Cluster recognition in spatial-temporal sequences: the case of forest fires

**Carmen Vega Orozco · Marj Tonini · Marco Conedera ·
Mikhail Kanveski**

Abstract Forest fire sequences can be modelled as a stochastic point process where events are characterized by their spatial locations and occurrence in time. Cluster analysis permits the detection of the space/time pattern distribution of forest fires. These analyses are useful to assist fire-managers in identifying risk areas, implementing preventive measures and conducting strategies for an efficient distribution of the firefighting resources. This paper aims to identify hot spots in forest fire sequences by means of the *space-time scan statistics permutation model* (STSSP) and a *geographical information system* (GIS) for data and results visualization. The scan statistical methodology uses a scanning window, which moves across space and time, detecting local excesses of events in specific areas over a certain period of time. Finally, the statistical significance of each cluster is evaluated through Monte Carlo hypothesis testing. The case study is the forest fires registered by the Forest Service in Canton Ticino (Switzerland) from 1969 to 2008. This dataset consists of geo-referenced single events including the location of the ignition points and additional information. The data were aggregated into three sub-periods (considering important preventive legal dispositions) and two main ignition-causes (lightning and anthropogenic causes). Results revealed that forest fire events in Ticino are mainly clustered in the southern region where most of the population is settled. Our analysis uncovered local hot spots arising from extemporaneous arson activities. Results regarding the naturally-caused fires (lightning fires) disclosed two clusters detected in the northern mountainous area.

Keywords Forest fires · Cluster analysis · Point patterns · Space-time scan statistics · Permutation model

C. Vega Orozco (✉) · M. Tonini · M. Kanveski
Institute of Geomatics and Risk Analysis, University of Lausanne, Bâtiment Amphipôle, 1015 Lausanne, Switzerland
e-mail: CarmenDelia.VegaOrozco@unil.ch

M. Conedera
WSL Swiss Federal Research Institute, Insubric Ecosystems Research Group, via Belsoggiorno 22, 6500 Bellinzona, Switzerland

1 Introduction

Forests cover about 30 % of the total planet land area and have a major role in the life on earth. For instance, they are essential for atmospheric regulation of CO_2 and O_2, soil conservation, regulation of the hydrological cycle, while simultaneously providing habitat for many species and useful products (biomass, wood, food) [3]. Forest distribution and composition are mostly affected by *fires*. Such events can profit forest ecosystems by regulating species composition and influencing plant growth and reproduction; however, fires can severely threaten the underlying environment when they are uncontrolled or when the forest ecosystems are non-fire-adapted. Forest fire phenomena exhibit complex and irregular spatio-temporal distributions which make it difficult for modelling and predicting their behaviour [32].

Switzerland is generally not subjected to big or frequent forest fires; nevertheless, approximately 200 fires take place every year in its territory, destroying up to 300 ha of forests which are valuable to the country's socio-economic sectors and for the protection against other natural hazards. The majority of the forest fires in this country are located in the Alpine region, comprising the Cantons of Valais, Uri, Grisons and Ticino [44].

Forest fires frequency, intensity and distribution are controlled by the coincidence of different factors such as the presence of biomass (fuel), fire-conductive environmental conditions (e.g. meteorology) and ignition energy (lightning or humans) [16]. When fire events are recurrent and consistent in a particular area over a certain period of time, they usually result in a specific fire regime [17]. Today's fires are mainly the result of persistence conflicts in landscape exploitation as a consequence of anthropogenic activities.

A basic step towards achieving fire management goals by minimizing costs is to conduct a comprehensive assessment of the predominant fire regime, including the identification of existing fire clusters and the detection of their spatio-temporal extent and related risks [9, 13, 30]. Furthermore, this information is of paramount importance for fire management measures implementation in order to discriminate between structural fires and temporary hot spots. Structural fires belong to areas where forest fires are particularly recurrent due to permanent and concomitant presence of flammable fuel, ignition sources and fire-conductive weather conditions. Conversely, temporary hot spots are due to particular and extemporaneous local conditions such as repeated arsonist actions. In fact, when mitigating fire related problems, structural fire areas may need different preventive infrastructural and organizational measures than the strategies required in the extenuation of the temporary arson fires.

In this context, the principal objective of the present study is to detect spatio-temporal clusters of forest fires, as well as to analyse the ignition causes of the resulting clusters. This knowledge can further help fire managers and local authorities to define and to implement fire policy and control measures.

Generally speaking, spatial cluster analysis can be defined as a family of algorithms aiming at grouping spatially distributed objects showing a local overdensity. These algorithms can be grouped into global and local methods. The first group seeks to statistically define whether the events under study show a random distribution or a clustered pattern, e.g. Moran's I [24], *K*-functions [31], fractal dimension [23] and Allan factor [1]. The second group, local methods, intends to identify and locate clusters in space. Some examples are the geographical analysis machine (GAM) [27], the Turnbull's cluster evaluation permutation procedure (CEPP) [39], Density Based Spatial Clustering (DBSCAN) [8] and the Scan Statistic [20]. Several studies are related to forest fire cluster detection but they mainly carry out only spatial or only temporal clustering pattern analyses [22, 34–36]. Only a few studies perform a comprehensive spatio-temporal cluster analysis [11, 12, 40]; they have estimated

the degree of spatial clustering of fires for one-year frame periods, assuming, for each year, an independent distribution pattern of the ignition-points.

This paper widely differs from the aforementioned cluster analysis methods. This study proposes a methodology which detects clusters of forest fires in space and tests their statistical significance, as proposed by Tonini et al. [37] and Tuia et al. [38]. Building on these works, this study goes beyond the cluster identification and analyses the ignition-causes of the detected spatio-temporal clusters in order to identify the local trends of the fire cluster origins (structural fires and temporary hot spots).

The following section (Section 2) describes the study area, the data set collected by the Forest Services of Canton Ticino, and the scan statistics methodology applied in the development of this study. Section 3 portrays the results obtained after simulations. Sections 4 and 5 present a discussion and conclusions, respectively, drawn from the analyses.

2 Material and methods

2.1 Study area

Canton Ticino is located in Switzerland in the southern slope of the Swiss Alps (Fig. 1); it has a total area of 2,812 km² and a population of 320,000 inhabitants [10]. Topography in Canton Ticino is characterized by a marked altitudinal gradient (from 200 to 3,400 m above sea level—m.a.s.l.) with a rather heterogeneous geology, dominated by siliceous rocks originating in connection with the tectonics of the Alps. Depending on the elevation and the geographical location, the mean annual precipitation ranges from 1,600 to 2,600 mm and the mean annual temperature from 3 to 12 °C. The climate in this region is demarcated by dry and mild winters comprising few days (40 days per year on average) distinguished by strong gusts of a katabatic (descending) dry wind from the north (foehn), which consequently drops the relative humidity to values as low as 20 %. The high amount of summer rain (800 to 1,200 mm in the period June-September) is characteristic for the whole area. Throughout this season long periods without rain, or even drought, may alternate with thunderstorms and short spells of heavy precipitation [33].

