Modeling spatial dynamics of the Fani Maoré marine volcano earthquake data

Solym Manou-Abi, Said Said Hachim, Sophie Dabo-Niang, Jean-Berky Nguala

Modeling spatial dynamics of the Fani Maoré marine volcano earthquake data

Solym Manou-Abi[1,4*], Said Said Hachim[2], Sophie Dabo[3] and Jean-Berky Nguala[4,5†]

[1*]Institut Montpelliérain Alexander Grothendieck, UMR CNRS 5149, Place Eugène Bataillon, Montpellier, 34090, France.
[2]Conseil Départemental de Mayotte, Mamoudzou, Mayotte, 97600, France.
[3]Laboratory Painleve, UMR 8524, Cité scientifique, Villeneuve d'ascq, 59653, France.
[4]Centre Universitaire de Formation et de Recherche, 8 Rue de l'université, Dembeni, 97660, Mayotte, France.
[5]Laboratoire d'Informatique et de Mathématiques, EA 2525 Parc Technologique Universitaire, 2 rue Joseph Wetzell, La Réunion, 97490, Sainte Clotilde, France.

*Corresponding author(s). E-mail(s): solym-mawaki.manou-abi@umontpellier.fr;
Contributing authors: said.hachim@cg976.fr; sophie.dabo@univ-lille.fr; jean-berky.nguala@univ-mayotte.fr;
†These authors contributed equally to this work.

Abstract

This paper provides the outcomes of a work consisting in modeling and learning some earthquakes data collected during the Mayotte seismo-volcanic crisis of 2018-2021. We highlight the performance of some process data models in order to illustrate the spatial and temporal dynamic. Unsupervised clustering method, spatial pattern analysis, spatial density estimation through spatial marked point process; time series and spatio-temporal models are efficient tools that we studied in this paper to look for the spatial and temporal variation of such spatial data mainly driven by the detected underwater volcano around Mayotte called Fani Maoré. The dynamic of the magnitude and depth

events of the Fani Maoré with the use of the above mentionned models seems to perform the data. We present a discussion thoughout the presentation of the obtained results together with the limit of this study and some forthcoming projects and modeling developments.

Keywords: Classification, Spatio-temporal model, Time series, Spatial point pattern analysis , Earth Science, Spatial density smoothing

1 Introduction

A new emergence of natural seismic events occurred on May 10, 2018, in the Mayotte island, an oversea department of France, located in the Comoros archipelago. Only two days later, others earthquakes events were widely felt across the island and raised local people's fears of a past tsunami in 2004, see [1]. Since then, Mayotte is experiencing a major spread of magmatic events near its area ([2–4]), recording fairly intense earthquakes. During this new seismic wave, a strange low-frequency wave propagated across the planet. At that time, experts argued that earthquakes and this strange signal were probably related to the movement of molten rock; by arguing that the seismic episodes were the result of magma crashing into the subsurface. During these earthquake waves, several deep-sea fish were found dead on the sea surface of the island. A number of hypotheses have emerged in the scientific literature to explain such natural hazards during this crisis period. Among these, a probable volcanic eruption which would have modified the usual biotope of the underwater fauna and flora. In particular, a probable warming of the water temperature as well as a modification of the salinity which disturbed the species. Other hypothesis deal with methane trapped in marine sediments (from organic decomposition) or juvenile carbon monoxide (magmatic origin) accumulated at the base of the water column. The earthquakes must have disturbed the marine depths and released these to the surface. The seismic episode began when few monitoring instruments were present on the island (a few seismometers and GNSS stations) and they did not allow the area to be surrounded. In this context, the uncertainties related to the depth and location of earthquakes were very significant (Lemoine et al., 2019) and the hypotheses as to the origin of this phenomenon, difficult to substantiate. However, with limited monitoring of earthquakes; it took a year after the start of the crisis to begin to instrument the area before a real monitoring network volcanology and seismology of Mayotte (Revosima). The link to a volcanic activity was reinforced by an e-print of a study posted to the EarthArxiv server in February 2019 ([3]). This study attributed the seismic series to a huge magma chamber that was beginning to drain, in what could to be the largest submarine volcanic event ever detected. Thus, researchers finally discover an underwater volcano 50 km off the east coast of Mayotte linked to the ongoing seismic sequence. The basis of the edifice is about 5 km wide with 3.4 km of deep and covers more than a quarter of the area of the main island

(Grande-Terre). The population felt lot of earthquakes, mainly during the first month of the crisis with some buildings damages. Recall that, some relevant observed events during this period were the rise of dead pelagic fish and the indirect effects of subsidence associated with the deformation produced when the magma chamber empties, multiplies the episodes of submersion during high tides. Moreover, it is not known if the volcano is completely new or if it is only recent activity on an old structure. This seismic crisis now related to the volcanic eruption is comparable in volume to that of Lakki, see [5]. Between May and July 2019, nearly $5\,km^3$ of lava was deposited nearly $50\,km$ east of the island on the seafloor due to the emptying of a magma reservoir located at a depth of about $30\,km$, [6, 7]. However, this kind of study could not go further in its conclusions at the time and many questions remain unanswered while scientists analyze the data. It is, in this data analysis context, that we aim in this paper to study the spatio-temporal dynamic of the earthquake data (magnitude, depth) through the recorded data basis. The monitoring of the earthquakes dynamics both temporally and spatially can help to understand the data process evolution and the associate spatial risk. The paper also discusses on the seismic risk perception in Mayotte from the point of view of living populations on the island. Section 2 described the spatio-temporal observation of the geodynamic and seismicity in the study area. In section 3, we present five main analyis and modeling tools used to highlight the characteristics of the spatial and temporal evolution of the earthquake (magnitude, depth). More precisely, we conducted an unsupervised clustering method, spatial pattern analysis, spatial density estimation through spatial marked point process in a given delay; estimation methods through time series and spatio-temporal models. Section 4 consists of our results presentation and a detailed discussion that follows the main methodology applied to the data. Finally, conclusions, limit of the present study and future academic development and ongoing projects are presented in the last section.

2 Study area and observed Data

This section describes the study area of earthquake as well as the geodynamic story, data-sets involved and risk perception. Mayotte and it's surrounding region is used as the study area of the present work. Mayotte is an island of 374 km2 composed of two islands (the largest and main island called Grande-Terre) and located in the Indian Ocean, East of Africa, in the Mozambique Channel, separating Madagascar from Africa.

2.1 Seismovolcanic activity, vulnerability and risk perception

The genesis of the Comorian Archipelago is still discussed within the scientific community. Nevertheless, most authors agree on the existence of intraplate hotspots ([8]) which started to build volcanoes between 15 and 10 million years ago, with an emerging part between 8 and 10 million years ago ([9]). Some

authors propose a combined tectonic activity on transformation faults reactivated by a deformed lithosphere ([10, 11]). Although the region of Mayotte is not completely devoid of volcanic activity, Mayotte had remained peaceful for a long time: the last known eruption took place more than 4.000 years ago. Historically, the geological activity of Mayotte does not record strong and regular earthquakes (see Figure 1). The French seismic regulatory zonation (2010) classifies Mayotte's territory as a zone of moderate seismicity (Zone 3 out of 5). The seismicity story for the period before the global seismometer network was installed, is not well documented. Nevertheless, some events have passed through time without giving any indication of their intensity but only by the prevalence of effects via oral transmission. It is based on a set of testimonies collected from the population and their memories transmitted over several generations, gathered in [1]. This type of memory data collection is common to many African cultures with great certainty.

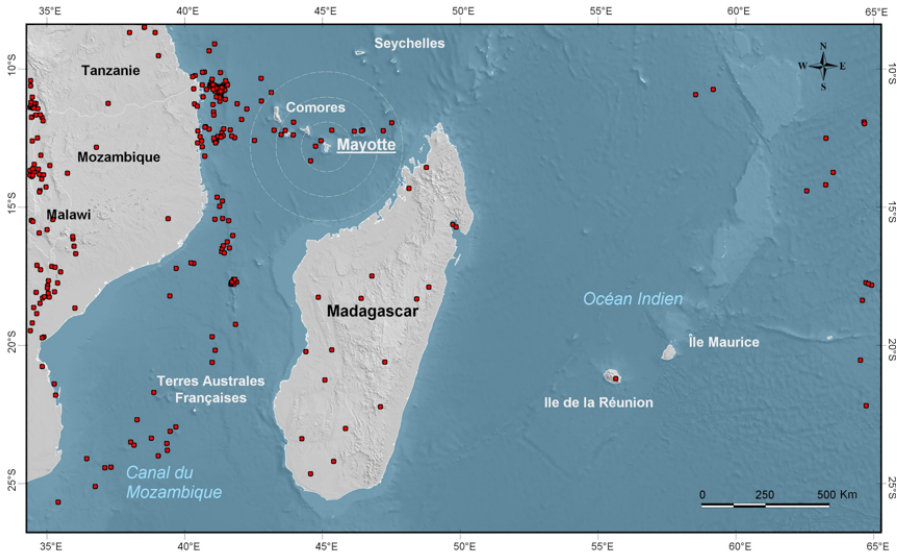

Fig. 1: Seismicity in the Mozambique canal since 1970 up to 2018

The power of an earthquake is measured (at given station) using a magnitude scale called the Richter scale which determines the magnitude level ($M_l = \log\left(\frac{A}{A_0(\delta)}\right)$ where A is the maximum excursion of the Wood–Anderson seismograph, the empirical function A_0 depends only on the epicentral distance δ of the station; see [12]). Thus, the Richter magnitude of an earthquake is determined from the logarithm of the amplitude of waves recorded by seismographs. The typical effects of earthquakes of various magnitudes near the epicenter can be found in [13]. In general, we noticed that the population of Mayotte was historically unprepared for earthquakes; even if historical

