Year : 2016

Analysis of Space-Time Environmental Data in Complex Regions

Orozco Vega Carmen Delia

UNIL | Université de Lausanne

Faculté des Géosciences et de l'Environnement
Institut des Dynamiques de la Surface Terrestre

Analysis of Space-Time Environmental Data in Complex Regions

Thèse de doctorat

Présentée à la

Faculté des Géosciences et de l'Environnement
de l'Université de Lausanne
par

Carmen Delia Vega Orozco

B.Sc Universidad del Valle, Colombie
M.A West Chester University of Pennsylvania, États-Unis
M.Sc Université de Lausanne, Suisse

diplômée en

Docteur en Sciences de l'Environnement

Jury

Directeur de Thèse	Prof. Mikhail Kanevski
Co-directeur de Thèse	Dr. Marj Tonini
Expert interne	Prof. François Bavaud
Expert externe	Prof. Mário G. Pereira
Expert externe	Dr. Michael Reinhard
Président du colloque	Prof. Suren Erkman

Lausanne, 2016

Unil
UNIL | Université de Lausanne
Décanat Géosciences et de l'Environnement
bâtiment Géopolis
CH-1015 Lausanne

IMPRIMATUR

Vu le rapport présenté par le jury d'examen, composé de

Président de la séance publique :	M. le Professeur Suren Erkman
Président du colloque :	M. le Professeur Suren Erkman
Directeur de thèse :	M. le Professeur Mikhail Kanevski
Co-directrice de thèse:	Mme le Docteur Marj Tonini
Expert interne :	M. le Professeur François Bavaud
Expert externe :	M. le Professeur Mario Gonzalez Pereira
Expert externe :	M. le Docteur Michael Reinhard

Le Doyen de la Faculté des géosciences et de l'environnement autorise l'impression de la thèse de

Madame Carmen Delia VEGA OROZCO

Titulaire d'un
Master en géographie, Université West Chester/Pennsylvanie, USA
et d'une
Maîtrise universitaire ès Sciences en géosciences et environnement, mention Analyse, monitoring et représentation des dangers naturels, Lausanne

intitulée

ANALYSIS OF SPACE-TIME ENVIRONMENTAL DATA IN COMPLEX REGIONS

Lausanne, le 21 janvier 2016

Pour le Doyen de la Faculté des géosciences et de l'environnement

S. E.

Professeur Suren Erkman, Vice-doyen

Analysis of Space-Time Environmental Data in Complex Regions

Carmen Delia Vega Orozco

Institute of Earth Surface Dynamics

Summary

This thesis develops a comprehensive and a flexible statistical framework for the analysis and detection of space, time and space-time clusters of environmental point data. The developed clustering methods were applied in both simulated datasets and real-world environmental phenomena; however, only the cases of forest fires in Canton of Ticino (Switzerland) and in Portugal are expounded in this document.

Normally, environmental phenomena can be modelled as stochastic point processes where each event, e.g. the forest fire ignition point, is characterised by its spatial location and occurrence in time. Additionally, information such as burned area, ignition causes, landuse, topographic, climatic and meteorological features, etc., can also be used to characterise the studied phenomenon. Thereby, the space-time pattern characterisation represents a powerful tool to understand the distribution and behaviour of the events and their correlation with underlying processes, for instance, socio-economic, environmental and meteorological factors. Consequently, we propose a methodology based on the adaptation and application of statistical and fractal point process measures for both global (e.g. the Morisita Index, the Box-counting fractal method, the multifractal formalism and the Ripley's K-function) and local (e.g. Scan Statistics) analysis.

Many measures describing the space-time distribution of environmental phenomena have been proposed in a wide variety of disciplines; nevertheless, most of these measures are of global character and do not consider complex spatial constraints, high variability and multivariate nature of the events. Therefore, we proposed an statistical framework that takes into account the complexities of the geographical space, where phenomena take place, by introducing the Validity Domain concept and carrying out clustering analyses in data with different constrained geographical spaces, hence, assessing the relative degree of clustering of the real distribution.

Moreover, exclusively to the forest fire case, this research proposes two new methodologies to defining and mapping both the Wildland-Urban Interface (WUI) described as the interaction zone between burnable vegetation and anthropogenic infrastructures, and the prediction of fire ignition susceptibility.

In this regard, the main objective of this Thesis was to carry out a basic statistical/-geospatial research with a strong application part to analyse and to describe complex phenomena as well as to overcome unsolved methodological problems in the characterisation of space-time patterns, in particular, the forest fire occurrences. Thus, this Thesis provides a response to the increasing demand for both environmental monitoring and management tools for the assessment of natural and anthropogenic hazards and risks, sustainable development, retrospective success analysis, etc. The major contributions of this work were presented at national and international conferences and published in 5 scientific journals. National and international collaborations were also established and successfully accomplished.

Analysis of Space-Time Environmental Data in Complex Regions

Carmen Delia Vega Orozco

Institut des Dynamiques de la Surface Terrestre

Résumé

Cette thèse développe une méthodologie statistique complète et flexible pour l'analyse et la détection des structures spatiales, temporelles et spatio-temporelles de données environnementales représentées comme de semis de points. Les méthodes ici développées ont été appliquées aux jeux de données simulées autant qu'à des phénomènes environnementaux réels; nonobstant, seulement le cas des feux forestiers dans le Canton du Tessin (la Suisse) et celui de Portugal sont expliqués dans ce document.

Normalement, les phénomènes environnementaux peuvent être modélisés comme de processus ponctuels stochastiques où chaque événement, par ex. les point d'ignition des feux forestiers, est déterminé par son emplacement spatial et son occurrence dans le temps. De plus, des informations tels que la surface brûlée, les causes d'ignition, l'utilisation du sol, les caractéristiques topographiques, climatiques et météorologiques, etc., peuvent aussi être utilisées pour caractériser le phénomène étudié. Par conséquent, la définition de la structure spatio-temporelle représente un outil puissant pour comprendre la distribution du phénomène et sa corrélation avec des processus sous-jacents tels que les facteurs socio-économiques, environnementaux et météorologiques. De ce fait, nous proposons une méthodologie basée sur l'adaptation et l'application de mesures statistiques et fractales des processus ponctuels d'analyse global (par ex. l'indice de Morisita, la dimension fractale par comptage de boîtes, le formalisme multifractal et la fonction K de Ripley) et local (par ex. la statistique de scan).

Des nombreuses mesures décrivant les structures spatio-temporelles de phénomènes environnementaux peuvent être trouvées dans la littérature. Néanmoins, la plupart de ces mesures sont de caractère global et ne considèrent pas de contraintes spatiales complexes, ainsi que la haute variabilité et la nature multivariée des événements. A cet effet, la méthodologie ici proposée prend en compte les complexités de l'espace géographique où le phénomène a lieu, à travers de l'introduction du concept de Domaine de Validité et l'application des mesures d'analyse spatiale dans des données en présentant différentes contraintes géographiques. Cela permet l'évaluation du degré relatif d'agrégation spatiale/temporelle des structures du phénomène observé.

En plus, exclusif au cas de feux forestiers, cette recherche propose aussi deux nouvelles méthodologies pour la définition et la cartographie des zones périurbaines, décrites comme des espaces anthropogéniques à proximité de la végétation sauvage ou de la forêt, et de la prédiction de la susceptibilité à l'ignition de feu.

À cet égard, l'objectif principal de cette Thèse a été d'effectuer une recherche statistique/géospatiale avec une forte application dans des cas réels, pour analyser et décrire des phénomènes environnementaux complexes aussi bien que surmonter des problèmes méthodologiques non résolus relatifs à la caractérisation des structures spatio-temporelles, particulièrement, celles des occurrences de feux forestières. Ainsi, cette Thèse fournit une réponse à la demande croissante de la gestion et du monitoring environnemental pour le déploiement d'outils d'évaluation des risques et des dangers naturels et anthropogéniques. Les majeures contributions de ce travail ont été présentées aux conférences nationales et internationales, et ont été aussi publiées dans 5 revues internationales avec comité de lecture. Des collaborations nationales et internationales ont été aussi établies et accomplies avec succès.

"The ability to take data -to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it- that's going to be a hugely important skill in the next decades, ... because now we really do have essentially free and ubiquitous data. So the complimentary scarce factor is the ability to understand that data and extract value from it."

Hal Varian, Google's Chief Economist
The McKinsey Quarterly, Jan 2009

Acknowledgements

First, I would like to express my special thanks to my supervisors Prof. Mikhail Kanevski and Dr. Marj Tonini for your priceless guidance, trust and advices on the scientific work. I greatly appreciate all your contributions of time, ideas and funding to make my Ph.D. experience productive and stimulating and to allow me to grow as a research scientist. Your joy and enthusiasm for the research was contagious and motivational for me. I am really honoured to have been able to work with both you; you have been tremendous mentors for me.

I would also like to acknowledge honorary my committee members Prof. Jorge Mário Gonzalez Pereira of the University of Trás-os-Montes and Alto Douro (Portugal), Dr. Michael Reinhard of the Swiss Federal Office for the Environment and Prof. François Bavaud of the University of Lausanne, for your thoughtful reading, for your constructive and brilliant comments and suggestions, and mostly for letting my defense be an enjoyable moment.

I would also like to express my sincere gratitude to Dr. Marco Conedera from the WSL for supplying part of the data used in this Thesis, for enriching my work with professional knowledge on the forest fire topic and for the marvellous collaboration we carried out since 2009. I would also like to thank the CITAB team of the University of Trás-os-Montes and Alto Douro for the prosperous and agreeable research collaboration.

A special thanks to my colleagues from the University of Lausanne for the stimulating discussions, the conference trips, and all the delightful moments we had in all these years. There are also a significant number of people who are important and influential in my life. For that, I thank my friends from different parts of the world, my American Jackson family, my family-in-law and my Vega-Orozco families for your priceless and unsurpassed love, faith, and the enjoyable moments we always have together.

Lastly, the most important people in my life. I thank my mother Nora and my father Carlos for your immeasurable love, sacrifices, devotion and dedication during my endless 26 years of studies and pursuits. My unique lovely and best sister Paula for your great humour and laughs, confidence, motivation and your beautiful family (Robert and my nephews). And most of all, I thank my beloved husband Fernando for all the trust, encourage, loyalty, love, enthusiasm and the incentive to strive me towards my goals; and for sharing with me the adventure of being parents of our daughter Karla and the coming-soon baby girl.

Words cannot express how grateful I am to all of you!

Carmen D. Vega Orozco, February 2016

Financial Support

The author of the Thesis would like to kindly acknowledge the SNSF for the financial support. This thesis was partly supported by the Swiss National Science Foundation project "Analysis and Modelling of Space-Time Patterns in Complex Regions", under the grant number 200021-140658/1.

Contents

List of Figures

List of Tables

Glossary

Clustering
: When points tend to be closer together. Here, the concept of "clustering" does not imply that the points are organised into identifiable "clusters"; instead, it refers to points that are closer together than expected for a Poisson process (CSR) [12].

Complexity
: High spatial heterogeneity, high variability of spatial/temporal clustering, and multivariate nature of a studied phenomenon.

CSR
: Complete spatial randomness. The probability that an event can occur at any point is equally likely to occur anywhere within a bounded region A and that its position is independent of each any other event.

Event
: outcomes of a stochastic process. In spatial point process, events are represented by points and marks; where the points describe the locations of the events, and the marks provide additional information characterising the event further [143].

GIS
: Geographic information system.

Homogeneity
: Equal distribution across a particular space, with constant intensity.

Inhomogeneity
: Also known as "non-uniform". The intensity λ of an inhomogeneous process varies from location to location. Thus, λ is a function of the location x [12, 143].

Interaction
: It is the stochastic dependence between the points in a point pattern. Usually we expect dependence to be strongest between points that are close to one another [12].

m.a.s.l.
: Metres above sea level.

Marked point processes
: When additional information characterising the events further (marks) are attached to the points.

Point processes
: It is a stochastic process that generates a countable set of points x_i in a topological space D (a Hausdorff space) [74, 90, 91, 143, 145, 193]. Each point represents the time and/or the spatial location of an event.

Poisson process
: A homogeneous Poisson process is a process where, conditional on $N(A)$, the events in a bounded region $A \subset \mathbb{R}^d$ are independently and uniformly distributed over A.

RF
: Random forests algorithm.

Space-time permutation scan statistics (STPSS)
It is a scan statistics method for the detection and identification of local clusters.

Susceptibility mapping
A map of the probability that an event occurs in a specific area without considering an absolute temporal scale.

WSL
Swiss Federal Institute for Forest, Snow and Landscape Research.

Wildland-urban interface (WUI)
This term is broadly used for indicating the interaction zones where human infrastructures meet or intermingle with natural vegetation such as grassland, shrub vegetation and forests [59, 130, 231, 232, 281].

Validity domain (VD)
It represents the geographical space of interest which constrains the studied space and reduces the dimensionality of the analysed process [150].

Symbols

Variables			
c_{zd}	number of observed cases within a zone z in a day d	c_A	number of observed cases in a cylinder A
C	total number of observed cases		
n_i	number of events in the ith box	N	total number of events
δ	box size (length of the diagonal)	t	length of a time-interval (timescale)
$N(\delta)$	number of boxes of size δ	$N(t)$	number of time-intervals of length t
$N_\alpha(\delta)$	number of boxes of size δ having singularity strength α	$N(r)$	number of points in a distance r
$p_i(\delta)$	the probability distribution in the ith box of size δ	$p_i(t)$	the probability distribution in the ith time-interval of length t
Q	number of boxes of size δ	Q_t	number of time-intervals of length t
q	order moment	x	vector of the location of the observed points in 1, 2 or 3-dimensional space
$\lambda(x_i)$	intensity: mean number of events occurring at location x_i	$\lambda(x_j)$	intensity: mean number of events at location x_j
λ	spatial intensity	λ_t	temporal intensity
$\|x_i - x_j\|$	distance between points x_i and x_j	$w(x_i, x_j)$	proportion of the circumference of a circle centred at x_i passing through x_j falling inside the study region

Functions			
α_i	the singularity strength in the i box of size δ	$\alpha(q)$	Hölder exponent of moment q
df_{box}	the Box-counting fractal dimension		
D_q	the Rényi generalised dimensions	tD_q	the temporal Rényi generalised dimensions
$f(\alpha)$	the singularity fractal dimension	$f(\alpha_q)$	the multifractal singularity spectrum
GLR	Generalized Likelihood Ratio		
I_δ	the Morisita index	I_t	the temporal Morisita index
$I_q(\delta)$	the Rényi information of qth order		
$K(r)$	Ripley's K-function	$L(r)$	L-function
$K(t)$	the temporal K-function	$L(t)$	the temporal L-function
K_{inhom}	the inhomogeneous K-function	L_{inhom}	the inhomogeneous L-function
$\lambda(x)$	Intensity function	μ_A	Expected number of cases for a space-time cylinder A
μ_A	Expected number of cases in a cylinder A	μ_{zd}	Expected number of cases per zone z and day d
$\mu_i(q,\delta)$	the normalised probability in the boxes of size δ of the qth order	$\tau(q)$	Mass exponent function of moment q

Chapter 1

Introduction

1.1 Motivation

In environmental science, the continue proliferation of the amount of high dimensional and high resolution data makes the spatio-temporal modelling a challenging and demanding task. These data portray objects that are often the results of heterogeneous sources creating space-time structures at multiple scales and with complex patterns. In this sense, the development of methodologies with a strong theoretical foundation and real world applications are necessary to deal with the complexities encountered in these data. Their correct use and treatment allow assessing, enhancing and providing necessary tools to support and to assist environmental monitoring and management, policy and decision-making, risk assessment and sustainable development.

Many of these data can be modelled as sets of points distributed in the space and/or in time, and thus, from a statistical context, they can be treated as realisations of stochastic point processes where the geographical and/or time components are taken into account. Nowadays, modern spatial statistical analysis [112, 192, 257, 311] offers powerful techniques to both characterise and understand the behaviour of the observed patterns as well as to infer about their underlying processes, e.g., the processes that may have given rise to the observed distribution. The application of these statistical methods gained more interest and development in a wide range of disciplines, notably by the blooming of geographical information systems (GIS) [110] because of its capability to manage large spatial databases and its flexibility at integrating analytical functionality providing suitable and valuable tools for the visualisation, exploration and analysis of spatial data [111, 123].