At the beginning of the last century, Ticino was mainly a rural canton. This situation drastically changed in the post-war period when the canton experienced a large socio-economic transformation towards a more service-oriented economy that ensured prosperity and benefited a strong population growth. Nevertheless, this situation also caused the almost total abandonment of traditional agriculture, livestock breeding and landuse management activities. According to different available sources [2], the forest area grew to more than double its size, passing from *ca.* 62,900 ha in 1900 to around 129,000 ha in the year 2000, with a remarkable acceleration of the forest extension process since the post-war period [6]. Nowadays the forest cover is dominated at low elevations (up to 900-1,100 m.a.s.l.) by anthropogenic monocultures of chestnut tree (*castanea sativa*), occasionally interrupted by the presence of other broadleaved species such as *Tilia cordata, Quercus petraea, Q. pubescens, Alnus glutinosa, Prunus avium, Acer spp. or Fraxinus spp*. At medium elevations (900-1,400 m.a.s.l.) the forests mostly consist of pure stands of *Fagus sylvatica* followed by coniferous forests (*Picea abies*), and at higher elevations *Larix decidua*. On the south-facing slopes the beech belt is sometimes completely missing. The presence of *Abies alba* has been reduced to small patches on the north-facing slopes in the central part of the area. Pine forests are confined to very particular sites: *Pinus sylvestris* on dry south-facing slopes and *Pinus cembra* on the most continental areas of the upper regions [4].

Fig. 1 Forest-fires distribution in Canton Ticino. Dots represent forest fire ignition-points for the period 1969-2008

2.2 Forest fire geo-database

Basic information of the forest fires in Canton Ticino has been collected by the Forest Service since 1900. Starting from 1969, the data has been organized in a relational geo-database [28] that stores geo-referenced information of single fire events: XY-coordinates of the ignition points, date of first alarm, date of fire extinction, ignition cause, burned area, slope, altitude, etc.

For the present study, we analyzed all forest fires recorded between January 1st 1969 and August 31th 2008 comprising a total of 2,401 fire events. From this geodatabase, several datasets were extracted to perform different simulations regarding fire-origins (anthropogenic

and naturally caused fires) with different spanning periods. Only the two more significant studies are reported in this paper: one analysis using all the events enabling an overall analysis of all fires, and the second analysis considering only the events due to lightning.

Because the method proposed in this paper does not distinguish clusters due to an increase risk of fire or due to a different geographical event distribution at different times [19] (details in Section 2.3), the first dataset, which includes all fires, was split into three groups (datasets I, II and III) in order to obtain more homogeneous fire regime conditions as possible for a sound statistical analysis. Dataset I contained 833 fires occurring between 1969 and 1978, dataset II held 762 events happening between 1979 and 1990 and dataset III comprised 806 fires burning between 1991 and 2008. The definition of these three datasets was based on different factors that had conditioned the distribution and frequency of fires in Canton Ticino in the last 40 years. As cited in Section 2.1, this canton experienced an increase in forest area during the post-war period and a rise in the risk of fire. Under this situation, the cantonal authorities gradually put into operation measures lessening the ignition occurrences of the anthropogenic wildfires and actions for early fire-fighting [6]. The most efficient fire-preventative dispositions considered in the analyses of this paper were: the major fire brigades reorganization implemented in 1978, the systematic use of helicopters for both transport of the fire fighters and aerial firefighting since 1980 [6]; and the implementation of two preventive legal acts (1989 and 1991) aiming at prohibiting burning activities in the open spaces.

The lightning fire dataset was separately analyzed over the entire study period (1969-2008) comprising a total of 175 events, given that their fire regime is quite different from the anthropogenic-caused fires, and that their ignition occurrences are not affected by the measures mentioned above.

As shown in Fig. 2a, the distribution of the annual number of fires during the whole study period is rather irregular, with a general drift towards lower fire frequencies (< 60 events per year) after 1990. The annual burnt areas in Fig. 2b are generally of small sizes (< 500 ha per year) with a tendency to lower values after 1976; nevertheless, fires can get out of control when fire-prone conditions are present through the year. Such situations were observed in 1970, 1973, 1981, 1990 and 1997 where fires resulted in extremely large burnt areas. It should be noted that 1973 is an exceptional year with 176 fire events that burned 7,273.88 ha, despite an average of 1,500 ha for the years mentioned above. This trend toward a diminution of the frequency and burnt area size of the fires is mostly the effect of the implementation of several firefighting dispositions achieved by the canton.

Looking at the fire regime along the year, see Fig. 3, all three considered sub-periods uncovered a major peak in March-April; a time-frame mostly characterized by surface fires (rapid spreading) of anthropogenic origins at low elevations (< 1,000 m.a.s.l.). In the so called winter period, from December to April, lightning fires are not present. In the summer season (May to November, corresponding to the vegetation period), fires are of both natural and anthropogenic origins, with a prominent peak of lightning-fires in July-August, especially in the period 1991-2008.

Regarding fires of known causes, in the period 1969-1978, human activities were the origin of about 97 % of the fire events, while only 3 % corresponded to lightning. In the period 1979-1990, this proportion shifted to 94 % and 6 % respectively; whereas, in the third period 1991-2008, human behaviour was the cause for the ignition of about 87 % of the total fire events while lightning was responsible for the remaining 13 %. Concerning human-ignited fires, negligence and arson are the origins of the majority of forest fires as shown in Fig. 4. The implementation of preventive measures mitigating the risk of fire-ignition can explain the small presence of other anthropogenic-caused fires such as railways, army
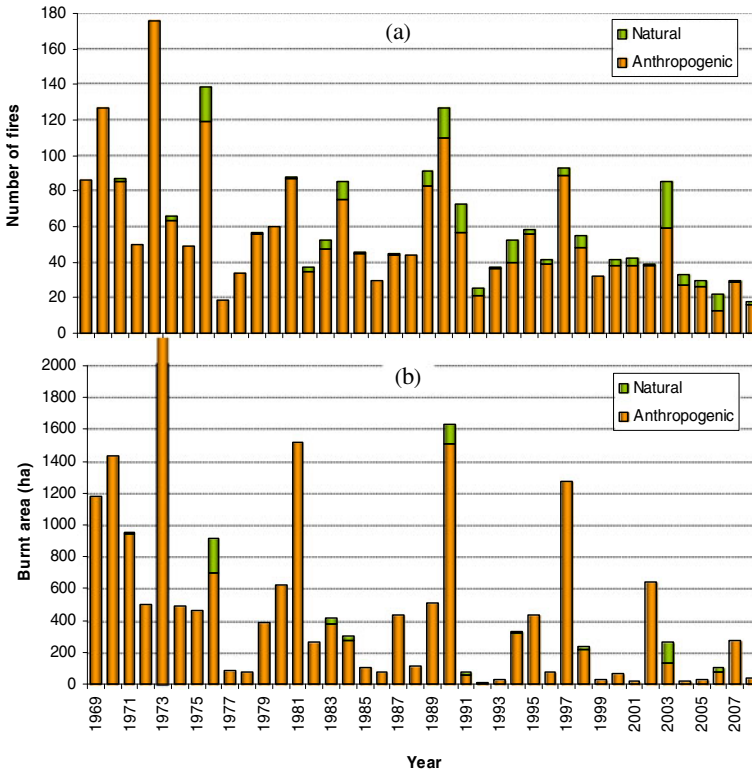
Fig. 2 Annual number of fires **a** and burnt area **b** in canton Ticino (1969-2008)

activities and electrical lines [5]. Among these minor ignition causes, fires detonated by electric lines and other sources are experiencing a blooming trend.

2.3 Space-time scan statistics

Scan statistics represents a collection of methods, used and adapted in many domains, to search for local excesses of events (clusters) in both space and/or time. The main purpose is to determine whether or not an observed cluster, assumed to belong to a randomly distributed pattern, is statistically significant or has rather occurred by chance. The method was first developed in health science by Naus [25, 26], and more recently, once more in the health domain, Kulldorff developed spatial [18] and spatio-temporal extensions [20, 21]. Nowadays, these methods are implemented in a large variety of fields.