records confirm that damaging earthquakes had occurred in the past in this region. The most ancient information goes back to the 17th century with four earthquakes that caused damage to the mosques of Mtsamboro (1606), a city in the north-west of Mayotte as mentioned in [1] and taken up in [14]. This seismic event is linked to the reconstruction of the Tsingoni mosque (he oldest active mosque in France), one of the few if not the only building that has stood the test of time in Mayotte since its construction in the 16th century. Information collected evokes reconstructions following the earthquakes that would be produced later in 1691. In 1829, [15] relates an earthquake that would have occurred 50 years earlier, in December 1829. This event would have led to the collapse of part of the mountain of Koualé. The village of Koualé is now abandoned, no doubt, due to the many natural hazards (ground movements, floods) that have hampered its development. Some seismic events have occurred in this period without a given indication of their intensity. The traces of the 1829 seismic events are mentioned in testimonies in the villages of Sada and M'Tsapéré, see [1]. Note that The catalog of SisFrance (see [16]) earthquakes, mentions also some events that occurred in 1936, 1941, and 1953. A well-known reference earthquake from the population of Mayotte which caused significant damage in the north and south of the main island (but no deaths reported) is the earthquake reported on April 23, 1993, with a magnitude $M_l = 5.0$. In [1], the author found testimonies from the village of M'Tsangamouji telling that this earthquake had destroyed a certain number of housing. The latest earthquake near Mayotte was felt on September 9, 2011, at 4:25 p.m. local time and was located 35 km west of Mamoudzou, the main town. This earthquake reached a magnitude of $M_l = 5.0$. A complete historical seismicity catalog integrating seismicity within the Comoros archipelago, the region of Madagascar, Tanzania, and the Mozambique Channel, can be found in [14] and references therein. On 13 May 2018, a magnitude $M_l = 4.6$ earthquake was widely felt by the island population. This was the first notable event in the earthquake swarm, which is estimated to have started on 10 May 2018 at 8.14am local time. It was the first of the ongoing seismic sequence in Mayotte. The population felt many of them that caused slight damage to buildings. This unusual seismic activity quickly caused panic among the inhabitants, who were not unprepared for such earthquakes events as discussed above. On 15 May, a powerful earthquake was recorded reaching a magnitude of $M_l = 5.62$. One month after the start of the crisis, more than 1.500 seismic events were recorded by measuring instruments, some of them exceed a magnitude of $M_l = 5$, and 140 reached a magnitude of more than $M_l = 4$. This has created an unbearable psychosis among a large part of the island's population. Consequently, prevention policies have to consider preparation for a destructive earthquake on the island as an important component of action plans. Physical damage is visible on many public buildings and some private homes. The authorities prevention policies measure was taken such as closing some of them as a precautionary measure. The school environment was particularly affected, with the closure of several classrooms in secondary

schools. In some schools, timetabling measures were put in place to allow students to continue attending classes despite the constraints. Faced with this situation, evacuation measures were taken, leading to the rehousing of many families whose homes had been destroyed, [17]. Considering the difficult living situation in this seismic crisis period, with especially pressure from the local political actors, the national authorities reacted quickly. From May 2018, a major scientific mobilisation through research resources have been mobilised to better understand this phenomenon. A constellation of several French research organisations were appointed to observe the phenomenon. The first mission called MAYOBS1, took place on board of the Marion Dufresne, a vessel of the French Oceanographic Fleet operated by Ifremer. The objective of this mission was to recover and redeploy six seabed seismometers (OBS) installed three months earlier, as well as to acquire data on the relief and nature of the sea floor. It is during this mission that, the active new marine volcano volcanic edifice was observed, at a depth of $3.5\,km$ and located from $50\,km$ east of Mayotte, [6]; see Figure 2. The volcano culminates at almost 820 m and the diameter of its base is about 5 km. The island itself appears to be shifting 1.5 cm eastward and sinking about 1cm every month since mid-July 2018. The volcanic activity has progressively moved northwest and is now located in the Comoros archipelago. It caused the island to stop growing which, coupled with active tropical erosion due to heavy rainfall, led to its subsidence and the slow construction of a coral reef around Mayotte. The island's inhabitants wonder about the possible dangers that this volcano could represent. Further studies confirm new threats such as the possibility of a tsunami affecting the coast of Mauritius, [18]. An exploratory analysis of the seismovolcanic activity was done in [19] in order to highlight the seismo-volcanic activity. Teleseismic receiver-functions and Rayleigh-wave dispersion curves are used to quantify wave velocity profiles beneath the active volcanic zone off Mayotte. The authors show that the lithosphere in the east-northeast quadrant is composed of four main layers, interpreted as the volcanic edifice, the crust with underplating, the lithospheric mantle, and the asthenosphere. The seismic activity recorded is between the new moho and the mantle. The authors identify a deep magma storage within the lithospheric mantle (between 38 and 54 km depth) whose dynamics have produced seismicity since 2018. This is the main magma reservoir composed of high-melting liquid (uncertainty of 2 km). There is a transfer of magma through a plumbing system from the large reservoir to the surface (new volcano) with intermediate magma chambers, located at 17 km and 28 km, feeding the activity.

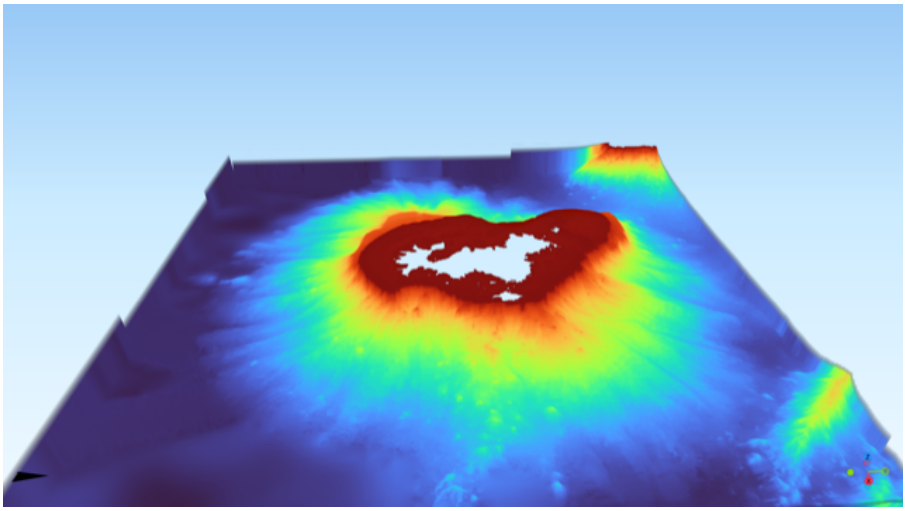

Fig. 2: Simplified reconstruction of the Mayotte volcano (Fani Maoré).

Population information and plan action is a fundamental issue for effective disaster volcano risk reduction. Some actions are undertaken to reinforce the safety of the population against these threats. Awareness-raising campaigns are being conducted to inform the population about what to do in the event of an earthquake or tsunami since january, 2021. Evacuation plans are developed and warning systems are put in place to ensure a rapid response in case of need, [20]. Exposure to high level natural hazards is important to control. People living under the threat of volcanic hazard need information about their vulnerability. In this way let us mentioned a major work on Atlas of natural risks and territorial vulnerabilities of Mayotte in [21], before the beginning of the earthquake swarm. This work deals with the understanding, spatial monitoring of natural hazards in Mayotte. It introduces a range of approaches and maps to explain the perceived and experienced risks, their consequences and their physical, socio-cultural, historical or geographical determinants. It uses multiple sources of geographical, geomorphological, climatic and historical information collected in the field, from institutions and populations. This is for the best of our knowledge the first synthetic monitoring tools that had ever been done in the department of Mayotte with respect to volcanic issues before 2018 for risk prevention. However, the framing of monitoring and risk management, as well as the strategies adopted to share information with the public, has evolved significantly over time, see [22]. Risk prevention and awareness actions in educational/school settings is actually important and being currently the subject of many lessons studies and rechearch projects involving primary and secondary schools of Mayotte, see as an example [23]. Note that, since the discovery of the underwater volcano as well as exploratory analysis of the seismovolcanic activity, other missions have followed until the last called MAYOBS23. Each mission brings new data for the understanding of the

phenomenon. The Marine Advanced geophysical Research equipment Mayotte multidisciplinary Observatory for research and response project (MARMOR) was the logical continuation of this mobilisation in order to better structure the French scientific community working on this subject for continuous seismological monitoring to provide it with continuous data, in real time, 24 hours a day, thanks to an underwater cable infrastructure.

2.2 Observed spatial data set of magnitude and depth

Since the emergence of the new swarm eartquake time series data in the area of Mayotte island, many efforts have been made by the above already mentionned volcanological and seismological monitoring network, Revosima, to monitor the occurrence of seismic events. Research now shows that the epicenter of the time series of seismic episodes were the result of magma crashing into the subsurface and the low-frequency wave, waves resonating in a collapsing magma chamber. The material of this work deal with earthquakes spatial time series data in the area of Mayotte presented spatially in Figure 3, 4 and in Figures 6 and 5 as concerned the time series of the weekly variation of the magnitude and the daily frequency from 2018 to the end of 2021. The time series sequence has an average magnitude equal to $M_l = 2.02$ with a minimum value of $M_l = 0.14$ and a maximum value of $M_l = 5.62$ which appeared on May 15, 2018. The spatial variation by month is shown in Figures 7 to 14 both for magnitude and depth. Note that, the data analysis and modeling aspects are implemented using R software.

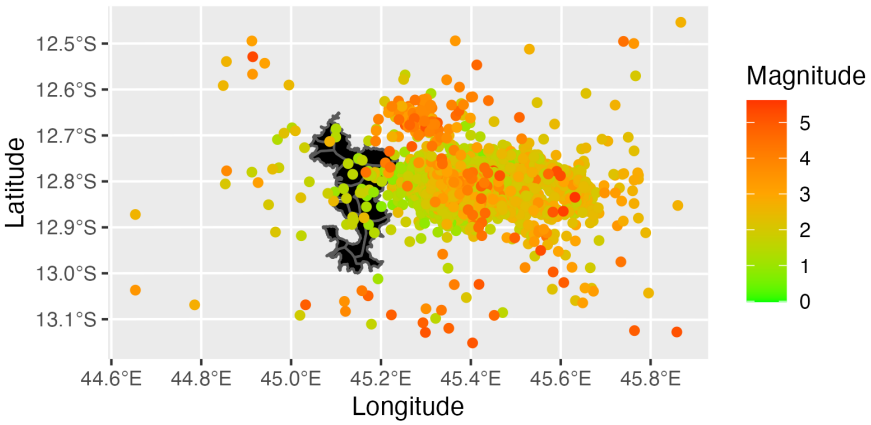

Fig. 3: The spatial magnitude variation from 2018 to 2021.