In this manner, we were motivated on developing an investigation that would tackle two major interests: on one hand, attending a scientific interest of developing a comprehensive spatial analysis by applying and adapting general methods of point processes to meet the needs of, i.e. environmental phenomena analysis, and to overcome unsolved methodological problems confronted when working with complex phenomena. And, on the other hand, the practical interest for applying this methodology in a real-world environmental events, in particular, forest fires. Several factors make this task difficult, i.e. the high spatial heterogeneity, high variability of the spatial/temporal clustering, and the multivariate nature of the phenomenon; from which arises the definition of complexity. Therefore, the success in the analysis not only depends on the availability of high quality data, but also on a qualified expert knowledge in both theoretical and application contexts.

This purpose impulsed us to creating the project named "*Analysis and Modelling of Space-Time Patterns in Complex Regions*"[1], from which the most important results are reported in the present thesis. Under this framework, this work portrays both the investigation of the spatial analysis for space-time pattern characterisation, and the description of how they can be implemented and adapted to the case of forest fire occurrences in the Swiss Alpine region. Although a broad range of methods for spatial analysis exist, we only focus on those we have found most useful to the studied events using different global and local measures with the use of statistical, fractal and second-order moments techniques. GIS tools are also applied for data preparation and visualisation. All these methods allow addressing fundamental questions about the space-time structures of the studied events and the factors triggering these pattern behaviours, such as: are the studied events clustered? and if so, what is their degree of clustering? where and when this clustering tends to occur? what are the processes that may have given rise to the observed event distribution (e.g. correlation with meteorological, environmental or socio-economic factors)?, etc. The application on forest fires illustrates many important features of interest for point process analysis, e.g. "clustering" caused by the spatial distribution of fire predisposing factors, "inhomogeneity" given by the various covariates, and "interaction" between the events (or occurrences).

On the other hand, with the knowledge gained by the statistical analysis in the first part of this investigation, a second study carried out in this project, exclusively related to the forest fire case and in collaboration with the Research Unit Community Ecology from the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) in Bellizona (see also Subsection 1.3.6 Collaborations), addresses the interest to defining and mapping critical and susceptible zones for fire ignition occurrences, for which fire managers currently focus their attention for fire management, monitoring, prevention and allocation measures. This area of interest is well known as the Wildland-Urban Interface (WUI) broadly used for indicating the interaction zones where human infrastructures meet or intermingle with natural vegetation such as grassland, shrub vegetation and forests [59, 130, 231, 232, 281].

In this regard, we propose a new GIS-based modelling approach for defining and mapping these subtle zones in the Swiss Alpine region. This also led us through the generation of a susceptibility map for fire occurrences, that is, a map of the probability of fire-ignition occurrence in specific areas without considering an absolute temporal scale. These two products are the first attempts for forest fires and WUI definition in a mountainous region as complex as the Swiss Alps.

[1]Supported by the Swiss National Science Foundation. Project No. 200021-140658/1.

Consequently, the spatio-temporal statistical analysis of forest fires and the WUI characterisation provide both an useful inside and information for understanding predominant fire regimes, underlying processes and for identifying fire vulnerable regions. This comprehensive assessment is fundamental to conduct fire management goals and strategies, and to support policy and decision-making for fire problem mitigation and risk monitoring.

1.2 Objectives

This thesis aims at developing, validating and adapting statistical tools for the detection and the characterisation of spatio-temporal point patterns of environmental data, i.e. forest fire occurrences, and at defining a new wildland-urban interface (WUI) concept, with a particular interest in the Alpine region. We believe that the proposed methodology benefits from a robust contemporaneous statistical and GIS frameworks adaptable to a broad range of environmental/natural/socio-economic point events. A solid theoretical background of traditional/modern statistical methods is developed here based on spatial and spatio-temporal statistics, fractal-based, point processes, machine learning approaches as well as GIS tools.

More general, we aim at proposing a response to the increasing demand for environmental monitoring and management tools for natural and anthropogenic hazard and risk assessment. In this sense, this Thesis intends to fulfil three general goals: 1) contributing to the theoretical adaptation and application of statistical techniques for space-time pattern analysis; 2) providing real-world solutions, particularly, for the case of forest fire occurrences, to both understand their spatio-temporal patterns and to detect clustered areas with a high rate of fire occurrences; and 3) defining, characterising and mapping the WUI and the fire-ignition susceptibility in the Alpine region.

The first objective of this thesis is intended for the monitoring, assessment and modelling of environmental point data. One critical step in environmental management is to conduct both comprehensive analyses and assessment of observed patterns. For instance, understanding and describing the interaction among points that could explain their location in space and/or in time, as well as learning about the underlying processes that may have caused the patterns. To this purpose, we provide a solid background on point processes, fractal, statistical and scan statistics methods and tools for detecting and characterising clusters in spatial, time and space-time data. Thus, this theoretical section comprises the development of global and local clustering measures for point process analysis.

The second objective addresses fundamental questions related to general problems arising when working with real-world applications, particularly, the forest fire phenomenon. In many cases, the application of these theoretical methods is not a straight forward task due to the complexity of both the studied phenomenon and the geographical space where the events take place, e.g. high spatial heterogeneity, high variability of spatial and temporal clustering, complex spatial/temporal constrains, and/or multivariate nature of the phenomenon. To overcome these issues, we develop the concept of validity domain (VD) where we estimated relative clustering measures in data confined in different constrained geographical and time spaces, and perform relevant statistical tests. Furthermore, we show how to apply these clustering measures in the case of forest fire occurrences.

The third and last objective tackles a topical issue that is particularly related to forest fire management which is the definition and characterisation of the WUI. This concept is presented as sensitive areas for human-environment conflicts where many vexing environmental problems take place. One of these issues is the forest fire occurrence from which the majority of recent scientific research and discussion have concentrated. Today's fires are mainly caused by the permanent conflicts of anthropogenic activities for landscape exploitation; hence, the protection of both forests and human infrastructures inside the WUI becomes a challenging task. In this context, we aim at proposing a robust, systematic, statistical and flexible approach for assessing and mapping the WUI in the Alpine conditions providing a scientific framework for land-use and ecological conservation planning and natural hazard monitoring. Accordingly, we develop a new statistical/GIS-based methodology to automatically mapping the WUI. This Thesis portrays the contribution part of the author in the collaboration project with the WSL Research Unit Community Ecology under the frame of the research project "Wildland-Urban Interface (WUI) and forest fire ignition in Alpine conditions - (WUI-CH)" (See Subsection 1.3.6 Collaborations).

1.3 Contributions

This Thesis makes several major contributions to the spatio-temporal pattern analysis for environmental data and to the application to the forest fire phenomenon for monitoring and risk assessment. The results have been presented and published in different national and international conferences and in 5 indexed journal articles. The following briefly reports the main contributions presented in Chapters 3, 4 and 5, software development (in Appendix B) and the lists of the publications and conference proceedings regarding each topic.

For a detailed presentation of these publications see Appendix A. Publications & Proceedings.

1.3.1 Chapter 3

This chapter proposes the theoretical development of a contemporary, comprehensive and robust methodology for the analysis, characterisation and detection of clusters in spatio-temporal point data by adapting and applying statistical point process methods for purely spatial, purely temporal and spatio-temporal analysis. In this framework, we present statistical approaches going from global measures for spatial clustering characterisation (i.e. defining whether the events are clustered or randomly distributed regardless of their location in space or in time [70, 214]) to local measures for cluster detection and identification (i.e. space-time location of statistical significant clusters).

Furthermore, this chapter also introduces the concept of validity domain (VD) which allows taking into account the complexity of the geographical space where real-world phenomena take place. The development of VDs allowed reducing the dimensionality of the mapping space of the studied events and, together with simulated data generated in these domains, enabled quantifying the relative degree of clustering of the observed patterns without making use of statistical tests to assert for statistical significance. This was possible by comparing deviation of the results of each statistical measure applied to the observed events with the results from the random patterns simulated in different constrained geographical spaces (or VDs). For more details see Section 3.5 in Chapter 3). In the same manner, these concepts were adapted to temporal analysis where two temporal constrained observed windows were defined (see Subsection 3.5.1 in Chapter 3). The introduction of the VD concept demonstrated that we can deal with the phenomena complexities without involving a priori knowledge of their underlying processes.

The contributions of this chapter can be divided into three main cluster analysis as follows:

1. *Space clustering analysis*
   This contribution consisted on the adaptation and application of point process methods for purely spatial analysis. In this framework, we presented statistical, fractal and second-order measures for the global characterisation of the spatial clustering of a point pattern. More precisely, we applied the Morisita index (subsection 3.4.1), the Box-counting fractal dimension (subsection 3.4.2), the multifractal formalism (subsection 3.4.3), and the Ripley's K-function (subsection 3.4.4).

2. *Time clustering analysis*
   The proposed methodology was likewise adapted and extended to temporal analysis. For the time case, we not only applied statistical measures exclusively developed for purely temporal analysis such as the temporal K-function (subsection 3.4.4.2) but we also adapted tools that were developed for purely spatial analysis such as the Morisita index (subsection 3.4.1.1), the Box-counting fractal dimension (subsection 3.4.2.1) and the Rényi generalised dimensions for the multifractal analysis (subsection 3.4.3.2).

3. *Space-time*
   For the spatio-temporal case, we applied a local measure such as the space-time permutation scan statistics (STPSS, subsection 3.4.5). This scan statistical methodology uses a scanning window, which moves across the space and the time, detecting local excess of events in specific areas over a certain period of time. The

resulted potential clusters are evaluated through the Monte Carlo hypothesis testing to infer for their statistical significance. It allowed identifying and mapping hot spots of point event data.

1.3.2 Chapter 4

This chapter entirely comprises the application of the methodology, developed in Chapter 3, on a real-world complex environmental phenomena as it is the case of the forest fire occurrences in Canton of Ticino. In addition, one application of the STPSS method is also developed and showed for the case of the forest fires in Portugal.

The main contributions of this chapter are: the quantification and the characterisation of the space and time structures of the forest fires in Canton of Ticino through the application of global clustering measures for point processes; the detection and identification of spatio-temporal clusters by means of the local clustering measure STPSS in the case of the forest fires in Canton of Ticino; and the evaluation of the capability of this STPSS model to detect real clusters in aggregated data as it is the case of the forest fire database of Portugal.

The resulted work was presented and published as follows:

1. *Space clustering analysis*

    - M. Kanevski, **C. Vega Orozco**, M. Tonini, V. Timonin and M. Conedera. Spatial Analysis of Forest Fires Clustering. Case Study: Canton Ticino, Switzerland. In *Proceedings of Spatial Statistics: Mapping Global Change, Enschede (The Netherlands)*, page P2.31, **2011**.
    - **C. Vega Orozco**, M. Tonini, M. Kanevski and M. Conedera. Point Pattern Analysis of Forest Fire Occurrences in Canton Ticino (Switzerland). In *ICFBR 2011 International Conference on Fire Behaviour and Risk Focus on Wildland Urban Interface, pages 126–127, Sassari (Italy)*, **2011**.
    - **C. Vega Orozco**, M. Kanevski, M. Tonini, J. Golay and M. Conedera. Multifractal Analysis of Forest Fires in Complex Regions. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page 1162, **2012**.
    - J. Golay, **C. Vega Orozco**, M. Tonini and M. Kanevski. Spatial Point Pattern Analysis of Environmental Data Using R. In *Symposium Proceedings, Open Source Geospatial Research & Education Symposium, Yverdon-les-Bains (Switzerland)*, pages 245–252, **2012**.
    - M. Tonini, **C. Vega Orozco**, M. Kanevski and M. Conedera. Spatio-Temporal Patterns of Forest Fires: a Comprehensive Application of the K-Function. In *European Geosciences Union General Assembly, Geophysical Research AbstractsCopernicus Publications, Vienna (Austria)*, volume 15, page 5461, **2013**.

2. *Time clustering analysis*

- **C. Vega Orozco**, M. Kanevski, M. Tonini, J. Golay and M.J. Pereira. Time Fluctuation Analysis of Forest Fire Sequences. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 5518, **2013**.
- M. Tonini, **C. Vega Orozco** and M. Kanevski. Spatio-Temporal Aggregation of Wildfires: from Global Cluster to Local Mapping. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.

3. *Space-time cluster detection*

- **C. Vega Orozco**, M. Kanevski, M. Tonini and M. Conedera. Patterns Mining in Spatial-Temporal Sequences: the Case of Forest Fires in Ticino (Switzerland). In *Proceedings of the International Symposium on Spatial-Temporal Analysis and Data Mining STDM, London (UK)*, **2011**.
- [305] **C. Vega Orozco**, M. Tonini, M. Conedera and M. Kanevski. Cluster Recognition in Spatial-Temporal Sequences: The Case of Forest Fires. *Geoinformatica*, vol. 16(4): 653-673, **2012**. doi:10.1007/s10707-012-0161-z.
- M.G. Pereira, L. Caramelo, **C. Vega Orozco**, R. Costa and M. Tonini. Space-time clustering analysis performance of an aggregate dataset: The case of wildfires in Portugal. *Environmental Modelling & Software*, vol. 72: 239–249, **2015**. doi:10.1016/j.envsoft.2015.05.016.

1.3.3 Chapter 5

This chapter proposes two systematic, statistical and flexible geospatial approaches for assessing and mapping the WUI and the prediction of fire ignition susceptibility in Canton of Ticino. The main contributions of this work are the development of the first methodological and decisional statistical/GIS framework for WUI characterisation in the Alpine context, and the generation of the first fire-ignition susceptibility map of this region by means of the random forests (RF) algorithm. This work was presented and published as follows:

- M. Conedera, M. Tonini, L. Oleggini, B. Pezzatti and **C. Vega Orozco**. Wildland-Urban Interface and Forest Fire Ignition in Alpine Conditions. In *ICFBR 2011 International Conference on Fire Behaviour and Risk Focus on Wildland Urban Interface, Sassari (Italy)*, pages 47–48, **2011**.

- M. Leuenberger, M. Kanevski and **C. Vega Orozco**. Forest Fires in a Random Forest. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 3238, **2013**.

- M. Tonini, **C. Vega Orozco** and M. Conedera. A Robust GIS Approach to Map WUI in the Alpine Regions. In *ICFFRMM - International Conference on Forest*

Fire Risk Modelling and Mapping: Vulnerability to Forest Fire at Wildland Urban Interface, Aix-en-Provence (France), **2013**.

- **C. Vega Orozco**, M. Leuenberger, M. Tonini and M. Kanevski. Anthropogenic Forest Fires Susceptibility Mapping using Random Forest Algorithm. In *ICF-FRMM - International Conference on Forest Fire Risk Modelling and Mapping: Vulnerability to Forest Fire at Wildland Urban Interface, Aix-en-Provence (France)*, **2013**.

- M. Leuenberger, **C. Vega Orozco**, M. Tonini and M. Kanevski. Random Forest for Susceptibility Mapping of Natural Hazards. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.

- **C. Vega Orozco**, M. Tonini and M. Kanevski. Analysis of the Dynamic of Urban Areas and of their Interaction with Forest Fires. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.

- [68] M. Conedera, M. Tonini, L. Oleggini, **C. Vega Orozco**, M. Leuenberger and G.B. Pezzatti. Geospatial Approach for Defining the Wildland-Urban Interface in the Alpine Environment. *Computers, Environment and Urban Systems*, vol. 52: 10–20, **2015**. doi:10.1016/j.compenvurbsys.2015.02.003.

1.3.4 Scripts development in open source software

As a by-product of this Thesis, we developed some functions in the R statistical software, an open source software environment for statistical computing and graphics [230]. This free environment allows data manipulation, calculation and graphical display. The R base can be extended via packages available through the Comprehensive R Archive Network (CRAN) which covers a wide range of modern statistics, or through the implementation of customised functions as we have done along this work. The following functions were developed and customised by the author of this Thesis for the application and the adaptation of four clustering measures for both space and time analysis:

- The Morisita index.
- The Box-counting fractal dimension method.
- The Rényi generalised dimensions (multifractality).
- The multifractal singularity spectrum (multifractality).
- The multiquadrats counting.
- The envelope function.