For the purely spatial scan statistics, the region under study is scanned by a circular window centered on each event. Each window moves across the entire area, varying its radius continuously from zero up to a fixed upper limit. Each circle takes the nearest neighbour events location falling inside and compares them with those lying outside. Under the null hypothesis of spatial randomness, these events are inferred to be distributed according to a known discrete-state random process (Poisson distribution) which parameters can be estimated. Given this assumption, it is then possible to test whether or not these events are randomly and independently distributed in a specific area.
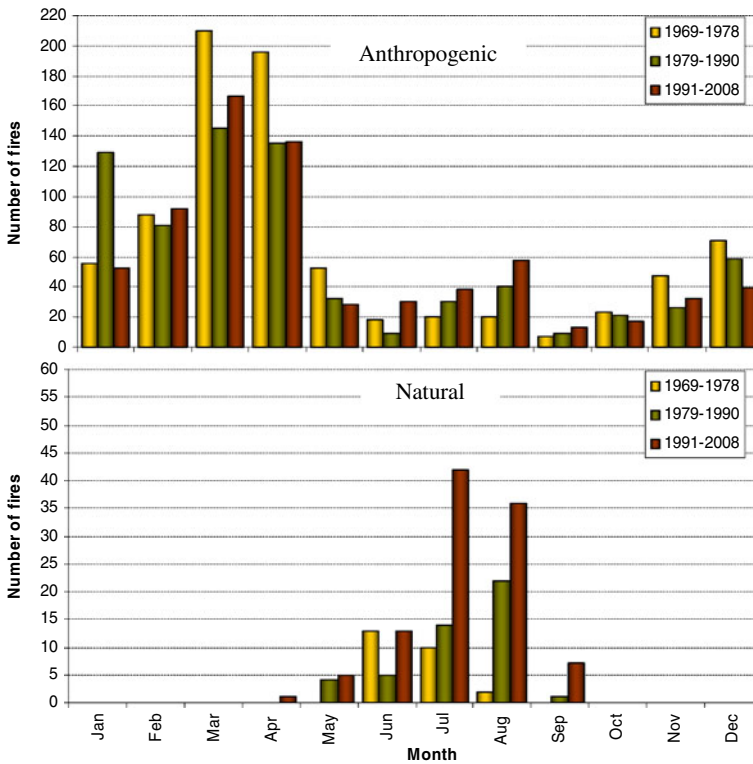
Fig. 3 Monthly distribution of forest fires in the 3 sub-periods for natural (lightning) and anthropogenic ignition-causes

The likelihood function, representing the probability that a specific zone contains a cluster, is computed for every possible scanning window and the one maximising the function represents the most likely cluster, and so on. The statistical significance of the retained potential clusters is then evaluated in order to test whether or not they have occurred by chance. For this purpose, Monte Carlo hypothesis testing was performed with a large number of random replications of the dataset generated under the null hypothesis [19]. The rank of the maximum likelihood from the real dataset is compared with the rank of the maximum likelihood from the random datasets. For instance, if the likelihood ratio for the most likely cluster exceed 95 % of the values in the Monte Carlo simulations, then, the cluster is considered to be significant at the 5 % level (p-value =0.05) [21]. In this way, it is possible to reject a cluster when the corresponding p-value is above the fixed threshold value. Based on the desired threshold, the number of Monte Carlo simulations is established.

The computational method of the space-time scan statistics is an adaptation of the purely spatial scan statistic in space and time. The circular window is replaced by a cylinder with the circular base representing the geographic space and the height corresponding to the time period of the potential clusters. Cylinder's sizes can increase from zero up to a maximum value in both space (radius) and time (height). As in purely spatial scan statistics, each cylinder visits each event geographical location and, additionally, it visits each possible time period.

The two standard models (Poisson for discrete data and Bernoulli for binary data) demand the definition of a control population in order to compute the expected number of cases
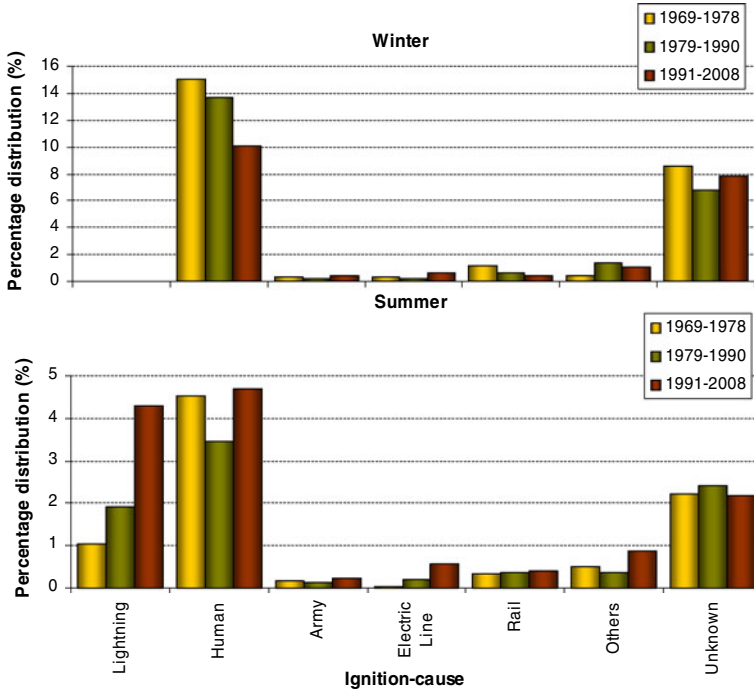
Fig. 4 Ignition sources according to fire season (winter and summer) (100 %=entire database)

inside each scanning window. However, when the control population data is not available or not known, the problem is overcome using the *space-time scan statistic permutation model* (STSSP). This model only requires case data which corresponds to the single observations of the events under study. Thus, the expected number of cases is estimated on the base of the observed cases under the assumption of no space-time interaction, meaning that the spatial and temporal locations of all events are independent of each other. A complete explanation of the permutation model can be found in Kulldorff et al. [21]; here, the statistics is briefly exposed.

Let C be the total number of observed cases and c_{zd} the number of cases observed within a zone z in a day d. The expected number of cases μ_A for a space-time cylinder A can be estimated as the sum of μ_{zd} (the number of expected cases per day and per zone) belonging to cylinder A:

$$\mu_A = \sum_{z,d \in A} \mu_{zd} \qquad where \qquad \mu_{zd} = \frac{1}{C}\left(\sum_z c_{zd}\right)\left(\sum_d c_{zd}\right)$$

Let c_A be the number of observed cases in a cylinder A. Inferring that this variable follows hypergeometric distribution and that C is large compared to $\sum_{z \in A} c_{zd}$ and $\sum_{d \in A} c_{zd}$, then c_A can be considered to be Poisson-distributed with mean μ_A [37]. Thus, a Poisson Generalized Likelihood Ratio (GLR) can be computed as follows:

$$GLR = \left(\frac{c_A}{\mu_A}\right)^{c_A}\left(\frac{C - c_A}{C - \mu_A}\right)^{(C - c_A)}$$

This ratio is calculated and maximized for every possible cylinder and Monte Carlo simulations are performed to test the statistical significance of the detected clusters.