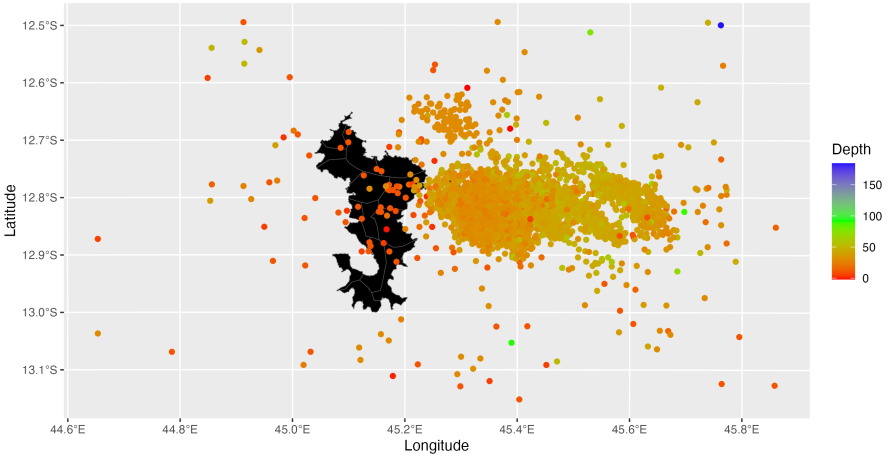

Fig. 4: The spatial depth variation from 2018 to 2021.