The main motivation and purpose of this software development is the usefulness that these customised functions can provide for the scientific and academic community. The fact that these functions are developed in a open source software allows easily sharing them with the people that could be interested in teaching or in carrying out clustering analysis of point data. The developed R functions are presented in detail in the Appendix B. Software.

1.3.5 Other contribution studies

Other contribution work of this Thesis is the application of the described methodology, adapted and tested for forest fire occurrences, in other phenomena. These studies were carried out in collaboration with other researches, and they are presented as follows:

1. *The Morisita index*
   This work resulted in the proposition of two simple methodologies based on the multipoint version of Morisita index (m-Morisita index, developed by our colleague Jean Golay). The first methodology aimed at characterising the degree of clustering of an environmental monitoring network through the m-Morisita index. And the second methodology is proposed for the detection of structures in monitored phenomena by means of the extension of the m-Morisita index to functional clustering measures. The application cases were the Indoor Radon Monitoring Network in Switzerland and the Swiss population distribution. These studies were presented and published in:

   - J. Golay, M. Kanevski, **C. Vega Orozco**. Multipoint-Morisita Index for the Analysis of Spatial Patterns. In *Proceedings of geoENV2012, IX Conference on Geostatistics for Environmental Applications, Valencia (Spain)*, pages 129–130, **2012**.
   - J. Golay, M. Kanevski and **C. Vega Orozco**. The Multipoint Morisita Index for the Analysis of Geodemographic Data. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.
   - [120] J. Golay, M. Kanevski, **C. Vega Orozco** and M. Leuenberger. The Multipoint Morisita Index for the Analysis of Spatial Patterns. *Physica A: Statistical Mechanics and its Applications*, vol. 406: 191–202, **2014**. doi:10.1016/j.physa.2014.03.063.

2. *Fractal and multifractal analysis*
   This work portrays a fractal and multifractal analysis of the Swiss population distribution. This investigation enabled characterising the spatial patterns and degree of clustering of the population distribution in the three Swiss geographical regions (Alps, Plateau and Jura) and the entire country. Additionally, it also allowed quantifying the dissimilarities between the four patterns. This work is the first Swiss geodemographic study applying multifractal methods using high resolution data. This study was presented and published in:

- **C. Vega Orozco**, J. Golay and M. Kanevski. Multifractal Portrayal of the Distribution of the Swiss Population. In *Proceedings of Spatial Analysis and GEOmatics, Liège (Belgium)*, pages 392–407, **2012**.
- [306] **C. Vega Orozco**, J. Golay and M. Kanevski. Multifractal Portrayal of the Swiss Population. *Cybergeo: European Journal of Geography [Online], Section: Systems, Modelling, Geostatistics*, vol. 714, **2015**. doi:10.4000/cybergeo.26829, url:http://cybergeo.revues.org/26829.

3. *Ripley's K-function*
This work was accomplished for the spatial pattern analysis of rock avalanche occurrences in Argentina. The analysis was carried out using the Ripley's K-function to estimate the degree and the extend of the spatial clustering of the mentioned events. Additionally, an application of the cross K-function was also carried out for detecting spatial correlation with other geological events, e.g. earthquakes. This study was presented and published in:

    - I. Penna, M. Tonini, **C. Vega Orozco**, C. Longchamp, M.H. Derron and M. Jaboyedoff. Rock Avalanche Occurrence in the San Juan Province (Argentina): an Analysis of Their Spatial Distribution and Main Forcing Factors. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 10515, **2013**.

1.3.6 Collaborations

Two collaborations with two institutions were founded during this Thesis:

One collaboration was established with the Research Unit Community Ecology from the WSL under the frame of the research project "Wildland-Urban Interface (WUI) and forest fire ignition in Alpine conditions - (WUI-CH)", co-financed by the Swiss Federal Office for the Environment and the WSL research program *Forest and climate change - Phase I*. This project intended establishing a methodology for defining and mapping the WUI in the Swiss Alpine region, as well as on the reconstruction of the dynamic of the WUI in Canton of Ticino in the last 30 years by means of remote sensing image classification. The work resulted from this collaboration was presented and published in:

- The work presented in subsection 1.3.3 for WUI characterisation.
- R. Ceré, M. Conedera, G. Matasci, M. Kanevski, M. Tonini, **C. Vega Orozco** and M. Volpi. Wildland-Urban Interface Evolution Mapping Using Multi-Temporal Landsat Imagery. The Case of Forest Fires in Southern Swiss Alps. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page , **2012**.

- M. Kanevski, A. Champendal, **C. Vega Orozco**, M. Tonini and M. Conedera. A New Concept of Wildland-Urban Interface Based on the City Clustering Algorithm. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page 2092, **2012**.

- A. Champendal, **C. Vega Orozco**, R. Ceré, M. Kanevski and M. Tonini. A Geomatic Approach to Wildland-Urban Interface Detection. In *Proceedings of Spatial Analysis and GEOmatics, Liège (Belgium)*, pages 73–76, **2012**.

The second collaboration was carried out with the Centre for Research and Technology of Agro-Environment and Biological Sciences (CITAB) of the University of Trás-os-Montes and Alto Douro in Portugal. The work resulted from this collaboration was presented and published in:

1. The article [214] (mentioned in subsection 1.3.2 space-time analysis in forest fires in Portugal).

2. R. Costa, M.J. Pereira, L. Caramelo, **C. Vega Orozco** and M. Kanevski. Spatio-Temporal Clustering of Wildfires in Portugal. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page 12208, **2012**.

3. R. Costa, M.G. Pereira, L. Caramelo, **C. Vega Orozco** and M. Kanevski. Assessing SaTScan Ability to Detect Space-Time Clusters in Wildfires. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 14055, **2013**.

4. M. Pereira, R. Costa, M. Tonini, **C. Vega Orozco** and J. Parente. Influence of the input database in detecting fire space-time clusters. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, **2015**.

5. M. Tonini, M. Pereira, **C. Vega Orozco** and J. Parente. Multivariate cluster analysis of forest fire events in Portugal. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, **2015**.

1.4 State of art

When working with spatio-temporal events, fundamental questions concerning the analysis and modelling of their patterns arise such as whether there exist clusters; what is the overall pattern; how is the relationship between the events; what is the trend; where are the clusters located, etc. However, most of these issues can not be addressed by just looking at a simple graphical representation of the objects. The application of methods of spatial statistics are thereby indispensable in many physical and socio-economic disciplines to provide better information of the structures and the underlying factors of processes that are not perceptible to the bare eye.

Spatial statistics concerns a variety of statistical techniques to quantitatively analyse space-time data. Coupled with GIS tools, these techniques can be used for mapping, data quality assessment, sampling design optimisation and modelling of space and spatio-temporal data structures. These methodologies concern the location, area, distance and interaction between the study objects, and are based on the first law of geography (Tobler [285]): "everything is related to everything else, but near things are more related than distant things". However, the definition of "near" and "distant" may depend on the particular context of the phenomenon of interest for which the observed events need to be referenced in space and/or in time. The literature of spatial statistics is well established and extensive, going from theoretical to applied texts. The theoretical development provides techniques for point process analysis [14, 16, 55, 56, 72, 74, 74, 79–81, 89, 93, 108, 112, 143, 153, 192, 239–241, 267], and spatial interpolation, known as Geostatistics [23, 73, 74, 124, 144, 150, 151, 183–185]. Recently, several good books have been published giving an important overview and state-of-the-art description of spatial analysis in different domains of application, including Cello et al. [43], Diggle [89, 91], Finkenstädt et al. [100], Gelfand et al. [112], Glaz et al. [119], Illian et al. [143], Kanevski [150], Kanevski and Maignan [151], Ripley [243], Seuront [257] and Wiegand and Moloney [316].

In general terms, spatial cluster analysis, as a part of the spatial statistics, can be defined as a family of algorithms aiming at grouping objects showing a local over-density in space and/or in time. Here, the term clustering concerns only the spatial repartition of the events (points) and it is defined as the spatial non-homogeneity of the point distribution in the geographical space in which they are embedded in [290]. Thus, a spatial cluster is a region where the density of events is higher than expected in the surrounding area. The probability that a cluster is a real cluster or that it has occurred by chance is carried out through a comparison with random distributions (e.g. Poisson model) which allows rejecting the hypothesis of independence between the events via statistical tests (e.g. Monte Carlo test).

These algorithms can be classified into two main groups: global and local methodological approaches. The first group seeks to detect clustering by measuring the intensity of the point pattern and defining whether the events are clustered or randomly distributed, regardless of their specific locations in space or in time [70, 256, 305]. Its characterisation can be assessed by an ample number of measures [74, 91, 143, 151] and they can be classified as [120]:

- Topological measures, which evaluate the level of clustering based on the geometry of the object. Among these indices we find the Voronoi polygons and the Delaunay triangulation [151, 310].

- Statistical measures, discovering the presence of clustering by evaluating the distribution of the events: the Morisita index [78, 120, 142, 198], Ripley's K-function [241], the Moran's Index [194] and the variance-to-mean ratio [142], among others.

- Fractal measures, which evaluate how a pattern fills the space where it is embedded in and how this filling varies at different scales detecting the presence of clustering. Some of these methods are the Box-counting method [171, 258, 290], the Sandbox-counting method [99, 127, 273, 290], the Lacunarity index [5, 46, 178, 223], the Information dimension [137, 257] and the Rényi generalised dimensions [29, 109, 126, 137, 209, 257].

The second group of algorithms of spatial cluster analysis, the so called local methods, intends to identify the space/time location and extent of the detected clusters. Some examples are the geographical analysis machine (GAM) [207], the Turnbull's cluster evaluation permutation procedure (CEPP) [297], the Besag-Newell clustering test [26], the Scan Statistic [156] and the density-based algorithms such as DBSCAN [97, 251] and further developments of polygon-based clustering [147, 312]. Nevertheless, analysis from global methods should precede the use of local approaches.

Spatial clustering analyses of environmental events comprises a broad and growing literature in different application fields, for instance, landslides [37, 174, 201, 295, 317] ecology [89, 91, 107, 166, 245, 257, 316], geography [28, 125], forestry [266], forest fires [6, 22, 27, 36, 189, 211, 215, 274–277, 279, 295, 296], pollution [150, 151, 167, 168], crop yield [322], natural hazards [43], landscape [39], urban geography [1, 10, 11, 19–21, 45, 85, 103, 104, 148, 208, 250, 271] among others.

1.4.1 Space-time pattern characterisation of forest fire occurrences

From a statistical point of view, forest fires can be treated as a stochastic marked point process where events are characterised by points and marks, with points describing the locations of the events in space and/or time [291] and the marks providing additional information such as burnt area, ignition-cause, duration, altitude and slope, etc. As many natural phenomena, forest fires are found to be clustered in space and/or in time rather than being randomly distributed. This has been confirmed in several studies. For instance, Moreno et al. [195], Podur et al. [224], Rorig and Ferguson [247], Telesca et al. [278], Tuia et al. [291], Vázquez and Moreno [301, 302] and Yang et al. [319] carried out spatial clustering analysis by applying statistical, topological and fractal techniques to study the spatial distribution of the events without specifying the exact spatial location of clusters. The spatial variability of forest fire long-history has been studied by Ali et al. [4] and Carcaillet et al. [41] using Ripley's K-function which also allowed determining the extent of the role of large- (e.g. climate) and local-scale processes (e.g. relief, slope aspect, human history, etc.) on fire pattern. In the case of temporal clustering analysis, one can find the work of Corral et al. [69], Ghermandi et al. [115], Lasaponara et al. [164], Telesca and Lasaponara [274], Telesca and Pereira [275], Telesca et al. [276, 279]. Most of this work concerns the application of time-fractal approaches to characterise the temporal distribution of forest fire sequences estimating the degree of their temporal clustering behaviour.

For spatio-temporal analysis, while some studies investigate the clustering behaviour such as Genton et al. [114], Hering et al. [139], Telesca et al. [277] and Tuia et al. [292], just few work concerns the detection and identification of statistically significant clusters e.g. Pereira et al. [214], Tonini et al. [286], Tuia et al. [293] and Vega Orozco et al. [305], which applied the space-time permutation scan statistic model [160] for the detection of high rate occurrences areas.

Much attention has also been concentrated in the analysis of power-law behaviour between frequency and size of fire distribution [173, 235, 238]. This behaviour is consistent with self-organised criticality (SOC) systems which is used to understand the dynamic of fires [52, 138, 295] by considering the phenomena as an extended dynamical system operating at states of critical equilibrium [17]. This has led to practical implications to forecast the risk of large fires based on the occurrence frequency of small and medium fires [259].

Works concerning the implication of environmental and socio-economic factors in the distribution of forest fires are found in the paper of Díaz-Avalos et al. [88] and Latham and Williams [165] which revealed the role of topography and fuel characteristics in defining the spatial patterns of lightning-induced fires. Likewise, the works of Cardille et al. [42], Guyette et al. [131], Hessburg et al. [140], Prestemon and Butry [226], Prestemon et al. [227] and Veblen et al. [303] assessed the influence of human factors on modern human-caused fire regimes.

Other researches on forest fires deal with several important and difficult tasks, including the modelling of this phenomenon taking into account physical, meteorological and other processes [212, 225], statistical modelling using cellular automata [152, 320], fire occurrence mapping by means of kernel density approaches [7, 8], fire ignition probability modelling using logistic regression [42, 181, 190, 304], fire hazard probability based on fire frequency [175], and fire risk assessment and mapping through generalised linear mixed models [88, 122], classification tree analysis [7] and GIS-based approaches [38, 51, 121]. In addition, one can find the application of data mining and advanced machine learning algorithms to characterise forest fire regions and patterns [9, 49, 206, 321]. And more recently, stochastic models integrating meteorological dynamic and historical data for forest fires risk mapping and prediction [95].

1.4.2 Characterisation of the wildland-urban interface

Different definitions exist for Wildland-Urban Interface (WUI), but the most widely used is the definition of the US Departments of Agriculture and of the Interior [204] where WUI refers to the geographical areas where human development meet or intermingle with natural vegetation such as grassland, shrub vegetation and forests [59, 130, 231, 232, 281]. Probably the earliest contribution to this topic is the work of Vaux [300] and the book of Bradley [32]. Many environmental-ecological issues are encountered within these areas [3, 32, 221, 300]; however, since the 1990s with the work of Davis [83], the term WUI is almost exclusively used in the context of wildland fires [263]. Problems related to fire hazard and fire management are by far the most WUI-relevant issues [162, 263] as a result of the influence of human settlement in the regimes of anthropogenic-induced fires. In this sense, fire planning and management became of primary actions in these WUI areas [128, 222].

In the scientific domain, investigations mainly involve the development of suitable WUI definitions [221, 232], mapping methods [130, 161, 162], implementation of measures into specific fire management plans [2, 116, 134, 186], validation of the proposed approaches [188], and assessment of fire risk and intervention priority [132]. Most of these works concern the WUI in different fire-prone areas in the world, i.e. United States [18, 54, 188], Autralia [116], Mediterranean Europe [15, 161–163], and South America [86]. Nevertheless, no attempt has been made to investigate and elaborate a definition of the WUI in mountainous environments such as the Alps [68], where general socio-economic/environmental conditions, fuel types and wildfire regimes display different characteristics.

There is no a single standard definition of WUI commonly accepted around the world. As pointed out by Mell et al. [188], Platt [221] and Stewart et al. [263], the assessment of WUI and its related fire risk is mainly an issue of definitions and parametrisations adapted to the local conditions of the study area, the availability of data, the scale of reference, and the management purposes. Nevertheless, all of the existing definitions of WUI agree with the association of three basic elements: 1) the "anthropogenic component" given by human infrastructures (i.e. houses, roads, etc.) which can act as both the ignition sources [15, 42, 162] and the loss during the fire [57]; 2) the "wildland component" given by the vegetation (e.g. forest) different from cultivation (urban green areas, orchards or agricultural activities) [263]; and 3) the "interface area" between the two first components, representing the potential interactions and feedback effects in case of wildfire [68, 221].

Conforming to Davis [83], Haight et al. [132], Platt [221], Radeloff et al. [232], Stewart et al. [263] and Theobald and Romme [281], WUI may be defined according to the spatial organisation of the three mentioned components as: 1) "interface-WUI" when infrastructures are directly adjacent to wildland fuel; 2) "intermix-WUI" when infrastructures are scattered throughout a wildland or forest area; and 3) "occluded-WUI" when infrastructures are completely surrounded by wildland areas [68].