STSSP model is a statistical tool very useful for the analysis of the distribution of environmental data. The main advantage of this model is that it only uses the observed cases instead of requesting both the case data and the population-at-risk. In the case of forest fires, the identification of the control population referring these events is a thorny task. Biomass could be considered as the material risking to be burnt; yet, this element is quite complicated to quantify and to localize at high resolution level and over large areas, becoming the major limitation for the implementation of other scan statistical models.

In the present study calculations were performed with SaTScan™ software developed by Martin Kulldorff [19]. The program allows the user to indicate all the requirements to perform the analysis, such as input data, coordinates system, study period, number of Monte Carlo replications, etc. Concerning the input data and the study period, cluster analyses were conducted for four datasets as described in Section 2.2 and for each one a SaTScan simulation was performed: dataset I with fires that occurred from 1969-1978, dataset II with fires detected in the period 1979-1990, and dataset III with fires that burned during 1991-2008. The fourth dataset comprises only lightning fires and spans the whole study period (1969-2008) for a total of 175 events.

All fire events were specified as individual locations (X,Y-Cartesian coordinates referenced with the Swiss National grid) with each point representing one fire occurrence (case) and the related date of fire-ignition in the format of year-month-day. The results are non-overlapping clusters identified using the retrospective space-time permutation model.

For the definition of the scanning space-time window parameters in SaTScan, multiple simulations were executed using different maximum-size values; for instance, for the upper limit on the geographical size, circles with varying radius from 500 m to 5 km were used, and for the temporal cluster size, maximum lengths from 1-year to 50 % of the study period and time intervals from 1-week to 1-month length were studied. These hyperparameters were also compared with different comprehensive spatial and temporal structural analyses that were completed (out of the scope of this paper) in order to detect the degree of clustering of the forest fires in Canton Ticino. The temporal distribution was analyzed using Allan Factor statistics suited to detect scaling behaviour in point processes [35]. The spatial point pattern of the forest fires was analyzed using a wide variety of spatial structural analysis tools such as topological, statistical and fractal measures [14, 15, 41] and the Ripley's K-function [42].

Only the two most representative results, consistent with qualitative analyses from the forest fire experts in Canton Ticino, are presented in this paper. The scanning space-time window parameters used in these analyses were set as follows: 1) for the datasets I, II and III, the maximum spatial window size was set to a 3-km radius and the temporal window was set to a maximum size of 25 % of the time-length of each dataset, enabling the detection of clusters spanning several years, and a time interval of 1 month in order to detect clusters with monthly temporal trends. 2) For the lightning induced fires dataset the maximum spatial window size was set to be a 3-km radius, and the temporal window was set to a maximum length of one year and with a time interval of 15-days length in order to detect clusters within one season since lightning fires only take place each year in summer season.

The statistical significance threshold for cluster detection was fixed at 5 % level of confidence (i.e. p-value≤0.05) with the smallest p-value at 0.001. Therefore, 999 Monte Carlo replications were performed.

3 Results

Results of the STSSP model of the analysis completed for datasets I, II and III are presented in Table 1. Figs. 5, 6 and 7 display a GIS visualization of the localization of the resulted clusters for each sub-period. Ten significant clusters were distinguished for Canton Ticino: two clusters in the sub-period of 1969-1978, and four in each of the two sub-periods 1979-1990 and 1991-2008. In the three analyses, the uncovered clusters were predominantly localized in the hilly deciduous forest belt of the southern part of the study area (*Sottoceneri*), where the population density is highest. These clusters are mainly defined by 74 fires of anthropogenic origins, located in altitudes ranging from the 300 to 1,250 m.a.s.l. and with slopes varying from moderate to very strong. They mostly started in areas within 50 m of the forest edge (23 fires), broadleaves forests (17 fires), chestnut stands (14 fires) and the 21 remaining burned in coniferous forest with broadleaves and in areas within 50-100 m of the forest edge. From the seasonal point of view, six out of ten clusters ignited during the winter period (November – April), while the other four occurred in the summer season (May – November).

The ten detected clusters are not geographically persistent, i.e. they do not appear at the same location in different times. In spite of this, seven clusters were identified in two zones adjacent to three of the five major urban areas in the canton. The first zone is in the very southern part, around the cities of Mendrisio and Chiasso, where three clusters – one from each sub-period – are exhibited. These clusters are constituted by fire events spotted in the winter season, mainly burnt in the areas within 50 m of the forest edge and all started by anthropogenic actions. In this region, cluster 1 from the sub-period 1969-1978 (dataset I) has

**Table 1** Results of the SSSP analysis

Study period	Cluster	Radius (m)	Time frame	Observed cases	Expected cases	p-value
1969- 1978	1	2987	01-02-1973 28-02-1973	7	0.40	0.007
	2	40	01-11-1978 30-11-1978	3	0.03	0.050
1979–1990	1	2212	01-12-1986 31-01-1987	10	0.69	0.001
	2	2213	01-01-1984 31-01-1985	21	5.08	0.001
	3	1639	01-01-1981 31-01-1981	8	0.67	0.006
	4	1388	01-04-1981 30-06-1981	4	0.08	0.034
1991–2008	1	1440	01-10-1997 31-10-1997	7	0.13	0.001
	2	0.00	01-02-2001 28-02-2001	4	0.03	0.002
	3	2511	01-05-1997 31-08-1997	6	0.27	0.012
	4	999	01-03-1992 31-03-1992	4	0.07	0.018
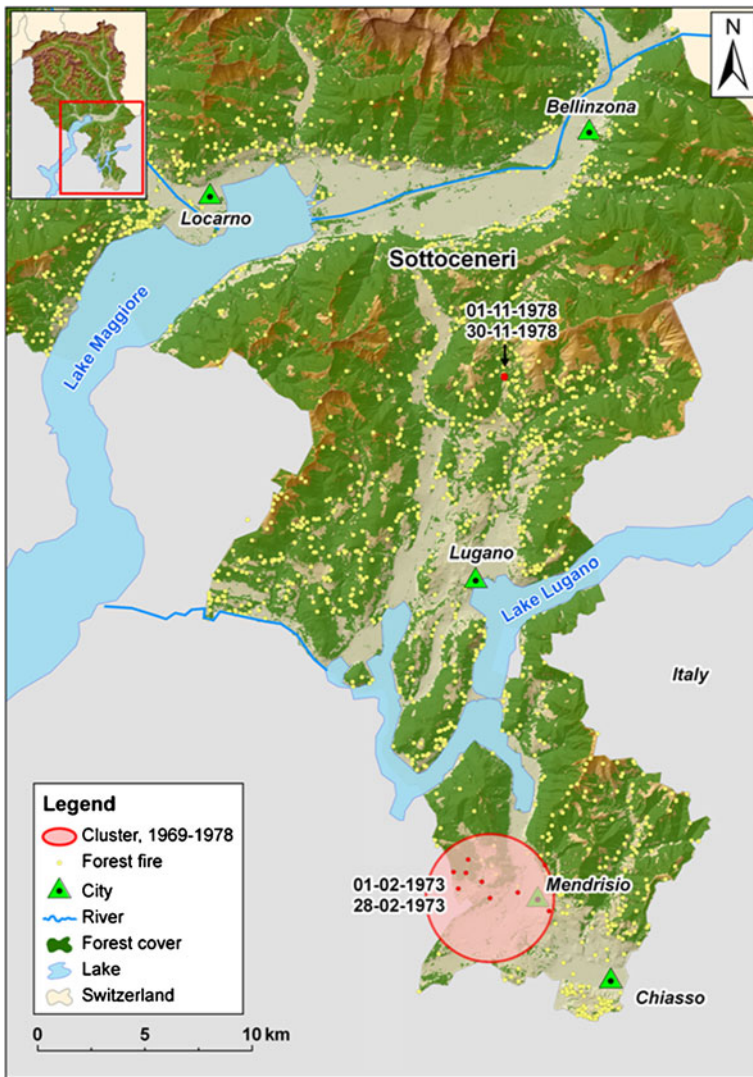
Fig. 5 Clusters of forest fires in canton Ticino in dataset I (1969-1979). STSSP model

the greatest radius range of all ten identified clusters with a value of 2,987 m spotted in February of 1973; see Table 1 and Fig. 5. The second cluster in this area is cluster 2 from sub-period 1979-1990 (dataset II) presenting the longest time frame: year-length. Though, it holds 21 observed events, only 5.08 cases were expected (Table 1).