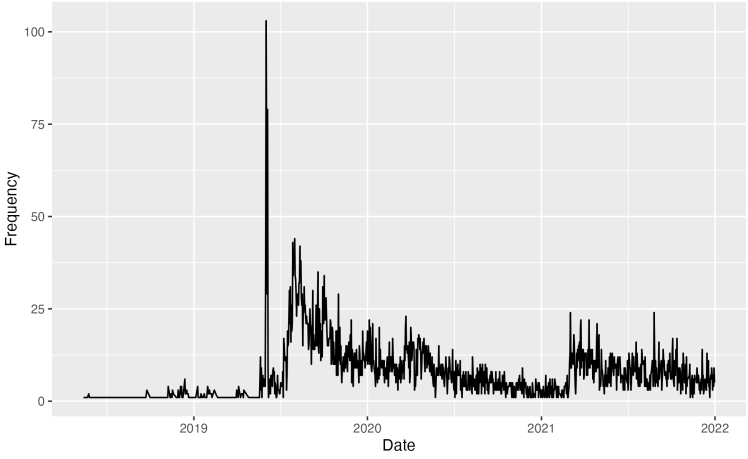

Fig. 5: Daily observation of earthquake

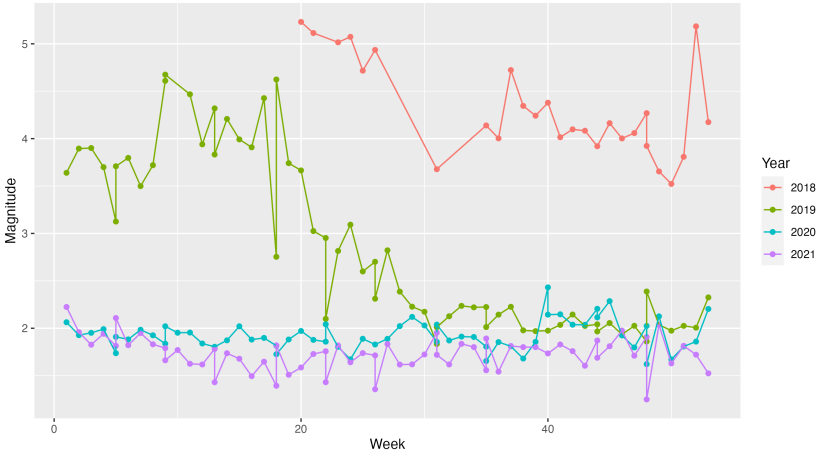

Fig. 6: Weekly observations of the magnitude

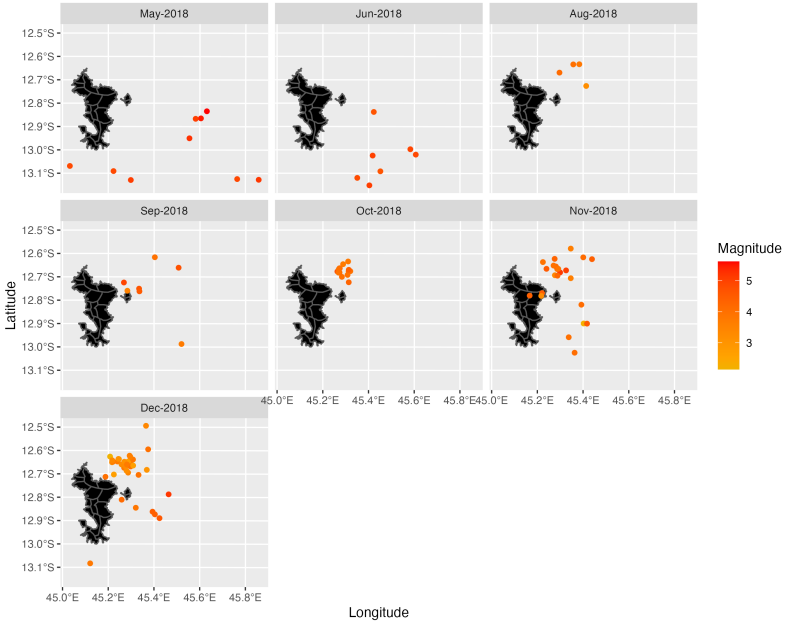

Fig. 7: Spatial monthly variation of the magnitude in 2018

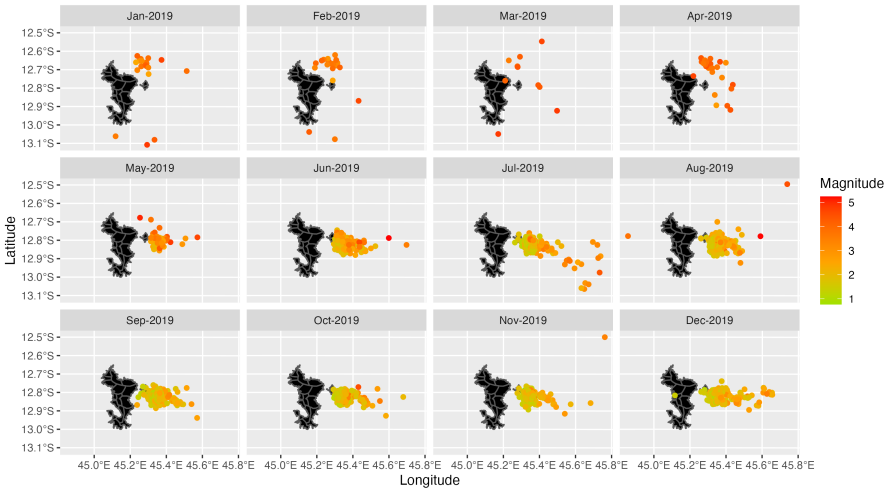

Fig. 8: Spatial monthly variation of the magnitude in 2019

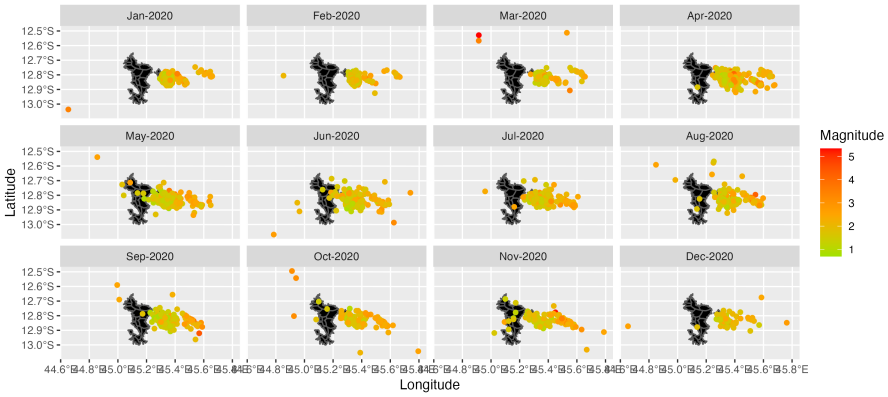

Fig. 9: Spatial monthly variation of the magnitude in 2020

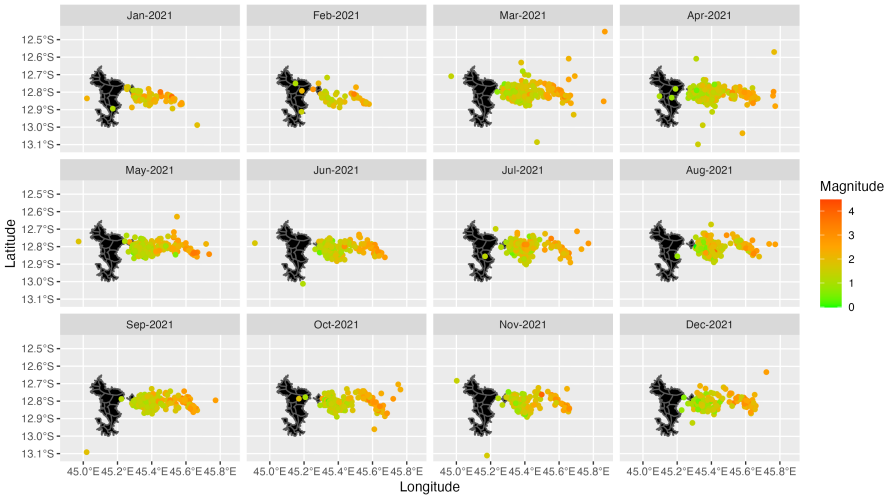

Fig. 10: Spatial monthly variation of the magnitude in 2021

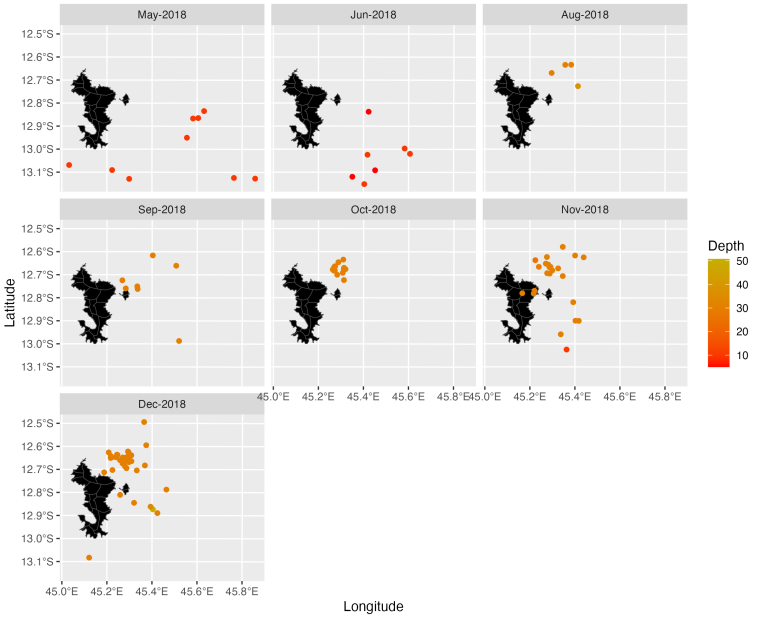

Fig. 11: Spatial monthly variation of the depth level in 2018

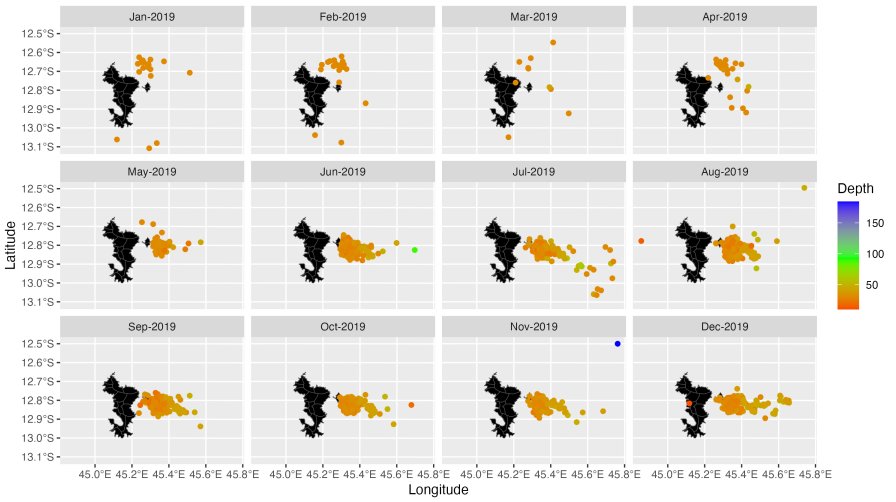

Fig. 12: Spatial monthly variation of the depth level in 2019

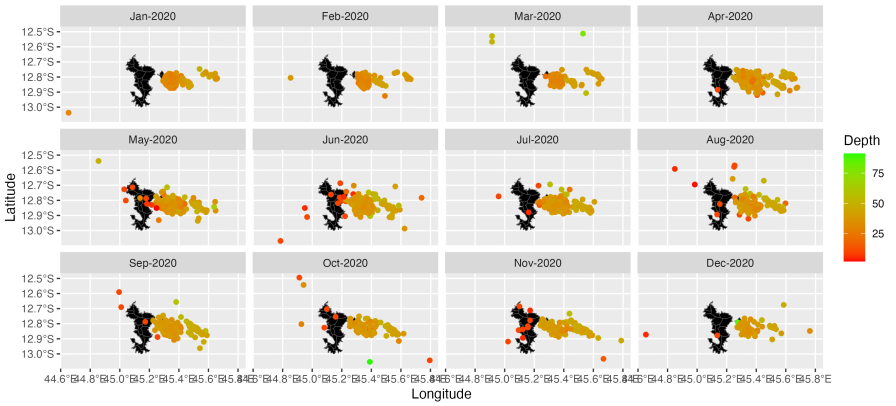

Fig. 13: Spatial monthly variation of the depth level in 2020

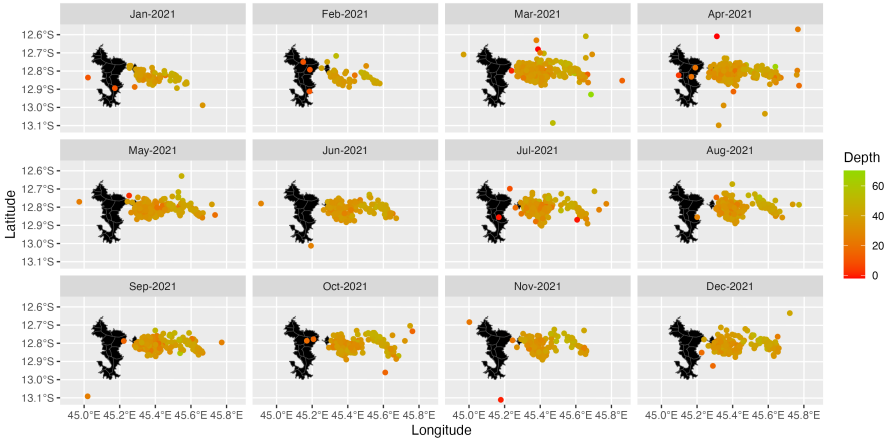

Fig. 14: Spatial monthly variation of the depth level in 2021

3 Modeling method

In this section, we introduce theoretical backgroundand relevant tools for the data analysis and modeling process.

Muldimensional exploration analysis

We conduct principal component analysis (PCA) (see [24]; [25]) on the earthquakes data parameters (longitude, lattitude, depth, magnitude, month/year). This analysis is a pre-processing step before moving on to clustering. The unsupervised classification designates a corpus of methods whose objective is to draw up or find an existing typology characterizing a set of n observations, from p characteristics measured on each of the observations. We use the complementarity between clustering and principal component methods to better highlight the main feature of our data set. The objective is to obtain a hierarchical tree, in other words a sequence of nested partitions going from a partition in which each individual is considered as a group to one in which all the individuals belong to the same group. The first dimensions of PCA will extract most of the information while the last ones are limited to noise and the number of dimensions retained for the grouping can be chosen with several methods. The hcpc function of R software allows us to visualize the tree, the partition and the main components in a convenient way.
An unsupervised clustering is done after based on the principal components

in this paper. The results are plotted in the map to correlate the obtained clustered with the spatial distribution of the earthquakes events.

Spatial point patterns analysis and density smoothing

The spatial distribution of the magnitude can be modelled through spatial point patterns analysis [26] and from marked point process in space-time. The complete spatial randomness (CSR) process corresponds to a distribution of an homogeneous Poisson process (constant intensity function). In general, such a process has the properties of homogeneity (absence of spatial preference) and independence (realizations in one region have no influence on the others). A basic measure proposed by B.D. Ripley (see [27, 28]). is the function $\hat{K}$ defined as the ratio between the average of the number of events expected on a radius r and the intensity λ:

$$K(r) = \frac{\mathbb{E}[N_0(r)]}{\lambda},$$

where $\mathbb{E}[N_0(r)]$. Given an estimation of intensity function at a point x, λ, we can estimate K by a non-parametric method (kernel method) as follows:

$$\hat{K} = \frac{1}{D} \sum_i \sum_{j \neq i} \mathbb{I}_{x_i - x_j \leq r} c(x_i, x_j, r),$$

with $D = \frac{n(n-1)}{\mathsf{W}}$, w stands for the study area and $c(x_i, x_j, r)$ corresponds to the correction of edge effects. The integer n is the total points of observations and $\mathbb{I}_{x_i - x_j \leq r}$ is an indicator function which is equals to 1 if points x_i and x_j are at a distance less than r and 0 otherwise. In the case of complete spatial randomness, the empirical function $\hat{K}$ deviates relatively little from πr^2. Otherwise, if $\hat{K}(r) > K_{pois}(r)$, the points are clustered, else they are in a regular distribution. The Chi-squared test of CSR using the quadrat analysis method (see [29, 30]) allows us to check/confirm the clustered configuration. Whenever it is so easy to see that the homogeneity of the earthquake space is not verified or optimal the use of an in-homogeneous marked point pattern process [31], is necessary in order to give a fitting model. Let

$Y = \{(x_1, m_1), ..., (x_n, m_n)\}, x_i \in W, \quad m_i \in M$, be an in-homogeneous marked point pattern process, where $W \subset \mathbb{R}^2$ is the study region and M a set of possible marks. Denote by $\lambda(x, m)$ on $\mathbb{R}^2 \times M$ the joint intensity for locations $x \in \mathbb{R}^2$ and mark values $m \in M$. The probability density f of the in-homogeneous Poisson marked point process Y with joint intensity function $\lambda(x, m)$ at location $x \in W$ and mark $m \in M$ is given by :

$$f(y) = \exp\Big(\sum_{m \in M} \int_W (1 - \lambda(x_i, m)) dx \Big) \prod_{i=1}^{n_m(y)} \lambda(x_i, m)$$

where $n_m(y)$ is the number of points in y having mark value m. To access spatial trends in the marks, the smoother mark value at location $u \in \mathbb{R}^2$ is

$$\hat{m}(u) = \frac{\sum_i m_i K(u - x_i)}{\sum_i K(u - x_i)}$$

where K can be the Nadaraya-Watson kernel smoother and m_i the mark value at data point x_i.