Few methods exist for mapping WUI, and although they are conceived using the same WUI components (human, vegetation and buffer distances), they mostly differ from their operational approaches and implementations [221, 264]. For instance, the definition of the human component may differ mostly due to the availability and resolution of data, e.g. estimates of housing density. The resulting settlement distributions may vary from isolated or scattered housing to dense or very dense agglomerations [161]. For the second WUI component, the vegetation could simply be classified into wildland and non-wildland fuel [68], or it can be divided into different categories, e.g. types, influence on fire ignition and intensity, etc. [221, 281]. The third WUI component is mostly defined by a buffer delineating area which criteria usually varies as a function of the fire characteristics and management purposes, e.g. intended to reduce fire ignition frequency for anthropogenic structure protection or to mitigate fire intensity in order to save fire-fighting conditions [57, 68]. No standard buffer distances exist in literature or in practice to define all WUIs. Currently, buffer distances may be defined using different empirical or theoretical approaches; for instance, by considering, empirically, an area that roughly corresponds to the range of different fire-fighting methods; or by estimating the distance at which radiant heat ignites homes or at which fire brands can fly [58, 68, 263]; or by applying variable buffer-width techniques to adapt WUI to the fire intensity produced by each vegetation type [68, 280].

1.5 Outline

The methodology of this Thesis can be divided into three parts as presented in Figure 1.1:

FIGURE 1.1: Mindmap of the implemented methodology.

The first part (in orange), corresponds to Chapter 2 and consists on the pre-processing and the exploration of the real study data. In this chapter, we present the state-of-art of the scientific work on the forest fires in Switzerland and we make a briefly history and description of the current situation of this phenomenon in Canton of Ticino.

The second part (in green), concerns Chapters 3 and 4. In Chapter 3 we introduce some basic properties of spatial point processes and the theoretical description of the statistical methods implemented in the methodology. This chapter comprises the development of clustering indicators (global and local methods). The group of global measures estimates the intensity of the point pattern defining whether the events are clustered or randomly distributed regardless of their location in space or in time [26, 70]. For that, we develop: the Morista index [198], the Box-counting method for fractal analysis, the multifractal characterisation by means of the Rényi generalised dimensions and the multifractal singularity spectrums [133, 137, 172, 177], and finally, the Ripley's K-function [239]. Simultaneously, simulated data with known patterns were generated as a benchmark for indicating how to estimate the clustering functions, how they behave in a theoretical context, and how they answer to the initial hypotheses. We introduce the VD concepts and simulate random patterns inside the defined spatial and temporal domains. In what concerns the group of local measures, which intends to detect and locate space-time clusters, we apply the space-time permutation scan statistics (STPSS) method [156, 160].

In Chapter 4, we deal with the adaptation and application of the previous statistical methods for purely spatial, purely temporal and spatio-temporal analysis to the case of forest fire occurrences in Canton of Ticino (Switzerland). One application to the case of the forest fires in Portugal is also presented using the STPSS model.

The third part of this Thesis (in purple) corresponds to Chapter 5 and it is related to the use of GIS and machine learning techniques for mapping WUI and predicting the fire ignition susceptibility in the Swiss Alpine context. In this framework, we implement the Random Forests method [34] (a machine learning algorithm) and GIS tools such as the ModelBuilder[2] of ArcGIS for the WUI definition.

Finally, in chapter 6 we present the most significant conclusions of this research and we state possible future research directions in the field of applied statistical analysis.

[2]ArcGIS Desktop - Graphical interface for creating, editing and managing models (workflows).

Chapter 2

Case study: the forest fire occurrences

2.1 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) [40]. One natural process defining the spatio/temporal dynamics of both forest distribution/composition and forest ecosystems is fires. These events are important ecological factors for forest ecosystems because they regulate species composition and influence plant growth and reproduction. Nevertheless, uncontrolled fires can have remarkable detrimental effects on non-fire-adapted forest ecosystems, landscape (desertification or flooding), environmental pollution and global climate change. Moreover, on humans they can have negative impacts on regional economies, human health and safety. [53, 129, 136, 199, 229, 288, 309].

Fires take place on Earth through the interaction of climate and biological factors (e.g. burnable biomass) [210]; however, since humans use fire, the evolution of such events also depends on the intensity and extent of the anthropogenic intervention on the landscape [31, 71, 228, 299]. In this context, the formation, frequency, intensity and distribution of forest fires are defined and controlled by the coincidence of four basic components (see Figure 2.1) such as: the presence of fuel (e.g. vegetative resources), the fire-conductive environmental conditions (e.g. meteorology), the ignition energy (e.g. lightning or humans) [130, 154, 220, 305] and the topography (slope, aspect) which plays an important role on fire distribution.

When these events are recurrent and consistent in a particular area over a certain period of time, they usually result in a specific fire regime [155, 305]. In this sense, fire regime is a concept that comprises the structure of fires [315] such as fire occurrence (i.e. number of fire outbreaks), burnt area (which crucially depend on landscape and topographic characteristics and the extent of the fire spread), and fire effects [313]. The

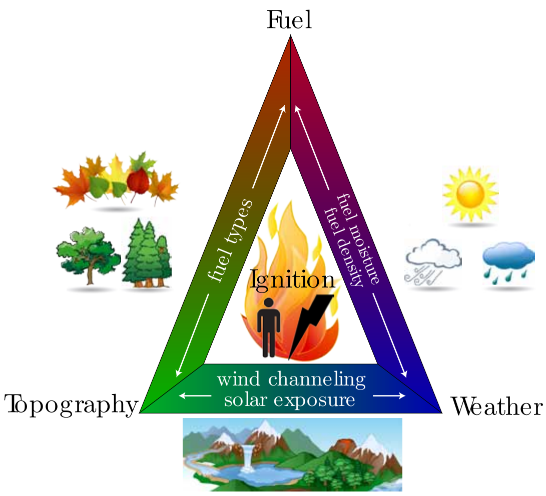

FIGURE 2.1: Components of wildfire ignition and distribution. Source: image modified from *hphanson*[1].

processes and components of fire regimes present consistently high variability in both space and time [42, 98, 197, 220], and are the source of the complexity and the irregularity of the forest fire patterns exhibited over a continuum of scales. Such complexity and scale-dependency of the fire structures are critical issues for both fire modelling and behaviour prediction [76, 260].

Some studies have proved, as the work from Pereira et al. [212], that meteorological conditions are crucial factors for fire development. Based on these factors, several modelling approaches have been developed and used for both forest fire ignitions and patterns predictions (e.g. Cardille et al. [42], Mandallaz and Ye [176]). For instance, a standard method is the calculation of a fire danger index such as the Canadian Fire Weather Index [265]. However, as showed by Viegas et al. [308], a common issue encountered when applying those models over broader temporal and spatial extents is that they are strongly dependent on the specific environmental conditions from where they were initially developed.

Likewise, other variables, such as anthropogenic and ground factors, are also highly relevant for fire ignition and fire distribution. The impacts of human activities (land use and fire management) on fire regimes had been also detected and studied around the world. Bowman et al. [30], Guyette et al. [131], Hessburg et al. [140] and Veblen et al. [303] have revealed that changes in the anthropogenic environment prompt substantial shifts on fire regimes as well; that is, changes on fire frequency, size, intensity and seasonality [155]. Reineking et al. [236] clearly demonstrated that differences in climate, weather, forest composition, and human activities resulted in different patterns of forest fire ignition, suggesting the need of using different model structures for specific situations, which is a non-trivial task.

[1]http://www.hphanson.com/fiction/environment/SERE/wildfire.shtml (consulted on 02.2015).

Moreover, due to the ongoing global changes (i.e. in land use, climate and society) [299], shifts on fire regimes are expected particularly with trends towards an increase of the frequency and severity of forest fires [101, 117, 254, 314]. This obliterated the general perception of the Alpine fires as minor and negligible [299], bringing new issues and challenges for fire management and fire-fighting purposes. Thus, the identification and understanding of both the spatio-temporal patterns of fires and their relationship with the human/environmental components became indispensable. On account of this, the aim of scientists is to provide responses and knowledge of the natural and cultural dynamics of forest fires and their underlying processes for the monitoring and management of forest fire hazard to better integrate fire prevention and mitigation strategies.

2.1.1 Forest fires in Switzerland

Switzerland is generally not subjected to neither frequent nor big forest fires (Figure 2.2); nevertheless, approximately 200 fires are registered every year in its territory, destroying up to 300 ha of forest areas valuable for the socio-economic sectors and for the protection against other natural hazards. Most forest fires in Switzerland occur in the southern part corresponding to the Alpine region. It comprises the Cantons of Valais (VS), Uri (UR), Grisons (GR) and Ticino (TI) [324] (Figure 2.3), and accounts for about 90% of the surface burnt at the national level [196].

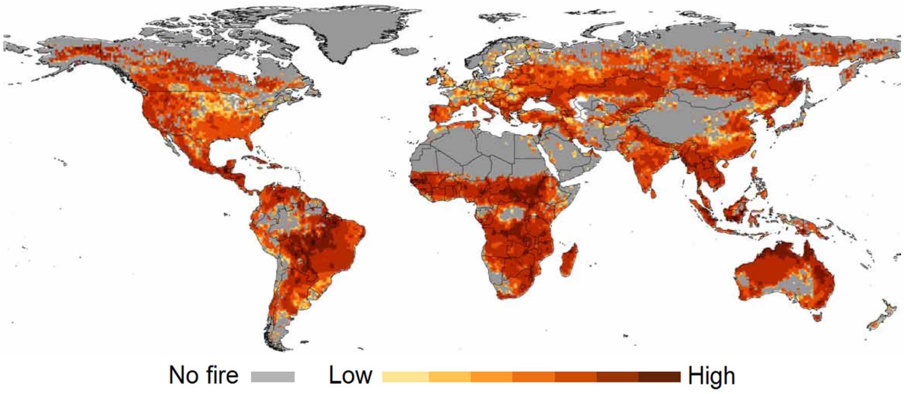

FIGURE 2.2: Worldwide counts of observed wildfire occurrence readings from combined MODIS and ATSR remote sensing products from 1996 to 2007. This figure modified from Moritz et al. [200].

Investigations on forest fires prediction in southern Switzerland started in the beginning of the 90's within the framework of the project *Climate change and Natural Disasters*, financed by the Swiss National Foundation [64, 180]. This project has led to the production of several scientific work studying fire regimes [63, 217, 313, 324], fire-driving factors [62, 325, 326], fire danger [84] and fire impact on forest ecosystems [196, 218, 282]. Nevertheless, a comprehensive analysis of the overall fire regime in the Swiss Alps is still to be fulfil.

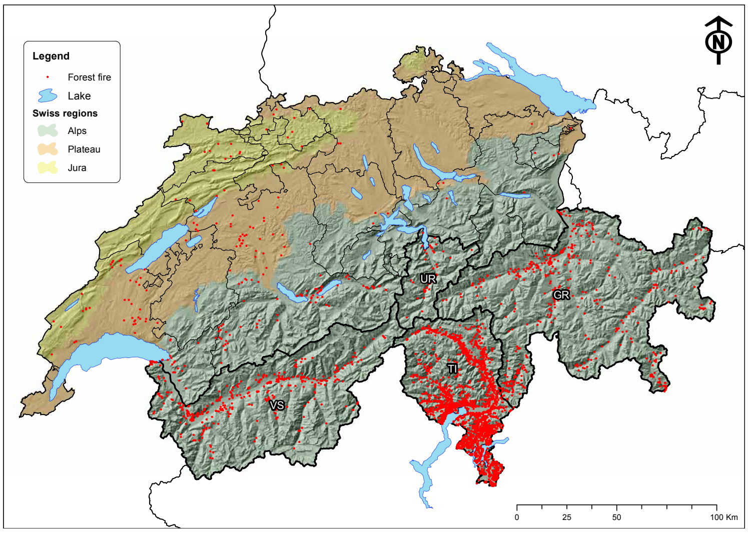

Figure 2.3: Forest fire distribution in Switzerland for the period 1969 to 2008.

Some studies have demonstrated that the history of wildfires in the Alpine region goes as far back as the Lateglacial-Holocene period (around 15,000 cal.yrs BC) [149] with fire regimes characterised by long fire return intervals, e.g. 300–1000 years [299]. At this period, as volcanoes in this region were already inactive, natural fires were only ignited by lightning [284]. However, for centuries, the long-lasting intensive human practices for land exploitation have prevented wildfires to develop in a natural way, and thus confining the knowledge of the natural course of fire regimes [41].

Tinner et al. [282] evidenced the role of fire in changing the composition of woodlands in the southern Switzerland for more than 7000 years. Conedera et al. [66] and Tinner et al. [283] also showed that lightning-ignited fires (also named as natural fires) exert a control on the subalpine coniferous forest ecosystems. Furthermore, Pezzatti et al. [218] analysed the selectivity of fires to forest vegetation classes both in terms of fire frequency and fire size in Canton of Ticino. Their findings suggested that natural fires are more frequently in spruce stands, and tend to be small in size in the beech and mixed forests; while, anthropogenic fires occur mostly at low elevations in chestnut stands, broadleaved forests and in the first 50 m from the forest edge.

On the other hand, the influence of human activities in shaping fire occurrences in the last century is not an exception in the Alpine region. For instance, Zumbrunnen et al. [326] proofed that land use practices contributing to removing biomass (i.e. livestock pasturing and wood harvest) were associated to fire occurrences in Canton of Valais. Furthermore, as suggested by Pezzatti et al. [220], recent short-termed shifts in fire regimes in the Alpine region are also influenced by the implementation of fire legislative and fire-fighting measures which have contributed to reduce fire frequency and size correspondingly.

Some alteration processes encountered in many parts of the world such as fuel build-up, stand density and connectivity of forest areas are also experienced across the entire

Alpine region due to ongoing changes on land use traditional practices, forest management strategies and the abandonment of former agricultural areas in the rural and marginal regions [113, 299, 326]. As a consequence, this situation might lead to fire management and social issues due to the lack of accessibility to fuels build-up areas, the unevenness of human population distribution in the Alpine region diversifying the structures of the wildland-urban interface [299], and the interaction with other natural disturbances, e.g. avalanches, rock-falls, relevant for the protection of forest areas and human settlements [33, 141].

The canton most affected by forest fires is Ticino with fire events occurring principally during winter season (December to April) when dry conditions are more permanent due to both low precipitation and strong foehn winds. This situation is more intensify by the accumulation of dry litter (tree's leaves) on the forest floor characteristic of the previous fall season. As the majority of modern wildfires on Earth, the main source of fire ignition in the canton is the anthropogenic intervention being responsible of more than 90% of the fire outbreaks [63]. Several work revealed an increase in the temperature and summer drought periods in this region in the last three decades (detected at the MeteoSwiss meteorological stations in Lugano and Locarno) [113, 234, 237, 252, 254, 323] and combined with landuse and socio-economic changes would make of fires a major hazard and disturbance factor in parts of the Swiss Alps where they were not so important in the past [325]. This is also predicted by Schumacher [253] in a prospective study on the dynamics of forested landscapes in the European Alps.

Despite all this work, a comprehensive knowledge of the relationship between fire regime and their anthropogenic and ground variables is still missing. This is of suitable interest due to substantial socio-economic and land uses changes undergone in this region and which are likely to continue in the future [113, 254, 326]. Thereby, in this thesis we focused the analysis on the characterisation of the forest fire regime structures in Canton of Ticino as number of fires, burnt area, fire seasonality, ignition sources (natural and anthropogenic) in the last 4 decades, and the influence of human and ground conditions.

2.2 Study area: Canton of Ticino

The Canton of Ticino (hereafter Ticino) is located on the southern slopes of the Swiss Alps bordering with Italy (Figure 2.4), with a total area of 2,812 km^2 [205] and a population of 346,539 inhabitants[2]. Topographically, Ticino is characterised by altitudes varying from 197 m.a.s.l around Lake Maggiore (Locarno) to 3,402 m.a.s.l. on the Adula Peak (northern Ticino, see Figure 2.4) and a rather heterogeneous geology, dominated by siliceous rocks originated in connection with the tectonics of the Alps [220, 305]. The northern and central regions consist of steep mountain slopes and coarsely populated valleys, while the south is a hilly region where most of the population is settled (see Figure 2.5).