The second zone is the central part of the *Sottoceneri* region near Lugano city. Four clusters have been recognized: two from dataset II and two from dataset III. These clusters were largely ignited by anthropogenic sources with a domination of criminal actions (arson), except for one fire ignited by lightning in cluster 4 in dataset II. It is the only fire in all identified clusters triggered by a natural source. Cluster 4 has the particularity of being detected in both winter and summer seasons with a time frame from April 1981 to June 1981. All clusters in this central region were principally burnt during the winter season in
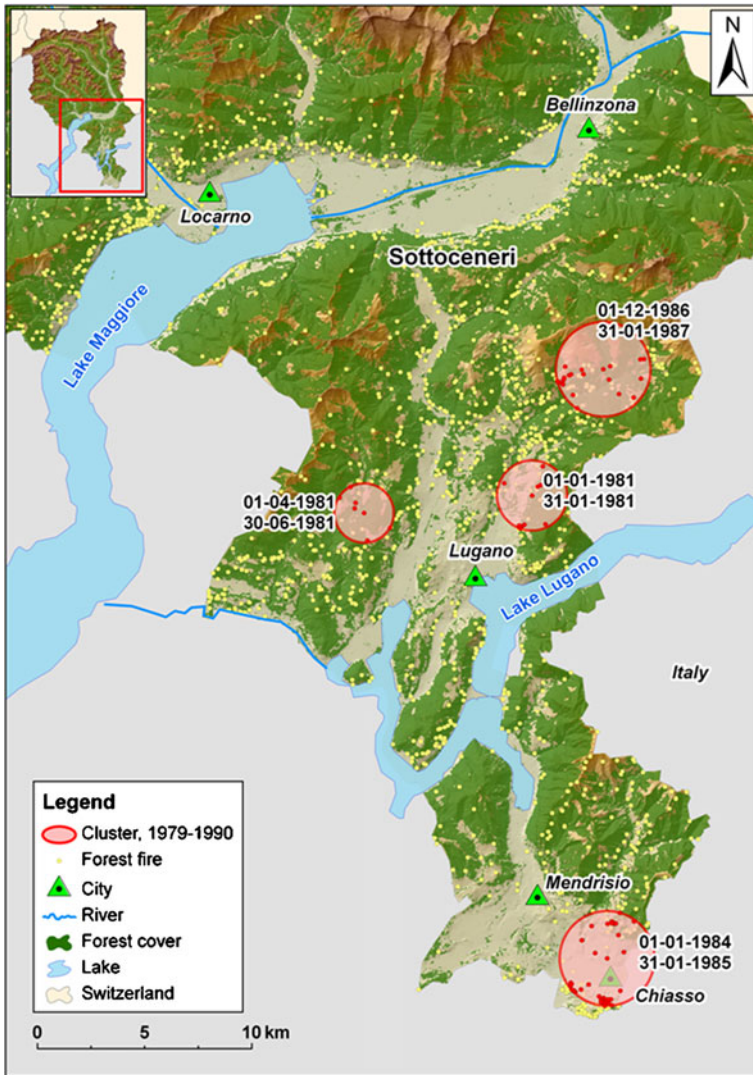
Fig. 6 Clusters of forest fires in canton Ticino in dataset II (1979-1991). STSSP model

broadleaves forests and chestnut stands. In this zone a particular cluster (cluster 2 from dataset III) is indicated by 4 fires, all detonated by arson actions at the same geographical location (cluster radius: 0).

Forest fires caused by lightning display very different spatio-temporal patterns from those of anthropogenic origins; consequently, it is indispensable to make a distinction between forest fires due to natural causes and to anthropogenic activities [29, 40, 43]. For this reason, cluster analyses for lightning fires were conducted in a separately dataset. The outcome clusters are reported in Table 2 and displayed in Fig. 8. Two significant clusters were detected in the mountainous coniferous forests in the northern part of the study area, known as *Sopraceneri*, where higher altitudes and slopes are found and less development is settled. It is important to call to mind that the original geo-database reports only lightning-fires
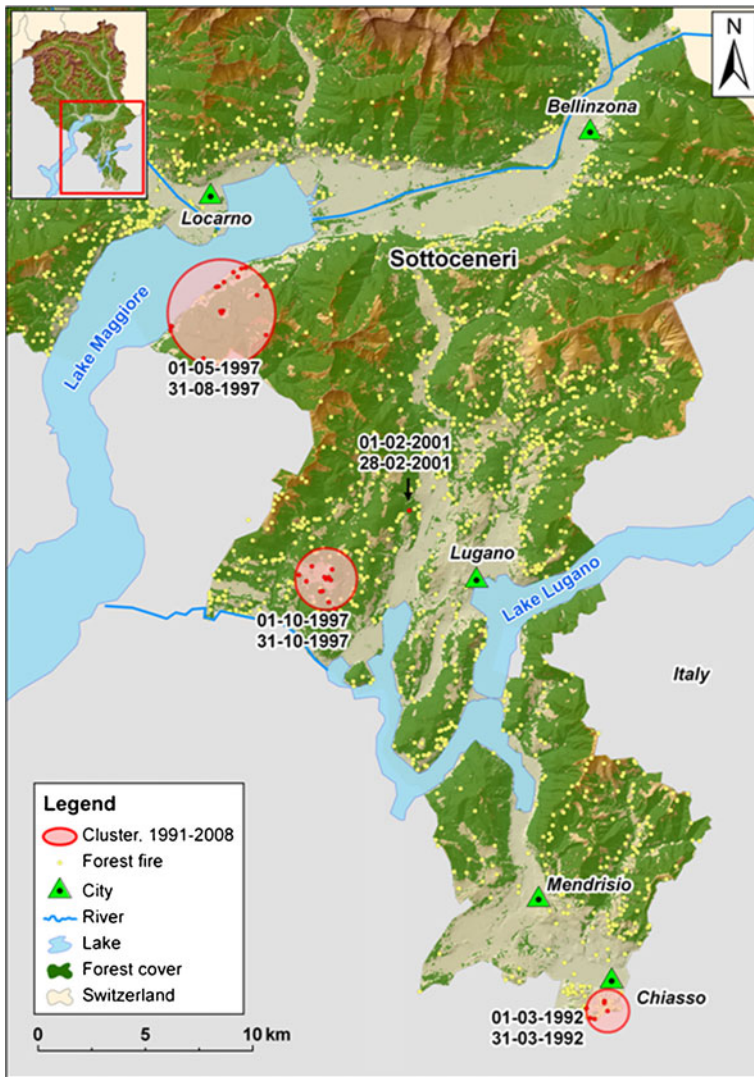
Fig. 7 Clusters of forest fires in canton Ticino in dataset III (1991-2008). STSSP model

