açores. PPERCAS Project, Report 3/2000, Azores University, Ponta Delgada, Portugal (in Portuguese), 2000.

Nunes, J.C., Forjaz, V.H., Oliveira, C.S.: Catálogo Sísmico da Região dos Açores. Lisbon, Portugal, (CDROM) (in Portuguese), 2004.

Pidd, M.: Computer Simulation in Management Science. John Wiley & Sons, 1987.

Ravindran, A., Phillips, D.T., Dolberg, J.J.: Operations Research principles and practice. John Wiley & Sons, 1987.

Reiter, L.: Earthquake Hazard Analysis: Issues and Insights. Columbia University Press, 1991.

Rodrigues, M.C.M., Oliveira, C.S.: Seismic Zones for Azores based on statistical criteria. Nat. Hazards Earth Syst. Sci, 13:9:2353-2367, doi:10.5194/nhess-13-2337, 2013.

Rubinstein, R.Y., Melamed B.: Modern Simulation and Modeling. John Wiley & Sons, 1998.

Russo, A., Soares, A.: Hybrid Model for Urban Air Pollution Forecasting: A Stochastic Spatio-Temporal Approach. Math. Geosci., 46:75-93, 2014.

Siegel, S., Castellan, N.J.: Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill Co, 1988.

Sopra, L., Patrizi, G.: The application of OR technics for the prediction and understanding of damages caused by seismic events. European Journal of Operational Research, 28:180-195, 1987.

Venables, W.N., Smith, D.M. & R Development Core Team: An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics. Springer-Verlag, 2011.