[2] Ufficio di statistica del cantone Ticino - Popolazione. Consulted on 31.03.2015

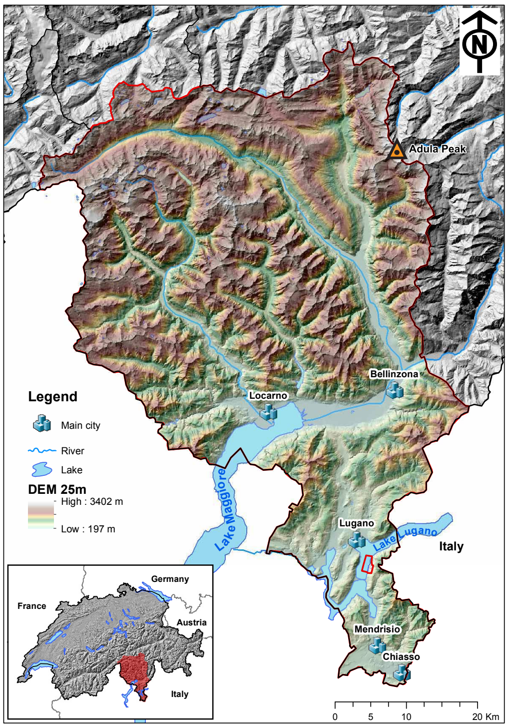

FIGURE 2.4: Geographical location of Canton of Ticino.

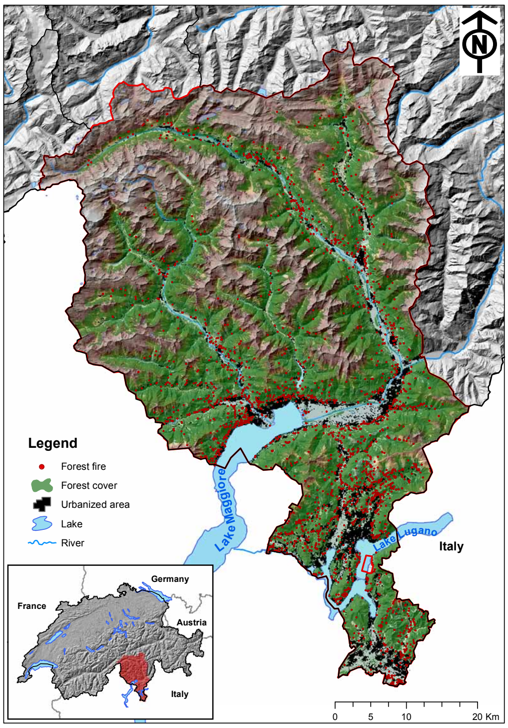

FIGURE 2.5: Forest cover (green areas), human settlement (black areas) and forest fire (red dots) distributions in Canton of Ticino.

Ticino counts with a generally warm-temperate and rainy climate. It presents dry and mild winters with 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 lower as 20% [218]. In the summer season (June–September), precipitation ranges from 800 to 1,200 mm and it is often concentrated in short and heavy spells alternated with long periods without rain or even drought [261]. Although, sunshine duration is high (1,800–2,150 hours/year), some valleys are shaded for several weeks during winter [218]. Depending on the elevation and the geographical location, the mean annual temperature ranges from 3 to 12 °C and the mean annual precipitation ranges from 1,600 to 2,600 mm (Figure 2.6).

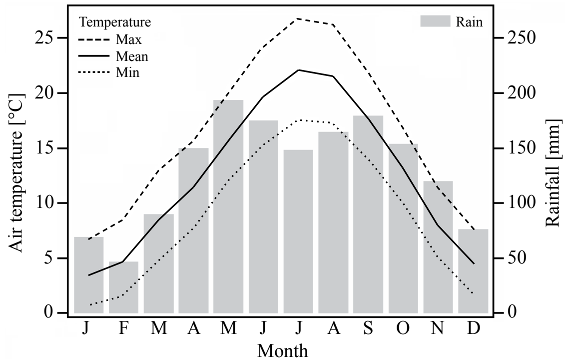

FIGURE 2.6: Monthly mean values of air temperature (left) and rainfall (right) for the meteorological station of Lugano from 1979 to 2008 [220].

The forest cover is proportionally high with about 46.4% of the total territory (Figure 2.5). At low elevations (up to 900–1,100 m.s.a.l.), forest vegetation is dominated by anthropogenic monocultures of chestnut tree (*Castanea Sativa*) occasionally interrupted by the presence of other broadleaved species. At medium elevations (up to 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 forest vegetation consists of *Larix Decidua* [44]. This wild-vegetation cover, as well as fire regimes, have been highly impacted by human activities for several centuries, as showed by Conedera and Tinner [62] who pointed out the effects of the interaction between human activities and forest fires since the Neolithic period in southern Switzerland.

During the 19^{th} century and at the beginning of the 20^{th} century, Ticino was a rural canton with an economy predominantly based on traditional agricultural activities. This condition brought substantial deforestation extended through out the entire territory. This situation drastically changed in the second half of the 20^{th} century, particularly after the World War II, when the canton experienced a marked socio-economic shift towards a more industrial and service-oriented economy system [65, 118, 268]. These changes, certainly, led to a high urban concentration in the low-lands of the main valleys (see Figure 2.7); meanwhile causing the abandonment of many traditional agricultural and forest practices which have favoured natural reforestation, connectivity of forest areas and fuel build up (biomass accumulation) and contributed to the generation of a more fire-prone landscape, in particular for those human infrastructures located in the wildland urban interfaces [161]. As a result, the risk of fire ignition has considerably increased specially in the marginal areas where wild vegetation intermixes with human activities.

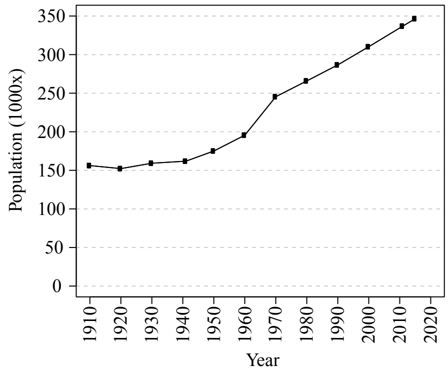

FIGURE 2.7: Evolution of the population in Canton of Ticino in the period 1910–2015. Source:di Statistica del Cantone Ticino and delle finanze e dell'economia [87].

2.3 Data

2.3.1 Forest fire database

The forest fire database used in this Thesis consists on 2,401 forest fire events from Canton of Ticino (Switzerland) for a period of 39 years from January 3rd 1969 to August 8th 2008 (see Figure 2.5). This database holds information about the X and Y coordinates of the ignition point, date and time of the fire outbreak (day-month-year), fire duration, cause of ignition, burnt area, fire type, forest type and topography of the ignition point (altitude and slope).

This data has been systematically collected by the local forest service since 1900. But it was only after 1969 when the dataset was enlarged to include drawings of the burned area on maps with a high accuracy of the location of the ignition point. The existing data has been organised in a relational database including the location of the ignition points and the fire perimeters in a GIS platform by the Swiss Federal Institute for Forest, Snow and Landscape Research WSL [217, 219].

2.3.2 Fire regime in Canton of Ticino

The distribution of the annual number of fires during the whole study period (1969–2008) is rather irregular (Figure 2.8, gray bars), with a general drift towards lower fire frequencies (< 60 events per year) after 1990.

The annual burnt areas (red line in Figure 2.8) are generally of small sizes (~ 590 ha per year) with a tendency toward lower values after 1976. Although, fires in Ticino can

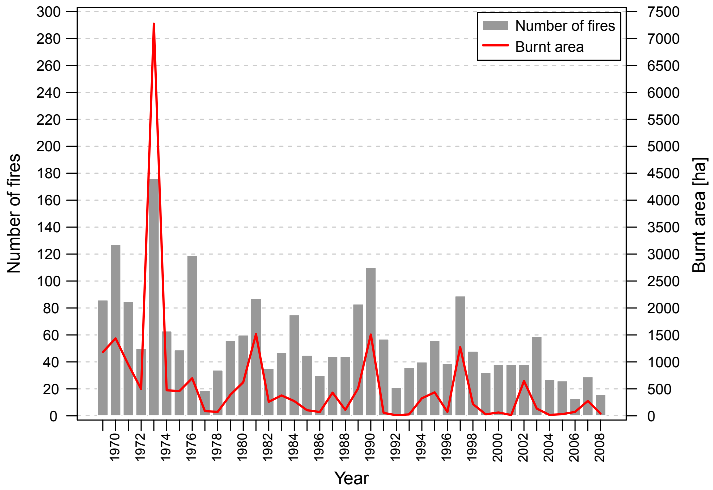

FIGURE 2.8: Annual number of fires and burnt area in Canton of Ticino from the period 1969–2008.

be considered small in size, when fire-prone conditions are present through the year, fires can get out of control resulting in extremely large burnt areas as observed, for instance, in 1970, 1973, 1981, 1990 and 1997. Indeed, the year of 1973 was exceptionally important with 176 fire events that burned 7273.88 ha, despite an average of 1500 ha for the years mentioned above. According to Pezzatti et al. [220], this reduction in anthropogenic forest fire, both frequency and sizes, is mainly the result of the legislative preventing measures and other fire fighting facilities that the canton has been implemented [220], e.g. the prohibition of burning garden debris in the open (Cantonal decree 1987) and the prohibition against fireworks and celebration fires on the Swiss National Day of August 1^{st} in case of high fire ignition danger (Cantonal decrees 1990) [84]. Yet, despite of this shift, fire activity greatly varies from year to year due to the dependence on weather conditions which also endure yearly fluctuations. Some extreme years as the ones mentioned above present burned areas greater than 1000 ha.

The current fire regime is characterised mainly by 3 patterns (Figure 2.9).

One pattern consists of fires caused by anthropogenic sources (hereafter anthropogenic fires) occurring during the fire winter season (December to April, hereafter called winter) with a major peak in March–April characterised by slope-driven surface fires of rapid spreading at low elevations ($<$ 1,000 m.a.s.l.) [218]. The second and the third patterns occur during the vegetation season (May–November, hereafter called summer) and consist of slow spreading anthropogenic and natural fires (ignited by lightning), respectively [66, 84, 236]; both with a peak in the months of July–August.

The monthly distribution of the burnt areas (Figure 2.9 b) presents a major peak in the winter season in the months of March–April due to the dryness generated by the foehn wind and solar radiation, which benefit quick advancing wildfires in this time of the year. In general, burnt areas are larger during the winter season than during summer time, probably because winter fires are principally of rapid spreading while

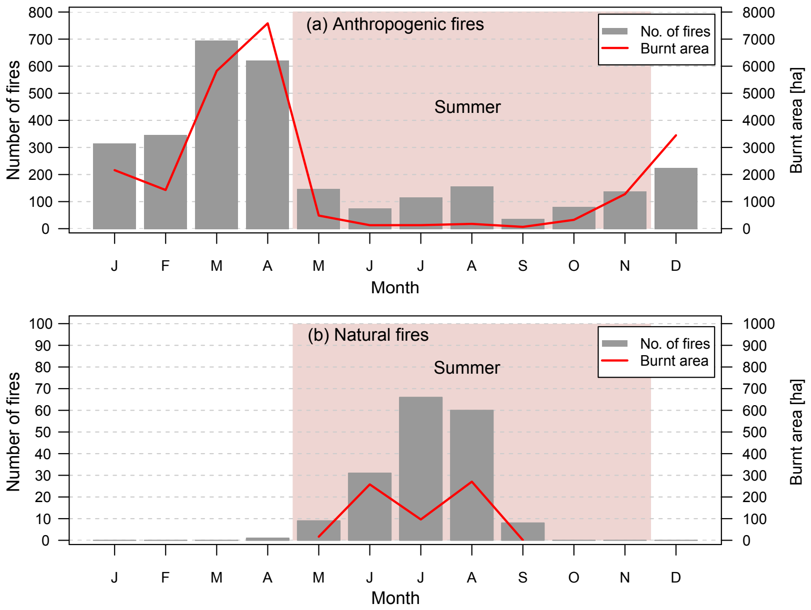

FIGURE 2.9: Monthly forest fire distribution in Canton of Ticino: (a) anthropogenic ignited fires and (b) natural (lightning) ignited fires.

lightning-induced fires are mainly of slow propagation.

Regarding the sources of fire ignition in Ticino, human activities are responsible for the majority of the fire events (92.71%), while only 7.29% corresponded to fire caused by lightning (Figure 2.10). Arson fires (or crime fires) and negligence fires are the origins of the majority of anthropogenic fires as shown in Figure 2.10 (14.20% and 41.69% respectively among human-caused fires). The implementation of preventive measures mitigating the risk of fire-ignition can explain the small presence of other anthropogenic-caused fires such as railways (Rail, 5.04%), army activities (Army, 1.42%) and electrical lines (1.92%) [60].

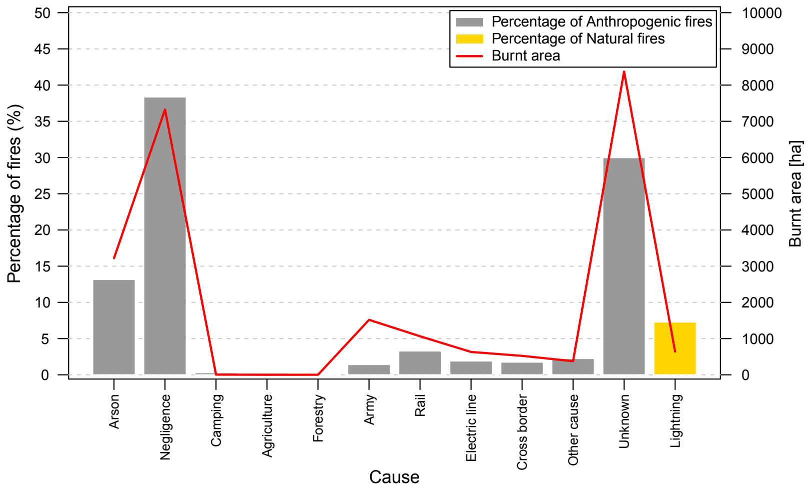

FIGURE 2.10: Percentage of forest fire ignition causes in Canton of Ticino during the period 1969–2008.

Anthropogenic and natural fires also present different geographic distributions (see Figure 2.11). According to Conedera et al. [66], natural fires tend to cluster in the coniferous forests at high elevations in the northern part of Ticino, at steeper slopes and lasting longer than those fires of human sources [84], while anthropogenic fires are distributed at lower altitudes across the entire territory. Because natural fires only occur during the summer season, those fires of unknown causes breaking-out during the winter season can be considered as anthropogenic fires (557 fires, that is, about 77% of the unknown fires - dark green points in Figure 2.11).

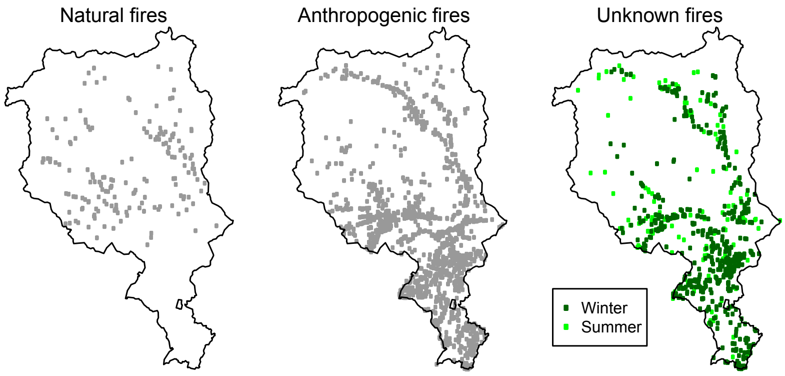

FIGURE 2.11: Geographical distribution of the natural, anthropogenic and unknown-caused forest fires in Canton of Ticino in the period 1969–2008.