3.1 Times series : Holt-Winters Model

In addition to the spatial distribution characteristics, the temporal relationship is also reflected in the time series data. In this article, a Holt-Winters model which is a forecasting method based on trends and seasonality, see [32, 33] is used. The Dickey Fuller test is used to check stationarity and the (partial) auto-correlation function (ACF and PACF) help to analyse the differences between the consecutive observations in the absence of stationarity, in order to make a time series stationary considering the delay.

We denote the observed time series by $Y_1, Y_2, \ldots, Y_n$. Denote by $\hat{Y}_{t+h/t}$ the forecast for h periods ahead of Y_{t+h} based on all the data up to time t. Holt-Winters' method is based on three smoothing equations: one for the level, one for trend, and one for seasonality. There are two different Holt-Winters' methods, depending on whether seasonality is modeled in an additive or multiplicative way.

Multiplicative Seasonality

The equations for Holt-Winters' multiplicative method are as follows:

$$\begin{cases} l_t &= \alpha \frac{Y_t}{s_{t-m}} + (1-\alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1-\beta) b_{t-1} \\ s_t &= \gamma \frac{Y_t}{l_{t-1} + b_{t-1}} + (1-\gamma) s_{t-m} \\ \hat{Y}_{t+h/t} &= (l_t + b_t h) s_{t-m} + h_m^+, \end{cases}$$

where l_t represents the level of the series, b_t denotes the growth (trend), s_t is the seasonal component, $h_m^+ = [(h-1) mod\, m] + 1$, and m is the length of seasonality. The parameters $\alpha, \beta,$ and γ are usually restricted to lie between 0 and 1. We refer the reader to [34] for more details.

Additive Seasonality

The basic equations for Holt-Winters' additive method are as follows:

$$\begin{cases} l_t &= \alpha(Y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\ s_t &= \gamma(Y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \\ \hat{Y}_{t+h/t} &= (l_t + b_t h)s_{t-m} + h_m^+. \end{cases}$$

3.2 A Spatio-Temporal data process model

In this section, we give a modified version of a deterministic Dynamic Spatio-Temporal process model, see [35]. It is the so called Inverse Distance Weighting and as the name suggests, the IDW model predicts the attribute value of a variable at positions where no samples are available based on the spatial distance between that position and other positions where samples have been collected. Observations that are spatially closest to the observation to be predicted receive a greater weight in the prediction, while distant observations will have a relatively weak influence on the prediction. The weight assigned to each sample is also controlled by a power parameter, often referred to as θ. As θ increases, the weight assigned to outlying samples decreases. This factor can be understood as a smoothing parameter. If the smoothing effect is too great, it can hide interesting variability or give the impression that the variable to be predicted is much more homogeneous on the plot than it really is. In fact, when θ is small, nearby and far samples exert a strong influence on the observations to be predicted. Consequently, the prediction risks being very smooth because it does not take into account only local phenomena. On the other hand, if θ is too large, only very local processes will be considered which may give the interpolated map a grainy, non-smooth appearance. An advantage the IDW technique is that it takes into account more observations to make the predictions, which is likely to improve the accuracy of the latter. Nevertheless, it is necessary to set the p-parameter correctly for the interpolated map to be relevant and reliable.

The novelty comes from the fact that we introduce a covariate (here the depth of the observed magnitudes) in the prediction. Note that the earthquake events we consider for the application are localised from from an underwater volcano. We consider magnitudes observations $\{M(s_{ij};\ t_j)\}$ for spatial locations (longitude and latitude) $\{s_{ij} : i = 1, \ldots, m_j\}$ at times times $\{t_j : j = 1, \ldots, T\}$. The following (exact) spatio-temporal predictor is built from [35] as follows on some location s_0 and time t_0:

$$\hat{M}(s_0;\ t_0) = \sum_{j=1}^{T} \sum_{i=1}^{m_j} w_{ij}(s_0;\ t_0) M(s_{ij};\ t_j)$$

where

$$w_{ij}(s_0;\ t_0) = \frac{\tilde{w}_{ij}(s_0;\ t_0)}{\sum_{k=1}^{T}\sum_{l=1}^{m_k}\tilde{w}_{lk}(s_0;\ t_0)}$$

and

$$\tilde{w}_{ij}(s_0;\ t_0) = K((s_{ij};\ t_j),(s_0;\ t_0)) \tag{1}$$

with $K((s_{ij};\ t_j),(s_0;\ t_0))$ a transition kernel function depending on the distance between the spatio-temporal location $(s_{ij};\ t_j)$ and the prediction location $(s_0;\ t_0))$ with the bandwidth parameter θ that specify the redistribution of weights for the process at the previous time over the spatial domain taking into account the contribution of the depth explanatory variable in the weights. We shall consider the case

$$K((s_{ij},\beta_{ij};\ t_j),(s_0;\ t_0)) = \frac{1}{d((s_{ij};\ t_j),(s_0;\ t_0))^{\theta}}, \tag{2}$$

At this step, it suffices to compute pairwise distance to obtain the kernel weight. But at this stage, we have practical implications, because not only, data may leads to very small values approaching zero of $d((s_{ij};\ t_j),(s_0;\ t_0))^{\theta}$ which causes problems in the above equation. Thus if values of 0 are present in the data, we derive the following transformation

$$Y^*(s_{ij};\ t_j) = \frac{Y^*(s_{ij};\ t_j)(N-1)+1/k}{N}$$

where N is the number of observations. This compresses the data symmetrically around .5 from a range of 1 to $(N-1)/N$, so extreme values are affected more than values lying close to 1/2. Additionally, we see that for N large enough the compression vanishes, that is, larger data sets are less affected by this transformation.
Note that Another classical example, among others, of a kernel function K is the Gaussian radial basis kernel

$$K((s_{ij};\ t_j),(s_0;\ t_0)) = \exp\Big(-\frac{1}{\theta}d((s_{ij};\ t_j),(s_0;\ t_0))^2\Big) \tag{3}$$

where the bandwidth parameter θ is proportional to the variance parameter in a Gaussian distribution. Usually, θ is chosen by cross-validation.

3.3 Performance criteria and Model selection

To evaluate the prediction models, we'll make use of the following statistical measure.

Mean Square Prediction Error (MSPE):

$$MSPE_N = \frac{1}{N}\sum_{i=1}^{N}\big(y_i-\hat{y}_i\big)^2,$$

where y_k denotes actual value and $\hat{y}_k$ denotes predicted value and N is the total number of testing samples. Since the use of the bandwidth parameters θ is necessary, we make use of a cross-validation principle for a better selection of the bandwidth parameter θ. We choose several values of the parameter θ and seeks to evaluate model predictions by splitting the data into a training sample and a validation sample, then fitting the model with the training sample and evaluating it with the validation sample. In K-fold cross-validation we randomly split the available data into K roughly equal-size components (or "folds"). Each fold is held out, the model is trained on the remaining $K-1$ folds, and then the model is evaluated on the fold that was held out. Specifically, for $k = 1, ..., K$ folds, we fit the model with the k-th fold removed, and obtain predictions $\hat{y}_i^{-k}$ of y_i for $i = 1, ..., m_k$ where m_k is the number of data in the k-th fold. We then select a MPSE metric by which we evaluate the predictions relative to the held-out samples. Note that, good choices for the number of folds are $K = 5$ and $K = 10$. We refer the reader to [36]; for more details on cross-validation methods.

4 Results

The purpose of this section is to apply the above methodology to the 2018-2021 earthquakes data in Mayotte area together with a detailled discussion.

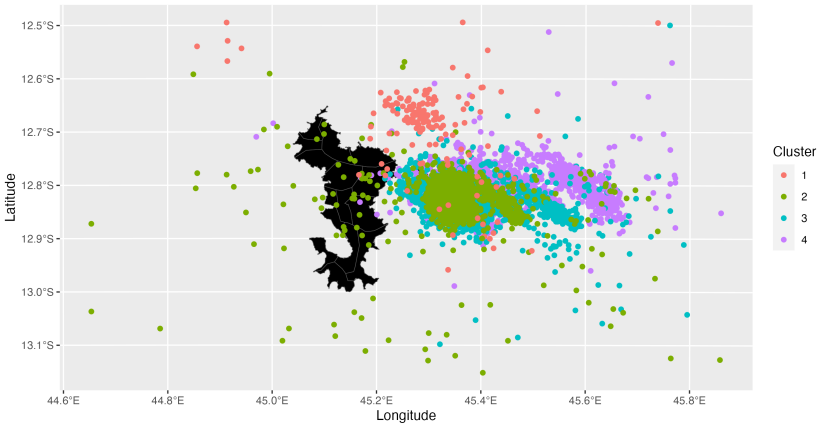

Fig. 15: Cluster position on the map: Cluster 1 (red color), Cluster 2 (green color), Cluster 3 (blue color) and Cluster 4 (pink color)

We try to explain the relative differences on the seismic sequence, based on the longitude, latitude, depth, magnitude, month and year, by using an Ascending Hierarchical Classification. Earthquakes data processed with a factor analysis shows the Clusters positions on the map in Figure 15. The cluster segregation shows that the earthquake data evolve in a random manner and are increasingly clustered towards the east of Mayotte. Spatio-temporal variation hence drives each cluster. Cluster 1 (158 events) is characterized mostly by magnitude level greater than $M_L = 4$ for 75% of them ; while the depth is between 10 and 53 with mean equal to 30. Such events happen between October 2018 and April 2019. In this cluster, events (minority proximal swarm) are located mainly in the north-east of Mayotte. This events corresponds to the period when the population felt the most seismic events. If in Cluster 2 (4560 events) events were recorded between 2018 and 2021, their are located mostly in the east of Mayotte with depths varying from 0 to 46 (it is the main proximal swarm). Note that 75% of the magnitudes are greater than 2.4 with a mean magnitude equal to $M_L = 2.08$. The cluster 3 (1882 events) segregates only events located at the east and increasingly to the east of Mayotte. They are recorded between december 2019 and 2021 with low and high depth values. The magnitude mean is around 1.94 and 75% of the magnitudes are greater than 2.27. Finally, Cluster 4 (2565 events) contains events that occurred between January 2020 and December 2021. It seems to be a distant swarm located mostly more than 25km from the east of Mayotte. The mean magnitude is equal to $M_L = 2.09$ and 75% of the magnitudes are greater than 2. The depth varying from 0 to 77 with a mean equal to 36.