Table 2 Results of the SSSP analysis for Lightning-induced fires

Study period	Cluster	Radius (m)	Time frame	Observed cases	Expected cases	p-value
1969–2008	1	1375	30-06-1989 12-08-1989	3	0.051	0.007
	2	200	17-08-1997 30-09-1997	3	0.069	0.032
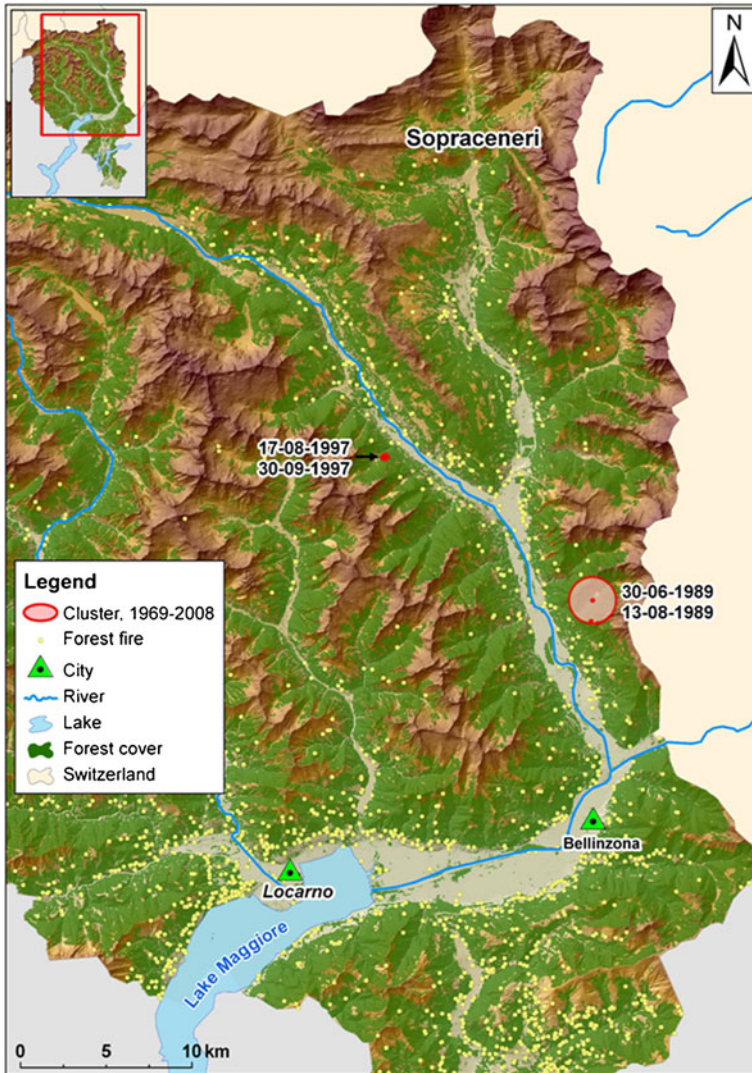
Fig. 8 Clusters of lightning-induced forest fires in canton Ticino in the entire period (1969-2008). STSSP model

during the summer period (May to November); consequently cluster time frames cannot fall outside of this temporal range. Cluster 1 comprises 3 fires reported in June-August 1989 with altitudes ranging from 1200 to 1720 m.a.s.l. at strong slopes (25-38 %); while cluster 2 is defined by 3 fires that occurred in August-September 1997, at lower altitudes than cluster 1, varying from 1017-1120 m.a.s.l., extreme slopes between 85 and 112 (percent) and all burned in spruce stands.

   As indicated in Section 1 (Introduction), the detection of the spatio-temporal extension of the fire risk, which was achieved with the scan statistics permutation model, is a very important task for better management and implementation of fire-fighting measures. However, the analysis of the predominant ignition-cause inside the detected clusters can

be of great utility when mitigating different fire related problems; for instance, repeated arson actions characterizing the structure of the ignition causes inside the clusters.

With the purpose of identifying temporary risk factors in the detected hot spots, due to particular and extemporaneous local conditions, the distribution of the ignition-causes in each single disclosed cluster was compared with respect to the whole area during the same time-frame of the cluster. Results of these analyses are shown in Table 3; the ratio value highlights the importance of the ignition-causes inside the cluster compared to the same causes in the same frame-period over the area outside the cluster.

Clusters from dataset I exhibit mainly frequencies of fires generated by negligent human actions. Dataset II reveals clusters with incidence of arson, negligence and unknown causes in almost all clusters with few exceptions. Looking at the ratio values for these ignition-sources, the two first clusters expose an arson frequency greater than the other causes, whereas, unknown causes are dominant in cluster 3 and lightning in cluster 4.

Dataset III puts in evidence a high prevalence of arson fires in clusters 1 and 2. In the first detected cluster, all fires were caused by arson actions; while in the second cluster, arson has an incidence of 8.6 times more than the average of the analysed sub-period. On the other hand, in the third and the fourth clusters, no arson is encountered.

4 Discussions

In the present study a cluster analysis was performed to explore forest fire ignition point pattern distribution in Canton Ticino (Switzerland), to detect spatio-temporal clusters at local scales and to evaluate their statistical significance. The space-time scan statistic permutation (STSSP) model made it possible to identify and to locate clusters that are space-time dependent using only the observed events. This paper highlights the efficacy of this method for environmental data analysis, particularly, when a population-at-risk is not easily detectable. The exhaustiveness of the fire-database from Canton Ticino allowed organizing the events in several groups taking into account the ignition-causes and the date of first alarm (ignition date).

Table 3 Relative frequency (RF%) of fire causes inside clusters and the ratio compared with the fire causes outside clusters for the same frame-period

Study period	Cluster	Arson		Private		Unknown		Lightning	
		RF (%)	Ratio	RF (%)	Ratio	RF (%)	Ratio	RF (%)	Ratio
1969–1978	1	0.0	0.0	85.7	1.4	14.3	0.5	0.0	–
	2	0.0	0.0	66.7	1.4	33.3	1.0	0.0	–
1979–1990	1	60.0	2.3	30.0	1.0	10.0	0.3	0.0	–
	2	33.3	1.9	52.4	1.5	14.3	0.5	0.0	–
	3	37.5	0.8	0.0	0.0	62.5	3.0	0.0	–
	4	0.0	0.0	0.0	0.0	75.0	2.3	25.0	5.5
1991–2008	1	100.0	[a]	0.0	–	0.0	0.0	0.0	–
	2	100.0	8.6	0.0	0.00	0.0	0.0	0.0	–
	3	0.0	0.0	33.3	1.6	16.7	1.0	0.0	0.0
	4	0.0	0.0	50.	1.6	0.0	0.0	0.0	–

[a] no arson outside the cluster

Two statistically significant clusters for dataset I (1969-1978) and four for dataset II (1979-1990) and dataset III (1991-2008) were unmasked and analyzed. All of them are localized in the southern part of Canton Ticino, called *Sottoceneri* (on the south of Ceneri Mountain), which encloses the urbanized areas with the highest population density in the Canton. The major cities of the canton in this zone are: Chiasso, Mendrisio, Lugano, Bellinzona and Locarno. The frame periods of all clusters fall mainly in the winter season (from December to April) and their major prevalent ignition-causes are of anthropogenic activities like arson or negligence fires. Only four out of the ten clusters fall in summer season, and only one fire inside these clusters originated from lightning.