Chapter 3

Measures for space-time pattern characterisation

3.1 Introduction

This chapter concerns a short description of the basic notations and fundamental theoretical concepts of the statistical methods relevant in the practice of this Thesis. We explain and illustrate several existing methods for the analysis of spatio-temporal data presented in a point set format such as $(x_i, t_i) : i = 1, ..., n$, where x_i denotes the spatial location of an event i and t_i the corresponding time of occurrence [90, 91]. The analysis of the patterns formed by those events can be described by a *point process*, a stochastic process which generates a countable set of points x_i in a topological space D (a Hausdorff space) [74, 90, 91, 143, 145, 193]. Thus, this approach aims at both analysing and describing the geometrical structure and the correlation of the patterns formed by events [143]. Initially, point pattern analysis only provided an assessment of whether an event distribution was completely spatially random, with implications on the underlying processes that may have caused the pattern. However, more recently developments allow considering bivariate, multivariate and marked patterns, and studying their structure across a range of scales [316].

Section 3.2 provides a short overview of the point process statistics. Section 3.3 describes 3 simulated point patterns, each of them presenting a different spatial structure, used as illustrative demonstration for the theoretical development of the spatial clustering methods implemented in this Thesis. Section 3.4 gives a brief theoretical description of the adaptation and implementation of the methods used in this Thesis (the Morisita index, the Box-counting method, the multifractal formalism and the Ripley's K-function) for purely spatial analysis illustrated with the 3 simulated point patterns of section 3.3; the adaptation and implementation of each of these spatial methods for purely temporal analysis; and lastly, the theoretical development of the space-time permutation scan statistics model for the spatio-temporal analysis. Finally, Section 3.5 introduces the concept of validity domain (VD), defines three spatial and two temporal validity domains for the case of forest fires in Canton of Ticino and describes the complete spatial random (CSR) distributions simulated within each specific validity domains.

3.2 Point process

A *point process* is a stochastic process that generates a countable set of points x_i in a topological space D (a Hausdorff space) [74, 90, 91, 143, 145, 193]. Each point represents the time and/or the spatial location of an event. The term "event" is used here to distinguish the location of an observation from any other arbitrary location within the study area [89, 110, 111]. These events are then described by *points* and also by additional information characterising the events further which are called *marks*. When marks are attached to the points, the process is called a *marked point process*. If only two types of points (events) are considered, the process is said to be *bivariate*, otherwise it is *multivariate* [143].

Mathematically, we can express a point process in terms of the number of events $N(D)$, characterising a particular phenomenon, occurring in a region D [74, 91, 193]:

$$N(D) = \#(x_i \in D) \tag{3.1}$$

Thus a dataset of this type (in statistics considered as a realisation) in a 2-dimensional space is called *spatial point pattern* while the underlying stochastic model for this data is called *spatial point process*. Likewise, these terms are also extendible for both purely time and spatio-temporal data being called time point pattern and spatio-temporal point pattern respectively.

In the temporal case, a point process is formed by a time-ordered sequence of events $t_i : i = 1, ..., n$, where t_i corresponds to the time occurrence of event i and n is the number of events that fall within the time-interval $[0, T]$. Similarly, in the spatio-temporal case, the most basic form of a point process data consists of a time-ordered sequence of events $(x_i, t_i) : i = 1, ..., n$ where x_i denotes the spatial location of event i, t_i denotes the corresponding time of occurrence of event i and n is the number of events that fall within a spatial region D and a time-interval $[0, T]$ [92]. An important assumption is that all events in the dataset are a complete record of the studied phenomenon occurring within the specific region D and the time-interval $[0, T]$ [90].

The aim of point process statistics is to describe and to analyse the geometrical structure of patterns formed by events distributed randomly in one-, two- or three-dimensional space [143]. With the point process is then possible to understand the interaction among the points while also providing information on the underlying processes that can explain the position of the events. This implies the analysis of the degree of clustering (attraction) or repulsion (inhibition) among the points, the space-time scales at which this behaviour occurs and explaining the position of the events in the study area [143].

The characterisation of the entire process may be an impossible tasks, however, we can characterise some properties representing important aspects of the behaviour of the process such as the so-called *first-order* and *second-order* properties. The first-order property describes the way in which the expected value (mean or average) of the process varies across the space (trends), and it is formulated in terms of the *intensity function*, $\lambda(x)$, of the process, which is the mean number of events per unit area at the point x [91]:

$$\lambda(x) = \lim_{|dx|\to 0} \frac{E[N(dx)]}{| dx |} \tag{3.2}$$

For a stationary process, $\lambda(x)$ is summarised by a single value λ, which is assumed to be a constant value representing the mean number of events per unit area [91, 143].

The second-order property, or the spatial dependence, describes the covariance (or correlation) between values of the process at different regions in space (relationship between events in pairs of sub-regions) [111]. This is formally defined in terms of a limit, the *second-order intensity* of the process [91, 111]

$$\lambda_2(x, y) = \lim_{|dx|,|dy|\to 0} \frac{E[N(dx)N(dy)]}{| dx || dy |} \tag{3.3}$$

For a stationary point process, $\lambda_2(x, y) \equiv \lambda_2(x - y) = \lambda_2(r)$. This implies that the second-order intensity depends only on the vector difference, $r = \| x - y \|$ and not on their absolute locations (the process is invariant under translations in $\mathbb{R}^2$). The process is further said to be *isotropic* if such dependence is a function only of the distance of this vector r and not of its orientation (the process is invariant under rotations about the origin in $\mathbb{R}^2$) [91, 111, 143, 193].

Here, we just provide the most fundamental part of the theory of point process relevant for this Thesis. For more detailed theory refer to Cox and Isham [72], Cressie [74], Daley and Vere-Jones [79, 80], Illian et al. [143], Møller and Waagepetersen [192, 193], Stoyan et al. [267] and Diggle [91].

3.3 Simulated data for theoretical demonstration

In spatial point process, the dependence between points (interaction) generally suggests three fundamental patterns in which distributions may be classified: randomness (independence), regularity (inhibition) and clustering (attraction).

A random realisation, also known as *complete spatial randomness* (henceforth CSR), implies no interactions among points, i.e., the probability that an event can occur at any point is equally likely to occur anywhere within a bounded region D and that its position is independent of each any other event. This property provides the standard baseline against which spatial point patterns are often compared and it is used as a dividing hypothesis to distinguish between "regular" and "clustered" patterns [74, 91, 112, 145]. In a regular distribution, events are more evenly spaced than would be expected under CSR. On the contrary, in a clustered distribution, points tend to be closer than would be expected under CSR.

The simplest theoretical model for a spatial point pattern is that of CSR. One important question is whether the observed events display any systematic spatial pattern: clustering, randomness or regularity. Other interesting questions also arise, for instance, is the observed clustering due mainly to natural background variation in the population from which events arise? Over what spatial scale any clustering occurs? are clusters merely a result of some obvious a priori heterogeneity in the region studied? are they associated with proximity to other specific features of interest? are events clustered in space also clustered in time?, etc. [111].

The most common, simple and useful CSR process is the Poisson distribution. The simulation of this process is necessary, in particular, for the calculation of intervals of confidence which helps testing the null hypothesis of randomness. Although, in practice, this idealise standard process is rarely attainable, most analyses in point patterns consider testing for departures from CSR process. This is mainly due to the fact that phenomena presenting a CSR distribution do not merit further analysis because the independence between the location of points makes their predictability impossible [91].

Thus, a Poisson process is characterised by two properties: (i) the number of events in any region D with area $\mid D \mid$ follows a Poisson distribution with mean $\lambda \mid D \mid$; (ii) given n events x_i in a region D, the x_i are an independent random sample from the uniform distribution on D. For a deeper theoretical development of this process, refer to Illian et al. [143].

For illustrative purposes, in order to facilitate the understanding of the theoretical development of the methods implemented in this Thesis, three distributions were generated as presented in Figure 3.1, based on the mentioned fundamental point structures. Each realisation was conditioned to have 100 points distributed in a square window of 100 x 100 dimension. The definition of a "window" where the events are observed is fundamental in point pattern analysis because it represents a source of complementary information on where points could potentially occur [193].

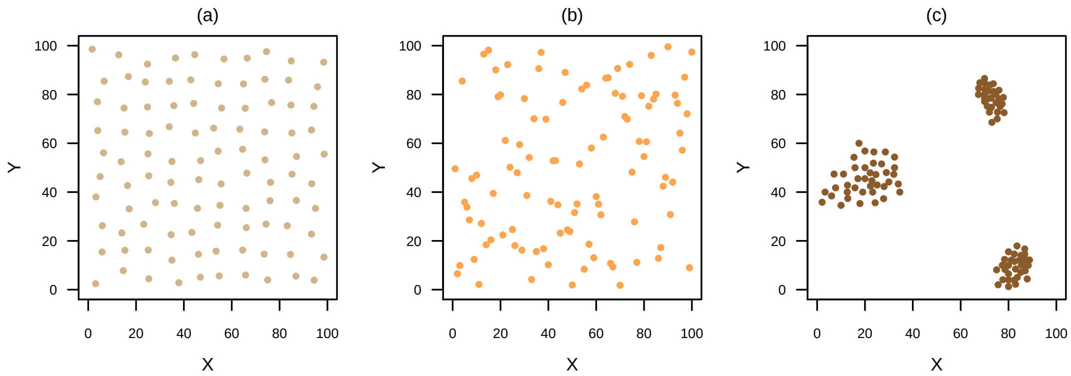

FIGURE 3.1: Three simulated point patterns with known spatial structures: (a) regular, (b) random and (c) clustered. Each pattern is a realisation of 100 points embedded in a bounding box of 100 x 100 dimension. They were generated within the R environment.

Figure 3.1 a shows a regular process over the square box where points tend to avoid each other. For the random distribution in Figure 3.1 b, we consider a homogeneous Poisson process where points are independently and uniformly distributed. And finally, Figure 3.1 c, shows a cluster realisation where points tend to be close together.

We note that for each pattern, we simulated one realisation process where events occur singly, i.e. we did not consider marked point processes. Other important clarification is that along this work, the concept of "clustering" refers to the dependence between points to be aggregated closer than expected for a Poisson process (i.e. spatial clustering); differing from classification concepts where "clustering" implies that the points are organised into identifiable groups (clusters) [12].

3.4 Methods

3.4.1 The Morisita index

The Morisita index of dispersion is a statistical measure used to characterise quantitatively the clustering (non-homogeneity) of a point set [151]. This classical index, I_δ, is obtained by dividing the study region into Q identical boxes of size δ (i.e. length of the diagonal, see Figure 3.2); first starting with a relatively small box size which in turn is increased until a chosen maximum value is reached. The number of events n_i within each box i of size δ is counted and related to the box size. The Morisita index is then computed as [290]:

$$I_\delta = Q\frac{\sum_{i=1}^{Q} n_i(n_i-1)}{N(N-1)} \tag{3.4}$$

where N is the total number of points. It is then possible to draw a plot relating every I_δ to its corresponding δ (see Figure 3.2).

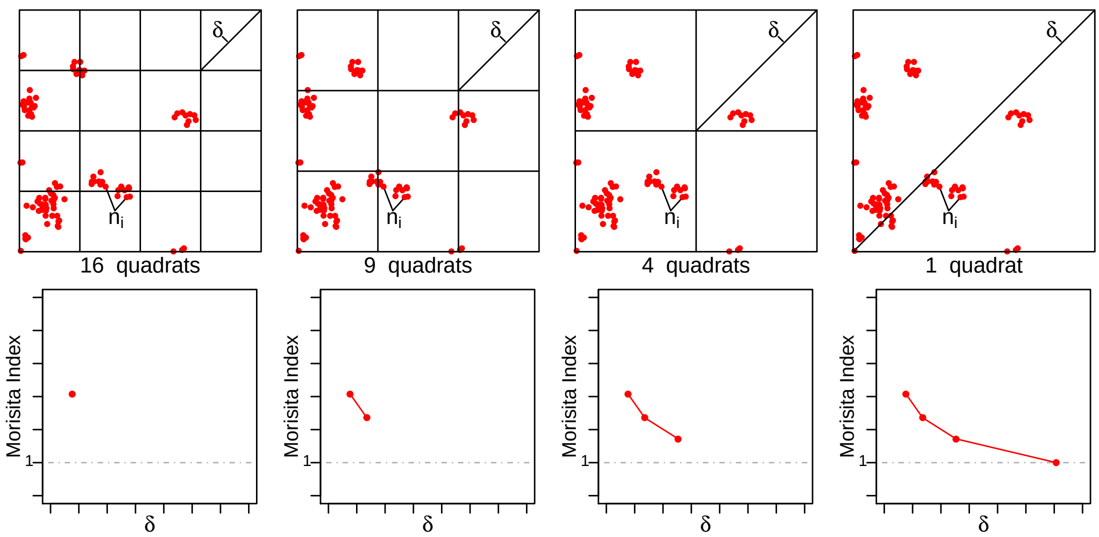

FIGURE 3.2: Graphical computation of the Morisita index.

Properly speaking, the Morisita index measures how many times more likely it is to randomly select two points belonging to the same box than it would be if the points were randomly distributed (i.e. Poisson process) [120]. In other words, this index can be viewed as a measure of the variability between patches assuming that a point process is a mosaic of patches of differing intensities with random spacing within patches [74]. Thus, the behaviour of the Morisita diagram reflects clustering of the event pattern at each scale. Notice that, in Figure 3.2, the boxes partly overlap from one scale to the next (i.e. the number of boxes used for the computation of the index throughout the scales does not follow a geometric series). In real case studies, this is done to give more importance to small scales where a change in box sizes is more likely to capture the characteristics of the point patterns than great changes at large scales (i.e. it is a kind of regularisation) [120].

Figure 3.3 displays the Morisita index estimation for the three simulated patterns.

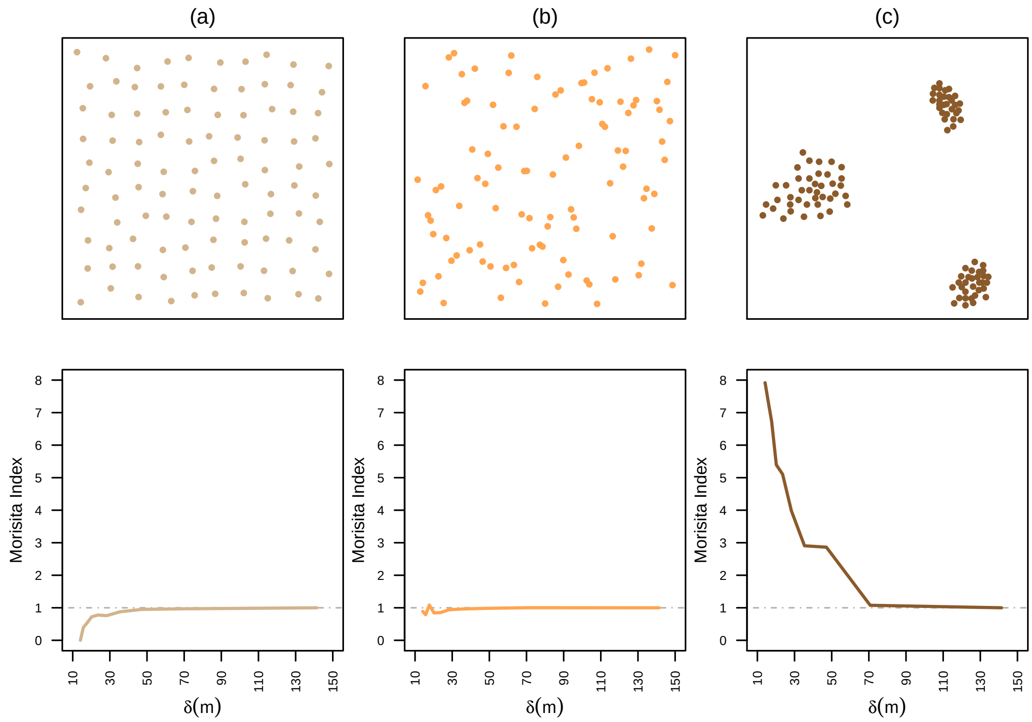

FIGURE 3.3: The Morisita index analysis for the three simulated patterns: (a) regular, (b) CSR and (c) clustered.

The diagram shows that for the random point distribution every I_δ value fluctuates around 1 (Figure 3.3 b). Keep in mind, that the Morisita index is a ratio between two probabilities: that from the analysed pattern and that from a Poisson distribution. If the analysed pattern (in this case the simulated random distribution) is a CSR distribution, then the probabilities to select two points in the same box is equal for both distributions, thus, the index equals 1 at every scale δ. The index also equals 1 when the box size is sufficiently large to cover the entire analysed pattern. This condition is independently to the type of the analysed pattern, due to the fact that the Morisita index does not take into account the location of the points inside the box, instead, it only considers the probabilities of encountering points, which in the case of large boxes is the same as it will be in a CSR distribution.