Now let's look at a spatial point pattern analysis which allow to abalyse the spatial structures through the distributions of the spatial points with a random distribution. By dividing the observation area into 5000 cells, the Chi-squared Test of the CSR using quadrat counts shows that the earthquake events distribution is clustered and not completely random. The estimation of the ratio function $\hat{K}$ between the average of the number of events expected on the radius r, the intensity λ and the CSR complete spatial randomness function $\hat{K}_{pois}$ is given in Figure 16. We find that the function $K(r)$ is almost above the complete spatial randomness function which induces again global clustered spatial points of earthquake events apart from May 2018 to May 2019 (where the distribution were completely random). Globally, we can say that the daily events of future earthquakes events depend on those of the past over the four years and are almost clustered.

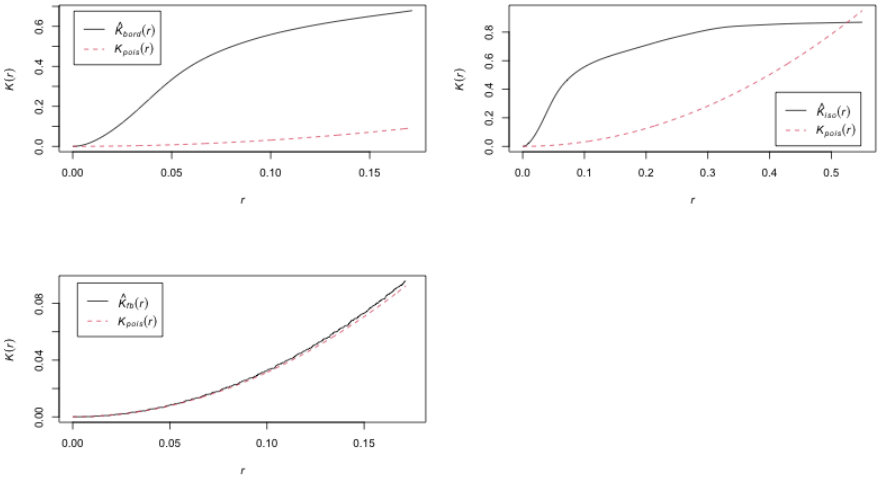

Fig. 16: The complete spatial randomness results

Since we observed an in-homogeneous and clustered spatial variation, we show here a spatial smoothing estimation effect obtained using a marked point pattern process. In Figure 17, we plot a spatial Kernel estimate of the observed magnitude density by year to get an impression of the local spatial variations. Globally, this shows that a spatial estimation of the magnitude density leads to a mixture of clustered and randomly spatial distributed earthquakes events. The spatial density smoothing results show a slightly spatial growing trend from the north and south toward the eastern part of Mayotte for slightly high and moderate magnitude level until April 2019; see Figure 21.

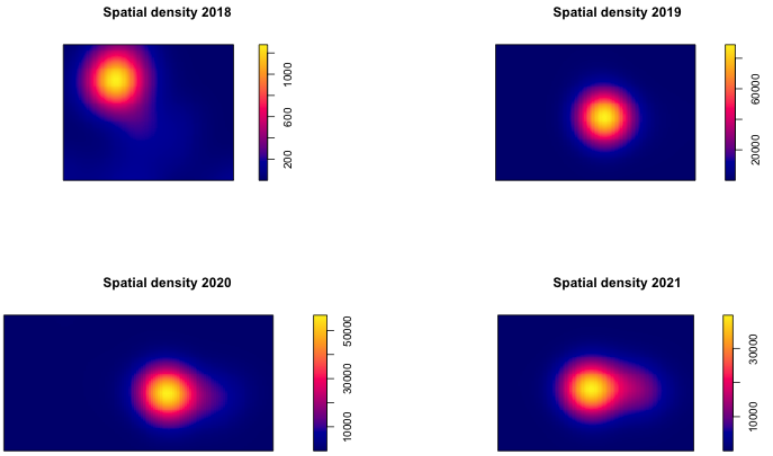

Fig. 17: Fitting the yearly spatial variation of the magnitude density

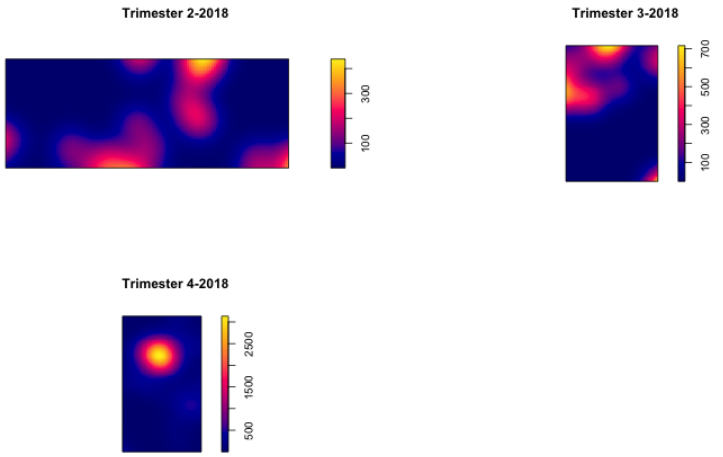

Fig. 18: Fitting the trimester spatial variation of the magnitude level events density over the year 2018.

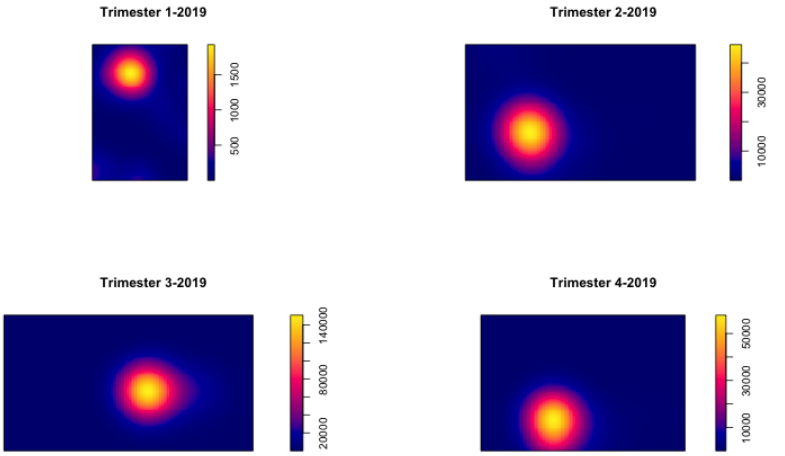

Fig. 19: Fitting the trimester spatial variation of the magnitude level events density over the year 2019.

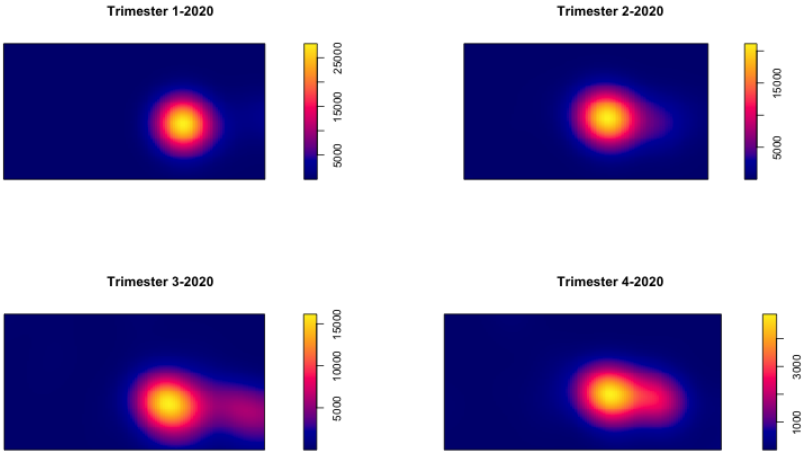

Fig. 20: Fitting the trimester spatial variation of the magnitude level events density over the year 2020.

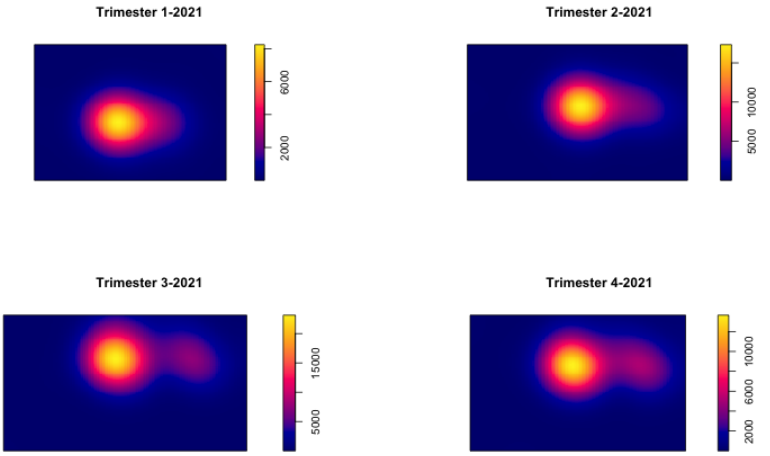

Fig. 21: Fitting the trimester spatial variation of the magnitude level events density over the year 2021.

In addition to the above spatial distribution characteristics, the temporal relationship is also reflected in the time series data. We show that the time series of the daily frequency and weekly magnitudes of earthquakes presents trends, cycles, and seasonality. The Dickey-Fuller Test show that with a p-value of 53.5% allow to conclude the non-stationarity of the time series. By applying the Holt-Winters models, we obtain approximately the same result in the additive and multiplicative methods for the forecasting results in Figure ??.