It is also noteworthy that two of the ten detected clusters have very small radii ranging from 0 m (cluster 2 in dataset III, see Table 1 and Fig. 7) up to 40 m (cluster 2 in dataset I, see Table 1 and Fig. 5). In the first case, the cluster consists of four arson strikes at the same site within 23 days (February 1st to February 23rd 2001). Fortunately all four fires were rapidly controlled by the fire brigades and their final burnt area never exceeded 600 m^2. In the second case, the fire brigade's reports referred to escaped-pasture fires from contiguous lots of land. These situations probably had, significantly, restricted the correspondent spatio-temporal clusters, constituting a temporary hotspot. In three cases, clusters 1 and 2 from dataset II and cluster 1 from dataset III, presumed serial arsonists were intercepted or even arrested.

In the analyses of the lightning fires, carried out over the entire study period (1969-2008), two clusters were detected: one in the summer of 1989 and the other one in the summer of 1997. Both of them are located along the valleys of the high east Ticino, not far from the path of Ticino's river.

The resultant clusters may be due either to an increase risk of fire or to an inhomogeneous increase/decrease of the geographical biomass distribution over the entire study period. This means that the detected clusters could be products of either a temporary increase in fire-ignition risk (e.g. arson activities, pasture-fire practices during non-fire-weather conditions, negligent behaviour) or from changes in structural-ignition conditions just for specific zones in a particular time-frame (e.g. reforestation, afforestation, changes in fuel load or fuel distribution, increased probability of anthropogenic fire ignition due to changes in population distribution and/or behaviour). These two typologies are not automatically distinguished by the STSSP model when dealing with data spanning several years and, consequently, final results can be biased.

Considering this model's constraints, cantonal authorities' actions that may have influenced the behaviour of fires (frequency and size) in the recent decades led to the selection of the three sub-periods: 1969-1978, 1979-1990, and 1991-2008. These legal acts and technical measures, aiming to reduce and to prevent the ignition of anthropogenic forest fires, were considered in the establishment of the dataset parameters (refer to the Section 2.2).

The prevalence of the human-negligence actions in the detected clusters from the sub-period 1969-1978 (dataset I) reflects the problematic condition of forest fire regime in Canton Ticino, in the period between earlier 50's to late 70's, when fire frequency increased dramatically, altering the risk of fire, as a consequence of the abandonment of traditional agricultural and land management activities [5, 6]. This situation of structured fires, led cantonal authorities' attention to focus on the development of preventative measures towards the mitigation of negligence-caused fires, and towards early fire detection and firefighting such as the implementation of fire guards, alert systems, information boards and announcements, prohibition of burning garden debris in the open, hydrant nets, among other technical measures [6].

The presented approach retrospectively detects temporaneous hotspot and structural fires, thus represents a useful tool for assisting authorities to better design and target fire pre-suppression and fire-fighting facilities and resources at local level.

5 Conclusions

The detailed fire database from Canton Ticino allowed the identification of the fires comprised in each of the detected ten clusters. For anthropogenic-triggered fires this investigation, combined with the results of the STSSP analyses (clusters location and frame period), provides very useful information concerning the characteristics of the clusters and their structural and temporal components. Similarly, the detection of lightning-fire clusters could be a starting point for a deeper analysis for fire prevention in the mountainous regions, in view of the dryer and lightning-fire richer summer periods in the future [7].

This methodology can be very computationally intensive, taking up to several days or weeks. The computing time depends on a wide variety of variables such as the size of the input dataset, the number of time intervals and the chosen analytical options. Depending on these requirements, the SaTScan user guide [19] provides formulas for an approximate calculation of the computing time and the memory requirements. Yet, the presented approach revealed to be very valuable and flexible, enabling analyses at different time and space windows, as well as, considering different ignition sources or frame periods separately. Compared with other cluster methods, space-time scan statistics has the advantage to detect both cluster's location and frame period, while also testing their statistical significance.

A practical functionality of the proposed methodology, to support decision-making, is the incorporation of the model outputs into a geographical information system, to map and identify fire-prone zones. Outcomes of the STSSP model simulations were integrated into a GIS environment allowing the exposition of the detected clusters. These analyses revealed valuable information of the structure of the forest fire clusters in Canton Ticino. Spatio-temporal components of the forest fires regime in Ticino, such as the frequency incidence of human and natural causes, were clearly identified. These results can assist the Forest Service and policy makers in the classification of hotspots into structural and temporary fires which is useful for a better definition of the infrastructural and organizational measures required to mitigate the fire risk in each of these types of fire-hotspots. Consequently, for fire risk management, this approach can also be advantageous to define and to spot at-risk areas.

In practical terms, and in the specific case of the study area, structural fires should be taken into consideration when planning permanent and expensive fire pre-suppression and fighting facilities, such as helicopter water points or water hydrant nets; whereas, detected temporary hot spots should be considered in the frame of preventative educational training, in the case of negligence, or intelligence-police actions in the case of repeated arson.

References

1. Allan DW (1966) Statistics of atomic frequency standards. Proc of the IEEE 54(2):221–230
2. Brändli UB (2000) Waldzunahme in der Schweiz—gestern und morgen. Inf sblatt Forsch bereich Landsch 45:1–4
3. Byrne K (2001) Environmental science (Bath Advanced Science). Nelson Thornes, Cheltenham
4. Ceschi I (2006) Il bosco nel Canton Ticino. Armando Dadò Editore, Locarno

5. Conedera M (2003) Incendi di boschi in Canton Ticino: dallo studio pionieristico di Ceschi all situazione attuale. Bollettino della Società ticinese di Scienze naturali 91(1–2):135–144

6. Conedera M (2009) Implementing fire history and fire ecology in fire risk assessment: the study case of Canton Ticino (southern Switzerland). Dissertation, Universität Fridericiana zu Karlsruhe

7. Conedera M, Cesti G, Pezzatti GB, Zumbrunnen S (2006) Lightning-induced fires in the Alpine region: an increasing problem. V Int Conf on Forest Fire Res, D.X. Viegas (Ed.):1-9

8. Ester M, Kriegel H-P, Sander J, Xu X (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Proc of 2nd Int Conf on Knowledge Discovery and Data Mining (KDD-96)

9. Finney M (2005) The challenge of quantitative risk analysis for wildland fire. Forest Ecol and Manag 211 (1–2):97–108

10. FSO (2009) Forestry in Switzerland, pocket statistics 2009. Federal Statistical Office, Federal Department of Home Affairs FDHA, Swiss Confederation

11. Genton MG, Butry DT, Gumpertz ML, Prestemon JP (2006) Spatio-temporal analysis of wildfire ignitions in the St Johns River water management district, Florida. Int J Wildland Fire 15:87–97

12. Hering AS, Bell CL, Genton MG (2009) Modeling spatio-temporal wildfire ignition point patterns. Environ and Ecolo Stat 16(2):225–250

13. Hessburg PF, Reynolds KM, Keane RE, James KM, Salter RB (2007) Evaluating wildland fire danger and prioritizing vegetation and fuels treatments. For Ecol and Manag 247(1–3):1–17

14. Kanevski M (Editor) (2008) Advanced mapping of environmental data: geostatistics, machine learning and Bayesian maximum entropy, iSTE and WILEY, Great Britain and United States

15. Kanevski M, Maignan M (2004) Analysis and modelling of spatial environmental data. EPFL, Lausanne

16. Krawchuk MA, Moritz MA, Parisien MA, Van Dorn J, Hayhoe K (2009) Global pyrogeography: the current and future distribution of wildfire. PloS ONE 4(4, e5102):1-12