For the dispersed distribution (Figure 3.3 a), $I_\delta < 1$ and approaches 0 at small scales [120, 151]. In this context, the values of the index indicate that the probability of selecting two points in the boxes is less than expected in a CSR distribution. For small size boxes, the index approaches 0 because only 1 or 0 points fall into the box.

Finally, opposed to the dispersed condition, the clustered distribution shows values of I_δ greater than 1 (Figure 3.3 c). This is given by the boxes with greater probabilities of selecting two points than expected in a CSR distribution. Thus, the great number of empty boxes at small scales increases the value of the index.

3.4.1.1 Adaptation of the Morisita index for temporal analysis

The adaptation of the Morisita index for a temporal analysis of clustering is performed based on the same principle as for the spatial Morisita index (see subsection 3.4.1). The time Morisita index is obtained by dividing the entire study period T into Q_t adjacent time-intervals of equal length t (timescale); first starting with a relatively small timescale, and then, increased until a chosen maximum value. The number of events n_i within each time-intervals i of length t is counted and related to the timescale. The time Morisita index is then computed as:

$$I_t = Q_t \frac{\sum_{i=1}^{Q_t} n_i(n_i - 1)}{N(N-1)} \tag{3.5}$$

where N is the total number of points. It is possible to draw a plot relating every I_t to its corresponding t.

As for the spatial case, the Morisita index reflects the clustering of a temporal pattern with dependence on the timescale. For a randomly distributed pattern, values of the Morisita index fluctuate around 1, while higher values indicate clustering.

3.4.2 The Box-counting fractal dimension

Fractal geometry is a fundamental approach for describing the complex irregularities of the spatial structure of point patterns. Introduced by Mandelbrot [177], the word fractal was first coined to describe objects (or sets of points) with abrupt and tortuous edges. An object called fractal means that it has some sort of self-similarity structure, i.e. a set of points whose any scale portion is statistically identical to the original object. In this manner, fractal measures describe the changes of an object over a variety of scales [94]. This can be characterised by a fractal dimension which refers to the invariance of the probability distributions of the object under geometric changes of scale [246]. Frequently, point processes exhibit a scaling behaviour indicating a high degree of point clustering over all scales [172] and thereby fractal tools can be used to characterise the intensity of their spatial clustering at a wide range of scales [48, 172].

According to Lovejoy et al. [171], Salvadori et al. [249] and Tuia and Kanevski [290], a fractal dimension can be used to analyse the clustering properties (non-homogeneity) across scales of point process realisations. If the studied distribution is embedded within a 2-D space (e.g. geographical space), its fractal dimension can range from 0 (being the topological dimension of a point) to 2 (the topological dimension of a geographical space), with 1 corresponding to the topological dimension of a line. It is called fractal because the space-filling capacity of the pattern (dimension) is characterised by a fractional number, instead of an integer value as the topological dimension. If the points are dispersed or randomly distributed within the 2-D space, the corresponding fractal dimension equals 2. However, this dimension decreases as the level of clustering of the pattern increases, reaching 0 if all the points are superimposed at one single location. Thus, fractal dimensions allow detecting the appearance of clustering as a departure from a dispersed or random situation while indicating the type of space they are filling, that is, if the patterns are near to fill a 0-, 1- or 2-D space.

A variety of fractal measures have been proposed such as the sandbox-counting method [77, 99, 127, 151, 273, 290, 307] and the information dimension [137, 257]. The fractal dimension method used in this Thesis is the Box-counting method [151, 171, 248, 290]. The computation of this method consists on breaking the fractal pattern (point set) into pieces by superimposing a regular grid of boxes of size δ_1 on the region under study (see Figure 3.4). Then, the number of boxes, $N(\delta_1)$, necessary to cover the pattern is counted; that is, one counts all occupied boxes despite the number of points in each box. Next, the linear size of the boxes is reduced, $\delta_2 (< \delta_1)$ and the number of boxes, $N(\delta_2)$, is counted again. The algorithm goes on until a minimum size δ_{min} is reached. For a fractal pattern, the scales (δ) and the number of boxes ($N(\delta)$) follow a power law:

$$N(\delta) \propto \delta^{-df_{box}} \tag{3.6}$$

where df_{box} is the fractal dimension measured with the Box–counting method [290] and it is obtained by calculating the slope of the linear regression fitting the data of the plot which relates $\log(N(\delta))$ to $\log(\delta)$ (see Figure 3.4). Thus, this method quantifies the dimensionality of a pattern by estimating the degree of the pattern spatial coverage of the space [290].

Figure 3.5 shows the Box-counting fractal dimension analysis for the three simulated patterns.

For the random point distribution the Box-counting fractal dimension df_{box} is 1.87 (Figure 3.5 b). Keep in mind that the Box-counting measures the spatial filling capacity of the pattern. Thus, theoretically, for an infinite point pattern, this value is exactly 2. However, the fractal dimension of this CSR distribution is not exactly 2 because the point set consists of a finite number of points which does not perfectly fill up the 2-D space in the marginal areas (edge effects).

For the regular distribution (Figure 3.5 a), the df_{box} is exactly 2 [151], and this is probably due to the fact that, although its finite number of points, the points closer to

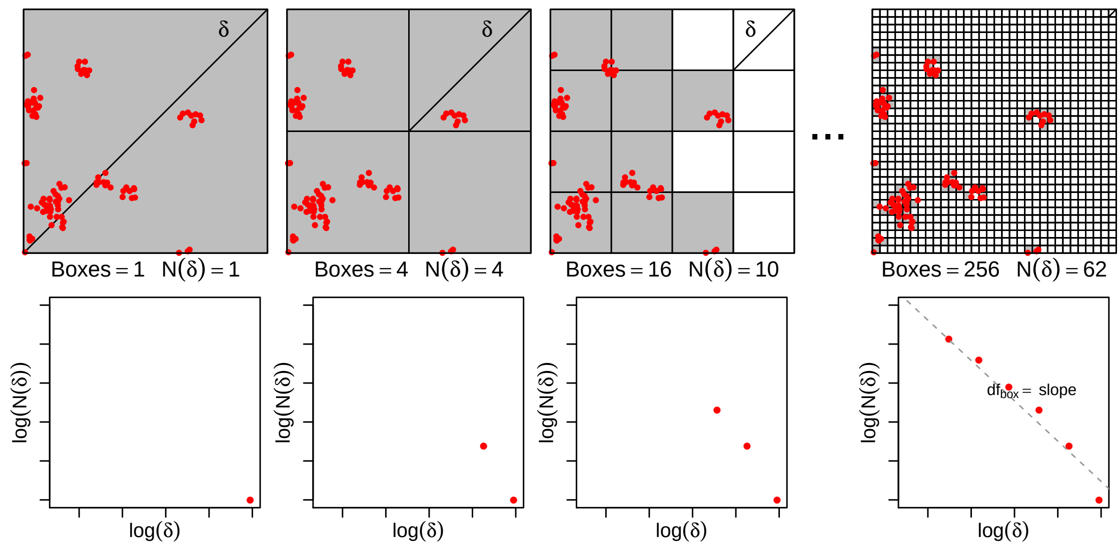

FIGURE 3.4: Graphical computation of the Box-counting method.

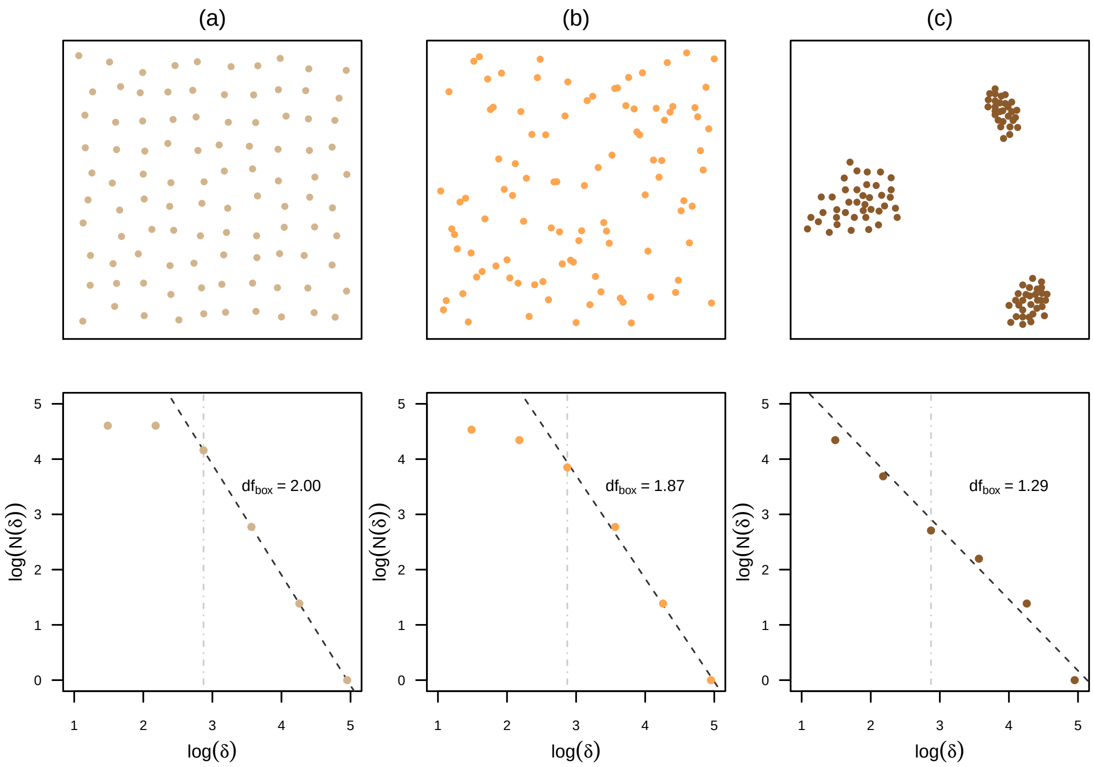

FIGURE 3.5: The Box-counting fractal dimension analysis for the three simulated patterns: (a) regular, (b) CSR and (c) clustered.

the margin areas better follow the bounding box shape reducing the effects of the edges.

For the clustered distribution, df_{box} is 1.29 (Figure 3.5 c) indicating a departure from a homogeneous condition with high concentration of points in few areas of the study region. This value means that the events are highly clustered where points are superimposed filling up a geometrical space of a linear object rather than a 2-D space.

3.4.2.1 Adaptation of the Box-counting method for temporal analysis

The adaptation of the Box-counting fractal dimension method for temporal point sequences is based on the same principle exposed in Subsection 3.4.2 but applied to a 1-D process. For that, we proceed by partitioning the entire studied period T into adjacent time-intervals of equal length t (timescale), and counting only the number of time-intervals ($N(t)$) with at least one event in it. For a temporal fractal pattern, the timescales (t) and the number of intervals ($N(t)$) follow a power law:

$$N(t) \propto t^{-df_{box}} \tag{3.7}$$

where df_{box} is the fractal dimension measured with the Box-counting method and it is obtained by calculating the slope of the linear regression fitting the data of the plot which relates $\log(N(t))$ to $\log(t)$.

In this temporal analysis, the studied point distribution is embedded within 1-D space, that is, a linear space. Thereby, the associated fractal dimension of the temporal pattern must lie between a lower limit of zero (the topological dimension of a point) and an upper limit of unity (the topological dimension of a line) [178]. Thus, the fractal dimension value of a randomly distributed temporal pattern equals 1. This value decreases as the level of clustering increases, reaching 0 if all the points are superimposed at one single time-interval.

3.4.3 Multifractality

The application of multifractal statistics to study a natural phenomenon provides an important window to explore subtle differences in a pattern and to infer about the mechanisms generating the process [47, 179, 294].

As stated by Mandelbrot [179], the notion of self-similarity can be extended to measures (spreading mass or probability) distributed on a Euclidean support (e.g. a point set). In this context, fractal sets might be described by not just one fractal dimension (as exposed in the previous section 3.4.2), but rather by a function [262] or a spectrum of interlinked fractal dimensions. Such fractal sets are said to be multifractal.

Two different approaches are used here to conduct a multifractal analysis: (1) the Rényi generalised dimensions [29, 126, 137, 209, 216, 257, 273] which can be viewed as a global parameter [45] examining how different densities are distributed in the space; and (2) the multifractal singularity spectrum [50, 133, 187, 262], viewed as a local parameter [45] which examines the regularity of the distribution of regions with similar scaling indices. Although, each of these two methods present a different approach, they both describe the same information. They are both based on the Box-counting method, where a regular grid of boxes of size δ is superimposed on the point set and, for both methods, a probability distribution is computed in each box [170]. Figure 3.6 shows the principle of the multifractal analysis based on the Box-counting method.

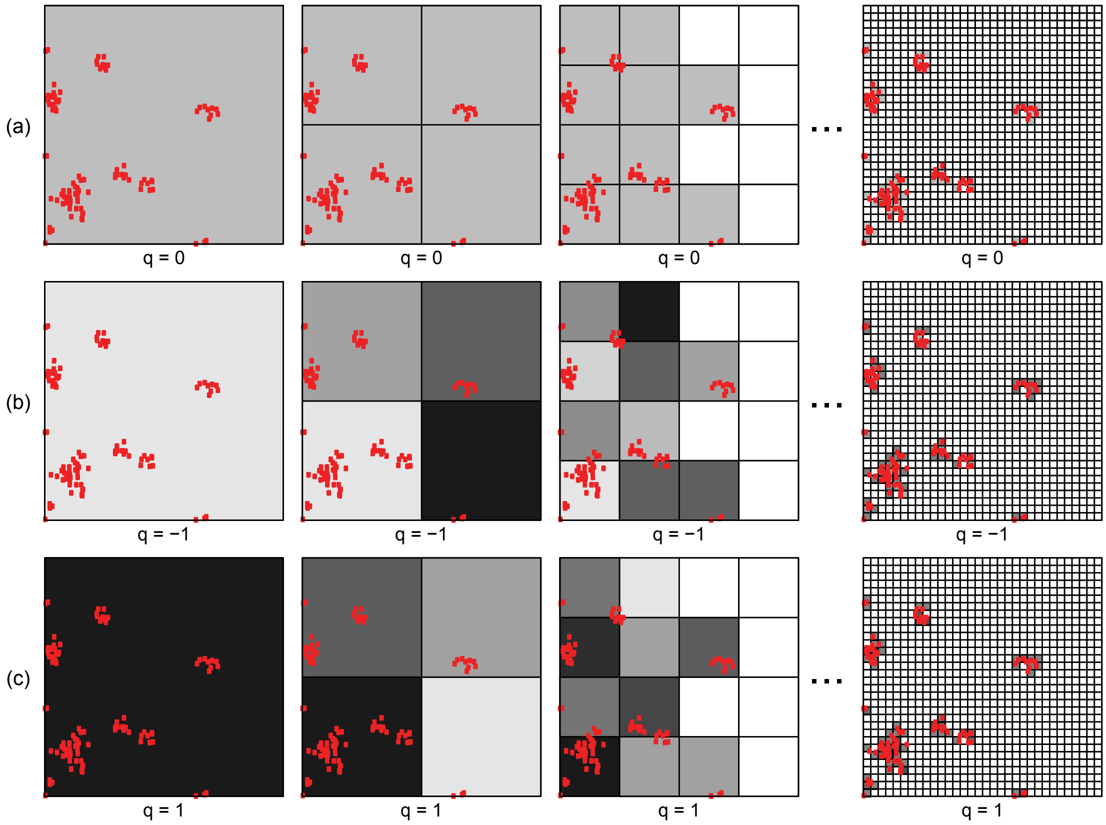

FIGURE 3.6: Graphical principle of the multifractality based the Box-counting method.

In the multifractal analysis, the measure inside each box is taken into account, while for fractal analysis all boxes are equally weighted (= 1) regardless the number of points in the box. In order to obtain a spectrum of fractal dimensions, the pattern distribution is scanned at different order moments q. When analysing $q = 0$ (Figure 3.6 a), the variation of the measure inside each box does not matter, being then equivalent to the Box-counting method. When analysing $q > 0$, the measure inside each box gains more importance as q increases, particularly for those boxes with larger number of events (Figure 3.6 c, where the darker the box the higher the measure). On the contrary, for $q < 0$, the boxes that are almost empty gain more importance (Figure 3.6 b, where the darker the box the lower the measure).