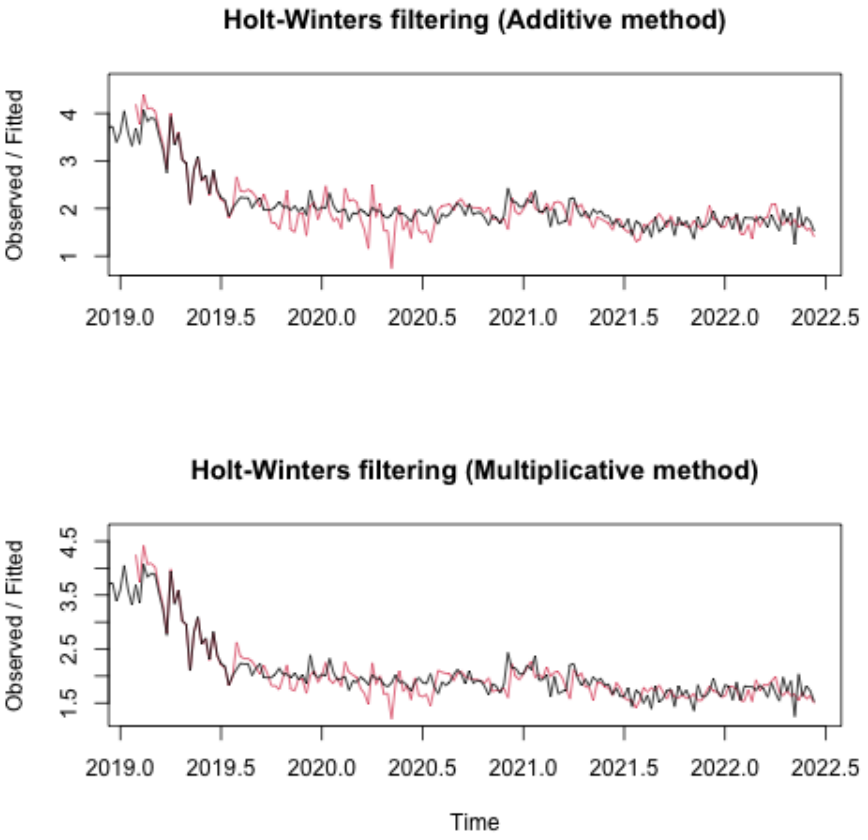

Fig. 22: Estimate frequency of daily earthquakes by Holt-Winters model

We observe a growing trend until September 2019 and a decreasing trend starting from this latter month. For seasonality, we observe seasonality during July over the four years. As a result, during this month, we observe an increase in the frequency of earthquakes. Globally, the frequency decreases significantly over the years for all types of levels. Between 2019 and 2020, there is a decrease of 13% for low magnitudes (magnitude level less than 3) earthquakes and 63% for medium magnitudes (magnitude level between 3 and 5), and from 2020 to 2021, there is an increase of 33% in low magnitude, compared to 74% for medium magnitudes. On the other hand, concerning the slightly high magnitudes (magnitude level between 5 and 6), note that from 2019 until the end of 2021, they are rare events.

Now, we present the results on the the predictions of the spatio-temporal data model introduced in Section 3.2. The spatio-temporal estimation that we

present is concerned with the magnitude (one could also liked to estimate the depth) and is given in Figures 23 and 24.

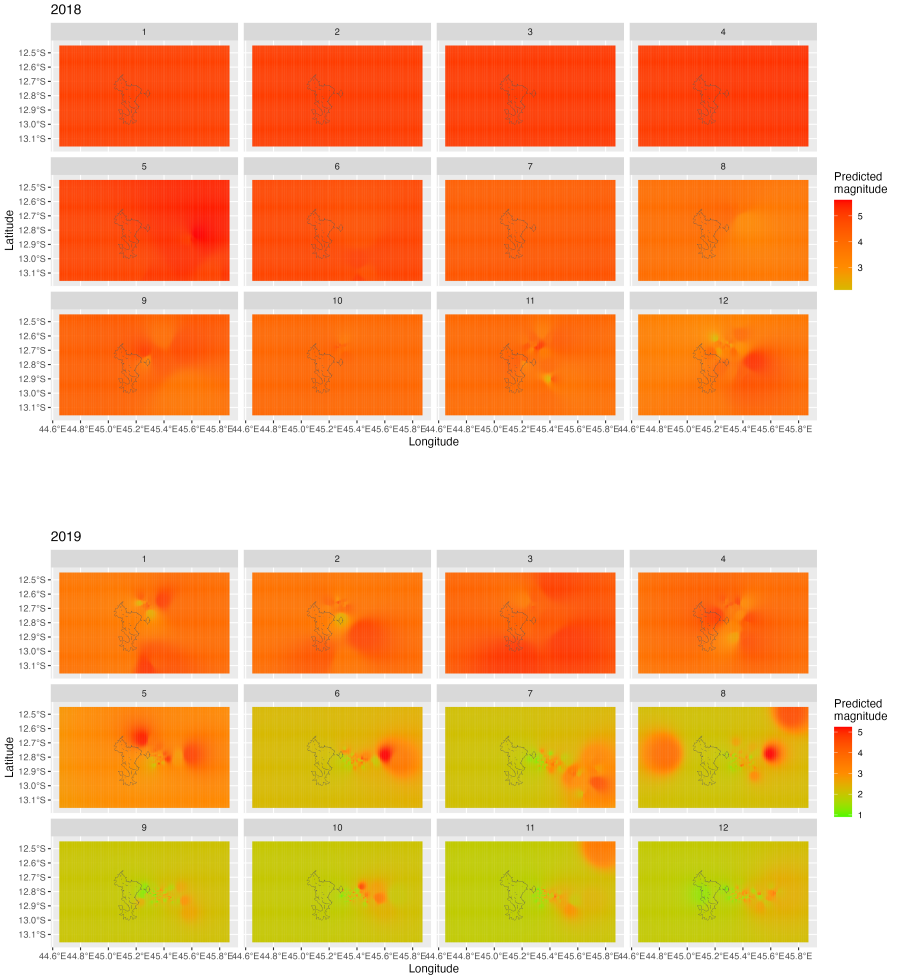

Fig. 23: Fitting the Spatio-Temporal Data model

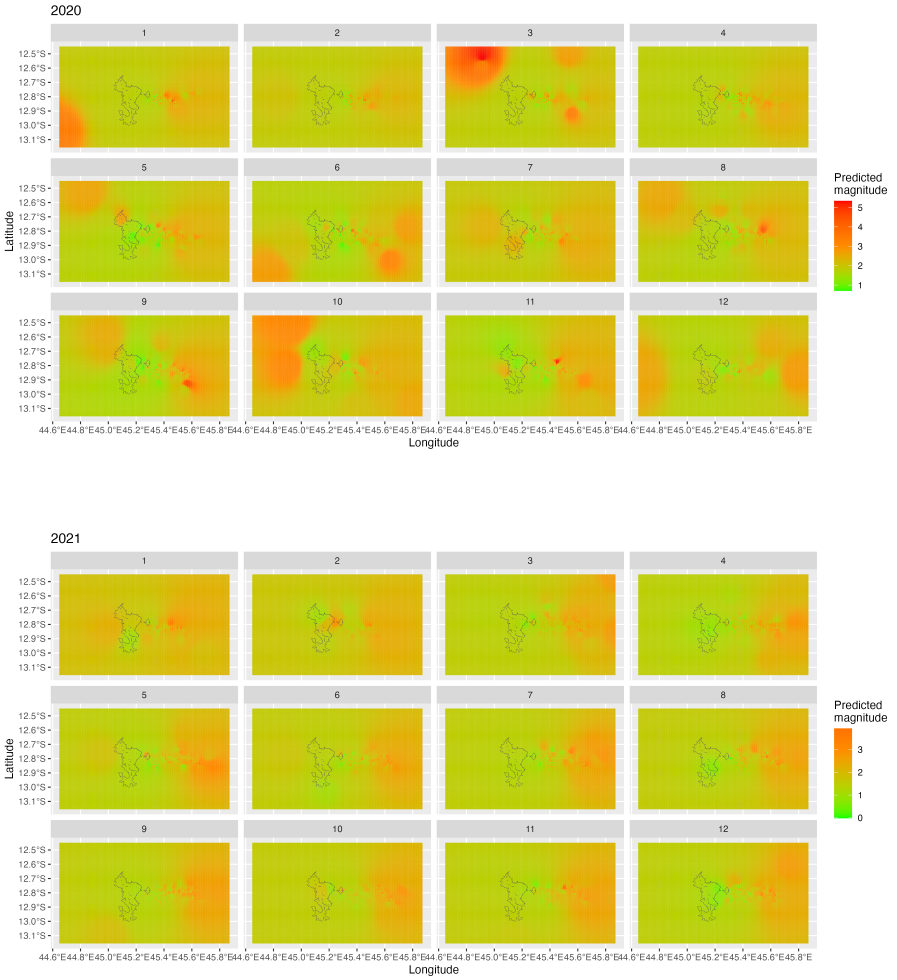

Fig. 24: Fitting the Spatio-Temporal Data model

Until May 2019, we observe a spatial evolution of earthquakes with magnitude remaining almost constant with moderate level and some week probability of appearance of slightly high magnitude around 5. We see later that from May 2019, this lead to a strong contagion in all spatial locations, especially for low magnitudes and a shift toward the eastern part of Mayotte. There appears to be an epicentre in the northwest first (2018). Thereafter, there was a massive concentration in the east of Mayotte, mainly with low magnitude levels (in the vicinity of the detected marine volcano). We also note a continuous propagation of earthquakes of average intensities approximately the same as in 2019

with in addition, a meteoric concentration of weak earthquakes. They spread mostly towards the east.