17. Krebs P, Pezzatti G, Mazzoleni S, Talbot L, Conedera M (2010) Fire regime: history and definition of a key concept in disturbance ecology. Theory Biosci 129(1):53–69

18. Kulldorff M (1997) A spatial scan statistic. Commun in Stat: Theory and methods 26(6):1481–1496

19. Kulldorff M (2009) SaTScan^{TM} User Guide for version 8.0. Technical Documentation

20. Kulldorff M, Athas WF, Feuer EJ, Miller BA, Key CR (1998) Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos. American J of Public Health 88(9):1377–1380

21. Kulldorff M, Heffeman R, Hartman J, Assunção R, Mostashari F (2005) A space-time permutation scan statistic for disease outbreak detection. PLoS Med 2:216–224

22. Lasaponara R, Santulli A, Telesca L (2004) Time-clustering analysis of forest-fire sequences in southern Italy. Chaos, Solitons and Fractals 24(1):139–149

23. Lovejoy S, Schertzer D, Ladoy P (1986) Fractal characterization of inhomogeneous geophysical measuring networks. Nature 319:43–44

24. Moran P (1950) Notes on continuous stochastic phenomena. Biometrika 37(1–2):17–23

25. Naus J (1965) Clustering of random points in two dimensions. Biometrika 52:263–267

26. Naus J (1965) The distribution of the size of maximum cluster of points on the line. J Am Stat Assoc 60:532–538

27. Openshaw S, Charlton M, Wymer C, Craft A (1987) A mark I geographical analysis machine for the automated analysis of point data sets. Int J of Geogr Inf Sci 1(4):335–358

28. Pezzatti GB, Reinhard M, Conedera M (2010) Swissfire: die neue schweizerische Waldbranddatenbank. Schweiz Z Forstwes 161(12):465–469

29. Podur J, Martell DL, Csillag F (2003) Spatial patterns of lightning-caused forest fires in Ontario, 1976-1998. Ecol Model 164:1–20

30. Reynolds KM, Hessburg PF (2005) Decision support for integrated landscape evaluation and restoration planning. For Ecol and Manag 207(1–2):263–278

31. Ripley BD (1976) The second-order analysis of stationary point processes. J of Applied Probability 13 (2):255–266

32. Sousa WP (1984) The role of disturbance in natural communities. Annu Rev Ecol Syst 15:353–391

33. Spinedi F, Isotta F (2004) Il clima del Ticino negli ultimi 50 anni. Dati Statistiche e Società 4:4–39

34. Telesca L, Amatulli G, Lasaponara R, Lovallo M, Santulli A (2004) Time-scaling properties in forest-fire sequences observed in Gargano area (southern Italy). Ecol Model 185(2–4):531–544

35. Telesca L, Kanevski M, Tonini M, Pezzatti G, Conedera M (2010) Temporal patterns of fire sequences observed in canton of Ticino. Nat Hazards and Earth Syst Sci 10(4):723–728

36. Telesca L, Pereira MG (2010) Time-clustering investigation of fire temporal fluctuations in Portugal. Nat Hazards and Earth Syst Sci 10(4):661–666
37. Tonini M, Tuia D, Ratle F (2009) Detection of clusters using space-time scan statistics. Int J Wildland Fires 18(7):830–836
38. Tuia D, Ratle F, Lasaponara R, Telesca L, Kanevski M (2008) Scan statistics analysis of forest fire clusters. Commun in Nonlinear Sci and Numer Simul 13(8):1689–1694
39. Turnbull BW, Iwano EJ, Burnett WS, Howe HL, Clark LC (1990) Monitoring for clusters of disease: application to leukemia incidence in upstate New York. Amercian J of Epidemiol 132(1):S136–143
40. Tuner R (2009) Point patterns of forest fire locations. Environ Ecol Stat 16:197–223. doi:10.1007/s10651-007-0085-1
41. Vega Orozco C, Kanevski M, Tonini M, Conedera M (2010) Geostatistical analysis of spatio-temporal forest fire data. Proc of the EGU General Assembly 2010, vol 12, EGU2010-3235
42. Vega Orozco C, Tonini M, Kanevski M, Conedera M (2011) Point pattern analysis of forest fire occurrence in Canton Ticino (Switzerland). Proc of the Int Conf on Fire Behavior and Risk Modeling 2011, Alghero, Italy
43. Wotton BM, Martell DL (2005) A lightning fire occurrence model for Ontario. Canadian J Forestry Res 35:1389–1401
44. Zumbrunnen T, Bugmann H, Conedera M, Bürgi M (2009) Linking forest fire regimes and climate—a historical analysis in a dry inner alpine valley. Ecosyst 12(1):73–86

Highlights

We performed a space-time cluster analysis of forest fire sequences

The forest fire geo-database stores ignition-points and burnt area information

Analyses were performed considering both anthropogenic and naturally caused fires

Several statistically significant clusters were detected

Risk factor analyses of the detected clusters were carried out identifying their ignition causes

Clusters are prevalently located near urban areas and are of anthropogenic origins

Carmen Delia Vega Orozco is a PhD at the IGAR institute, University of Lausanne (Switzerland). She holds a MSc in Environmental Geosciences at the University of Lausanne (Switzerland) in 2010 with a mention in analysis, monitoring and representation of natural hazards, a Master degree in Geography at West Chester University of Pennsylvania (United States) in 2007 with a mention in geographic information systems (GIS) and environmental planning, and a BS degree in Geomatic Engineering from the University of Valle of Cali (Colombia) in 2002. She has collaborated with private and public organizations engaged in environmental, urban and transportation systems studies. Her research experience focuses in the analysis of environmental phenomena throughout the application of statistical models, GIS and remote sensing.

Dr. Marj Tonini is specialist in geoinformatics supporting research and teaching at the Institute of Geomatics and Risk Analysis (IGAR), University of Lausanne (Switzerland). Dr. Tonini holds her PhD on agro-environmental modelling from the Sant'Anna School of advanced studies (Pisa, Italy) in 2002. She collaborated with different international research Institutes (e.g. Remote Sensing Laboratory of Ben Gurion University (Israel), Institute of Social and Preventive Medicine of Geneva University (Switzerland)) as well as private enterprises involved in environmental studies. Also co-supervised different master studies' at IGAR in cluster analyses.

Dr. Marco Conedera is the head of the Research Unit Community Ecology at the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) which includes also the Insubric Research Group located in Bellinzona. He has been working on fire history and fire ecology since 15 years, leading teams in the context of EU FP research projects (e.g., MinerveII, Prometheus s.v., Fire paradox) and Interreg Projects (Alpffirs). His research work mainly focuses on forest history and human impacts on forest fires, fire ecology, post-fire risk evaluation, and fire management. He is also member of the Forest fire Expert Group of the European Commission.

Prof. Mikhaïl Kanevski is professor at the Institute of Geomatics and Risk Analysis at the University of Lausanne (Switzerland) within the framework of the Faculty of Geosciences and Environment (FGSE). Professor Kanevski is author and co-author of 3 books and more than 150 publications. He has valuable experience and knowledge in relevant fields including geostatistics, machine learning algorithms (Artificial Neural Networks, Support Vector Machines), geomatics (Geographical Information Systems and remote sensing) and geosimulations providing algorithmic solutions to environmental problems, with application in environmental, socio-economic and financial data mining.