5 Conclusions and perspectives

In this study, several spatial analysis and data models were used to explore the dynamics of the observed spatial data. The use of unsupervised clustering method reveals that the data are increasingly clustered with a spatio-temporal variation towards the east of Mayotte. The spatial structure was check by a point pattern analysis method together with a spatial smoothing trough some time variation. In addition, the temporal distribution were reflected in the time series data with clearly a trends and seasonality and cycle. Finaly, a spatio-temporal process was introduce to capture both the space and time variation of the earthquake data driven by the Fani Maoré underwater volcano. Addressing other developed framework of this paper in future research may yield to non linear dynamic spatio-temporal process data model including more data information than the magnitude and depth. Other practical perspectives deal with the development of more reliable data collection tools. For example, the MAR-MOR project was able to monitor data that are becoming very big data hence relevant for a deep learning. In this context, some major projects are ongoing. Firstly, the implementation of an analytical platform to develop a data distribution based on web applications to analyse and identify some warning events. This will be usefull for the monitoring of the marine volcano evolution. In addition, given the wide variety of data sets from different missions (geophysical, marine biology, environmental, etc.), there is a need for computing and data storage infrastructure. The complexity of the modelling used for the analysis, processing, calibration and correlation of these data requires dedicated computing resources. Thus the establishment of a spatial data mining center in Mayotte, will allow the management and data analysis with statistical and spatial tasks. Indeed, it is important to minimise false alarms when analysing data to ensure reliable results.

6 Declaration

Supplementary information

Additional supporting information of the original contributions presented in the study are included in the article/supplementary material.

Availability of data and materials

The raw data supporting the conclusions of this paper is available in the volcanological and seismological monitoring network Revosima(see http://www.ipgp.fr/fr/revosima). The R-codes is available from the corresponding author on reasonable request.

Acknowledgments

We thank all of the people involved in this work. Part of this work received funding from the Conseil Departmental de Mayotte.

Author contribution

SMA and SSH performed the conception, experiment, interpretation, statistical analysis and the writing of the present paper. SD performed the conception, interpretation and the writing of the present paper. JBN performed the interpretation, writing and some section conception. All authors contributed to the article and approved the submitted version.

Declarations

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript. We certify that the submission is original work and is not under review at any other publication.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Funding

Not applicable.

References

[1] Hachim, S.: Catastrophes. Mayotte perd sa mémoire, Mémoire de DEA de Géographie de l'Université P. Valéry de Montpellier (2004)

[2] Paquet, F., Jorry, S., Deplus, C., Le Friant, A., Bernard Sr, J., Bremell-Fleury Sr, S., Feuillet, N., Gaillot, A., Guérin Sr, C., Thinon, I.: The

mayotte seismo-volcanic crisis: Characterizing a reactivated volcanic ridge from the upper slope to the abyssal plain using multibeam batymetry and backscatter data. In: AGU Fall Meeting Abstracts, vol. 2019, pp. 43–0219 (2019)

[3] Lemoine, A., Briole, P., Bertil, D., Roullé, A., Foumelis, M., Thinon, I., Raucoules, D., de Michele, M., Valty, P., Colomer, R.H.: The 2018–2019 seismo-volcanic crisis east of mayotte, comoros islands: seismicity and ground deformation markers of an exceptional submarine eruption. Geophysical Journal International (2020)

[4] Saurel, J.-M., Jacques, E., Aiken, C., Lemoine, A., Retailleau, L., Lavayssière, A., Foix, O., Dofal, A., Laurent, A., Mercury, N., *et al.*: Mayotte seismic crisis: building knowledge in near real-time by combining land and ocean-bottom seismometers, first results. Geophysical Journal International **228**(2), 1281–1293 (2022)

[5] Thordarson, T., Self, S.: The laki (skaftár fires) and grímsvötn eruptions in 1783–1785. Bulletin of Volcanology **55**, 233–263 (1993)

[6] Feuillet, N., Jorry, S., Crawford, W.C., Deplus, C., Thinon, I., Jacques, E., Saurel, J.M., Lemoine, A., Paquet, F., Satriano, C., *et al.*: Birth of a large volcanic edifice offshore mayotte via lithosphere-scale dyke intrusion. Nature Geoscience **14**(10), 787–795 (2021)

[7] Cesca, S., Letort, J., Razafindrakoto, H.N., Heimann, S., Rivalta, E., Isken, M.P., Nikkhoo, M., Passarelli, L., Petersen, G.M., Cotton, F., *et al.*: Drainage of a deep magma reservoir near mayotte inferred from seismicity and deformation. Nature geoscience **13**(1), 87–93 (2020)

[8] GRAVIOU, P.: Valorisation du patrimoine géologique de mayotte: un éventail d'actions complémentaires

[9] Debeuf, D.: Étude de l'évolution volcano-structurale et magmatique de mayotte, archipel des comores, océan indien: approches structurale, pétrographique, géochimique et géochronologique. PhD thesis, Université de la Réunion (2009)

[10] Nougier, J., Cantagrel, J., Karche, J.: The comores archipelago in the western indian ocean: volcanology, geochronology and geodynamic setting. Journal of African Earth Sciences (1983) **5**(2), 135–145 (1986)

[11] Michon, L.: The volcanism of the comoros archipelago integrated at a regional scale. Active Volcanoes of the Southwest Indian Ocean: Piton de la Fournaise and Karthala, 333–344 (2016)

[12] Richter, C.F.: An instrumental earthquake magnitude scale. Bulletin of

the seismological society of America **25**(1), 1–32 (1935)

[13] Hutton, K., Woessner, J., Hauksson, E.: Earthquake monitoring in southern california for seventy-seven years (1932–2008). Bulletin of the Seismological Society of America **100**(2), 423–446 (2010)

[14] Bertil, D., Mercury, N., Doubre, C., Lemoine, A., Van der Woerd, J.: The unexpected mayotte 2018–2020 seismic sequence: a reappraisal of the regional seismicity of the comoros. Comptes Rendus. Géoscience **353**(S1), 1–25 (2021)

[15] Gevrey, A.: Essai sur les Comores (1870). Pondichéry, fac-similé par l'Association malgache d'Art et d'Archéologie (1972)

[16] Jomard, H., Scotti, O., Auclair, S., Dominique, P., Manchuel, K., Sicilia, D.: The sisfrance database of historical seismicity. state of the art and perspectives. Comptes Rendus. Géoscience **353**(S1), 1–24 (2021)

[17] Hermann, D.: Danger au canelia palm: 65 logements évacués à cause des fissures. Les Nouvelles de Mayotte **8 juin 2018 Numéro 3041.** (2018)

[18] Lemoine, A., Pedreros, R., Filippini, A.: Scénarios d'impact des tsunamis pour mayotte. Technical report, Technical Report BRGM/RP-69869-FR. 169 p., 21 ill., 8 Tab., 61 ann (2020)

[19] Dofal, A., Michon, L., Fontaine, F.R., Rindraharisaona, E., Barruol, G., Tkalčić, H.: Imaging the lithospheric structure and plumbing system below the mayotte volcanic zone. Comptes Rendus. Géoscience **354**(S2), 47–64 (2022)

[20] Janvier, M.: La préfecture a présenté sa nouvelle communication sur les risques séisme et tsunami. FRANCE Mayotte matin **11 janvier 2021 Numéro 2563.** (2021)

[21] Leone, F., Péroche, M., Lagahé, E., Gherardi, M., Sahal, A., Vinet, F., Hachim, S., Lavigne, F.: Territorial accessibility modeling for tsunami crisis management in french-administered mayotte (indian ocean). In: Annales de Géographie, vol. 693, pp. 502–524 (2013). Armand Colin

[22] Devès, M., Lacassin, R., Pécout, H., Robert, G.: Risk communication during seismo-volcanic crises: the example of mayotte, france. Natural Hazards and Earth System Sciences **22**(6), 2001–2029 (2022)

[23] Nguala, J.B.: Club maths: une expérience pour re-donner confiance et développer un esprit critique des élèves des classes de lycée professionnel. Les Cahiers d'Education & Devenir (2021)

[24] Jolliffe, I.: Principal component analysis. Encyclopedia of statistics in behavioral science (2005)

[25] Praene, J.P., Malet-Damour, B., Radanielina, M.H., Fontaine, L., Riviere, G.: Gis-based approach to identify climatic zoning: A hierarchical clustering on principal component analysis. Building and Environment **164**, 106330 (2019)

[26] Gómez-Rubio, V.: Spatial point patterns: methodology and applications with r. Journal of Statistical Software **75**, 1–6 (2016)

[27] Ripley, B.D.: The second-order analysis of stationary point processes. Journal of applied probability **13**(2), 255–266 (1976)

[28] Ripley, B.D.: Modelling spatial patterns. Journal of the Royal Statistical Society: Series B (Methodological) **39**(2), 172–192 (1977)

[29] Gesler, W.: The uses of spatial analysis in medical geography: a review. Social science & medicine **23**(10), 963–973 (1986)

[30] Gatrell, A.C., Bailey, T.C., Diggle, P.J., Rowlingson, B.S.: Spatial point pattern analysis and its application in geographical epidemiology. Transactions of the Institute of British geographers, 256–274 (1996)

[31] Baddeley, A.J., Møller, J., Waagepetersen, R.: Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. Statistica Neerlandica **54**(3), 329–350 (2000)

[32] Hyndman, R.J., Khandakar, Y.: Automatic time series forecasting: the forecast package for r. Journal of statistical software **27**, 1–22 (2008)

[33] Billah, B., King, M.L., Snyder, R.D., Koehler, A.B.: Exponential smoothing model selection for forecasting. International journal of forecasting **22**(2), 239–247 (2006)

[34] Hyndman, R., Koehler, A.B., Ord, J.K., Snyder, R.D.: Forecasting with Exponential Smoothing: the State Space Approach. Springer, ??? (2008)

[35] Wikle, C.K., Zammit-Mangion, A., Cressie, N.: Spatio-temporal Statistics with R. CRC Press, ??? (2019)

[36] Ruppert, D.: The elements of statistical learning: data mining, inference, and prediction. Taylor & Francis (2004)