Although, the multifractal approach allows studying the scaling behaviour of sparser and denser regions by scanning the probability distribution at a large range of order moments ($-\infty < q < +\infty$), in this Thesis, we apply a multifractal formalism only on the positive values of q (≥ 0). Our reason is that we are interested in highly clustering behaviour, thus, we only focus on the characterisation of the areas with relative high density values. Moreover, we do not count with information at scales lower than the resolution of the data. Therefore, the theoretical illustration with the three simulated patterns is carried out only for $q \geq 0$.

3.4.3.1 The Rényi generalised dimensions

The spectrum of generalised dimensions, D_q, is estimated by computing the Rényi information, $I_q(\delta)$, of qth order [244]:

$$I_q(\delta) = \frac{1}{(1-q)} \log(\sum_{i=1}^{N(\delta)} p_i(\delta)^q) \tag{3.8}$$

where $p_i(\delta)$ is the probability distribution in the ith box of size δ, $q \in Z$, and the sum is taken for all non-empty boxes. When $q \to 1$, $I_q(\delta)$ is defined as:

$$I_q(\delta) = -\sum_{i=1}^{N(\delta)} p_i(\delta) \log(p_i(\delta)) \tag{3.9}$$

Then, when applied to multifractal sets, $I_q(\delta)$ follows a power law as:

$$exp(I_q(\delta)) \propto \delta^{-D_q} \tag{3.10}$$

And the Rényi generalised dimensions are defined as [126, 137, 209]:

$$D_q = \lim_{\delta \to 0} \frac{I_q(\delta)}{\log(1/\delta)} \tag{3.11}$$

D_q is obtained by the slope of the plot relating $\log(I_q(\delta))$ to $\log(1/\delta)$; then the spectrum of the D_q is plotted against q (Figure 3.7):

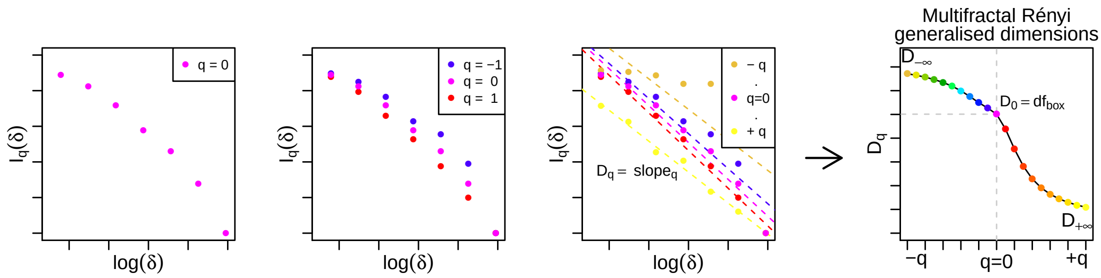

FIGURE 3.7: Graphical computation of the Rényi generalised dimensions.

For monofractal sets, D_q is equal for all q order moments, whereas, for multifractal sets, D_q depends on q and decreases as q increases characterising the variability of the measure ($p_i(\delta)$) [120, 137]. For $q > 0$ (see Figure 3.7 right plot $D_{+\infty}$), the function captures the scaling behaviour of the regions where the events are more clumped. While for $q < 0$ (see Figure 3.7 right plot $D_{-\infty}$), the function describes the empty areas.

Another important indicator of the spectrum is given by the range of the D_q values ($D_{-\infty} - D_{+\infty}$) [272]. This difference is an indicator of the heterogeneity of the densities in the distribution. That is, the variability between sparser and denser areas (boxes). Thus, the wider the spectrum, the more heterogeneous the distributions of the irregularities of the pattern. Moreover, the slope of the D_q spectrum is another important indicator of how much clustered are the events, where the steeper the slope, the more unevenly distributed are the densities and the more clustered are the areas with higher number of events [82].

It is also necessary to point out that particular cases are given within the multifractal analysis, for instance:

1. D_0, that is when $\lim_{q\to 0} D_q$, corresponds to the Box-counting dimension which considers how the number of boxes required to cover a pattern scales with the box size (see Figure 3.7 right plot $D_0 = df_{box}$) [151, 171, 248, 257].
2. D_1, that is when $\lim_{q\to 1} D_q$, corresponds to the information dimension which considers how the probability scales with the box size [257].
3. D_2, that is when $\lim_{q\to 2} D_q$, corresponds to the correlation dimension which considers how the number of distinct pairs of points scales with the box size [127, 257].

Figure 3.8 shows the Rényi generalised dimensions for the three simulated patterns. For the CSR distribution the spectrum of D_q (Figure 3.8 b) is almost constant (values of 2). The reason that the D_q is not exactly equal for all q is because the point set does not perfectly fill up the 2-dimensional space in the marginal areas (edge effects). Thereby, this slightly affects the value of D_q at higher moments. Similarly is the case of the regular distribution (Figure 3.8 a), where each D_q is almost 2. For the clustered distribution, D_q describes a multifractal behaviour (Figure 3.8 c) which designates a departure from a CSR condition and indicating a high concentration of points in few areas of the study region.

3.4.3.2 Adaptation of the Rényi generalised dimensions for temporal analysis

The adaptation of the Rényi generalised dimensions for temporal point processes is based on the same principle explained in Subsection 3.4.3.1 applied to a 1-D process. Thus, we can define the Rényi generalised dimensions for temporal sequences by adapting Equation 3.8 and replacing it in Equation 3.11 as:

$$D_q(t) = \frac{1}{(1-q)} \lim_{t\to 0} \frac{\log(\sum_{i=1}^{N(t)} p_i(t)^q)}{\log(1/t)} \tag{3.12}$$

where $N(t)$ is the number of time-intervals with at least one event, t the timescale, $p_i(t)$ the probability distribution in the ith time-interval of length t and q the order moment. Thus, multifractal spectrum is obtained by plotting $D_q(t)$ against q.

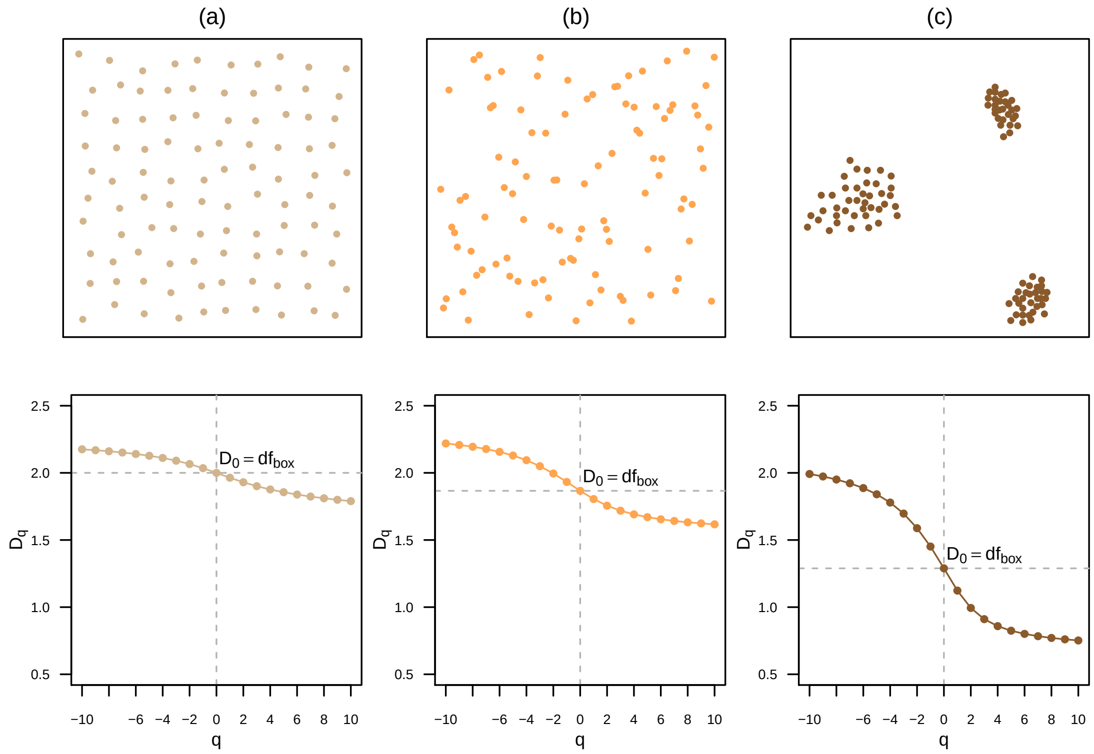

FIGURE 3.8: The Rényi generalised dimensions analysis for the three simulated patterns: (a) regular, (b) CSR and (c) clustered.

The interpretation of the spectrum of the Rényi generalised dimension is the same as for the spatial case, except that for the temporal case, $D_q(t)$ of a monofractal homogeneous distribution will be 1 for all q moments. Thus, for $q > 0$, the multifractal function captures the scaling behaviour of time-intervals where events are more aggregated, while for $q < 0$, the function describes the empty time-intervals. For multifractal sets, $D_q(t)$ decreases as q increases characterising the variability of the probability distribution in the time-intervals at different timescales.

Another important indicator of temporal homogeneity is given by the width of the spectrum of the $D_q(t)$ values ($D_{-\infty}$–$D_{+\infty}$) [272]. This difference is an indicator of the heterogeneity of the temporal aggregation of the distribution, that is, great differences in the number of events in the time intervals. A wider spectrum indicates a more heterogeneous distribution of the irregularities of the pattern. On the other hand, the slope of the $D_q(t)$ spectrum indicates how much clustered is the distribution in the time; where the steeper the slope, the more unevenly and aggregated are the distribution of intervals with high number of events.

3.4.3.3 The multifractal singularity spectrum

The multifractal singularity spectrum is a local parameter [45] that provides an alternative way to describe the scaling behaviour of a pattern through an interlinked set of

Hausdorff dimensions, $f(\alpha)$, associated to a singularity strength α [50, 133, 187]. This singularity strength is an index providing information about the local scaling behaviour of a measure (i.e. the degree of regularity of the measure). Thus, α_i is first calculated for each ith box of size δ as:

$$p_i(\delta) \propto \delta^{\alpha_i} \tag{3.13}$$

where $p_i(\delta)$ is the probability distribution inside the ith box of size δ. Then, the number of boxes, $N_\alpha(\delta)$, having a singularity strength in the neighbourhood of $\alpha(\alpha+d\alpha)$ (that is, boxes with probabilities with the same α), can be related to as:

$$N_\alpha(\delta) \propto \delta^{-f(\alpha)} \tag{3.14}$$

where $f(\alpha)$ is the singularity fractal dimension of the set of boxes with the same α [50, 133, 187]. By changing α, one must obtain a spectrum of $f(\alpha)$, named "singularity spectrum". Although, this spectrum can be computed directly from the data as explained below, it is also related to the Rényi generalised dimensions through a Legendre transform as:

$$(q-1)D_q = q\alpha(q) - f(\alpha(q)) \tag{3.15}$$

(for more details see [133, 179, 187, 257]). Chhabra and Jensen [50] proposed a method estimating the multifractal singularity spectrum directly from the data as a function of the qth order moments without the application of the Legendre transform. Let $\mu(q)$ be the normalised measure of the probabilities in the boxes of size δ, such as:

$$\mu_i(q,\delta) = \frac{[p_i(\delta)]^q}{\sum_i [p_i(\delta)]^q} \tag{3.16}$$

where, again, q provides a tool for exploring denser and rarer regions of the singular measure [257]. Then, α_q and $f(\alpha_q)$ can be computed as:

$$\alpha_q = \lim_{\delta \to 0} \frac{\sum_i^{N(\delta)} \mu_i(q,\delta) \log p_i(\delta)}{\log \delta} \tag{3.17}$$

and

$$f(\alpha_q) = \lim_{\delta \to 0} \frac{\sum_i^{N(\delta)} \mu_i(q,\delta) \log \mu_i(q,\delta)}{\log \delta} \tag{3.18}$$

Thus, the singularity strength α_q is obtained by calculating the slope of the linear regression fitting the data of the plot relating α_q to q (see Figure 3.9 a). The same is done for the singularity spectrum, which is obtained by estimating the slope of the linear regression fitting the data of the plot which relates $f(\alpha_q)$ to q (see Figure 3.9 b). Then, the multifractal singularity spectrum is obtained by relating in a plot $f(\alpha_q)$ with its corresponding α_q (see Figure 3.9 right plot).

Thus, the multifractal singularity spectrum provides a description of the singularities (degree of regularity) of the observed measure μ. The multifractal singularity spectrum is a powerful tool providing the variability in the local scaling properties of a pattern.

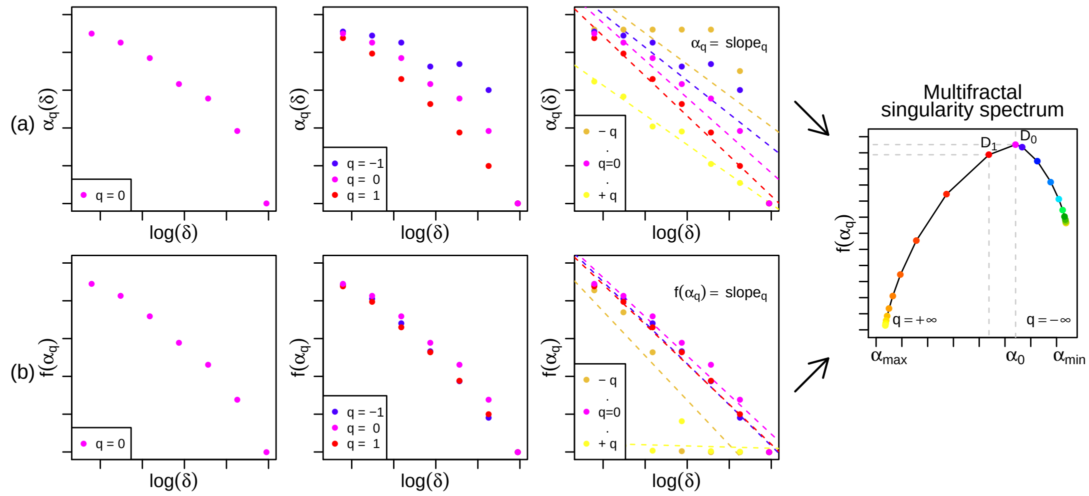

FIGURE 3.9: Graphical computation of the Multifractal singularity spectrum: (a) computation of the singularity strength α_q and (b) computation of the singularity spectrum $f(\alpha_q)$. Then, the slopes of α_q and $f(\alpha_q)$ are related in the same plot (right), the "Multifractal singularity spectrum".

The width of the spectrum ($\alpha_{max} - \alpha_{min}$) allows examining the heterogeneity of the measure (number of events) in an interval $[x, x + \Delta x]$ (Figure 3.9 right plot). The wider the spectrum, the higher the heterogeneity in the distribution of the measure which also indicates the presence of highly clustered areas. Likewise, the height of the spectrum corresponds to the degree of the scaling behaviour. For instance, small values of $f(\alpha_q)$ correspond to rare events (extreme values in the distribution that may be the product of highly clustering behaviour).

For a multifractal process, the spectrum is a bell shaped curve whose maximum value corresponds to the fractal dimension of the support (set of points). While for monofractal patterns, the spectrum is a point with equal values because the pattern presents the same fractal properties (i.e. the same singularities) at every scale and at every qth moment. For $q > 0$, the spectrum provides the behaviour of the strong singularities (clustered regions), while for $q < 0$, the spectrum describes the weak singularites (sparser regions). As for the spectrum of the Rényi generalised dimensions, the multifractal singularity spectrum presents particular cases as well. For instance, for $q = 0$, $f(\alpha_0)$ takes its maximum value and equals D_0, hence, it is also equivalent to the df_{box}; and for $q = 1$, $f(\alpha_1) = \alpha_1$ and equals D_1 which corresponds to the information dimension.

The multifractal singularity spectrums for the three simulated patterns are illustrated in Figure 3.10. Differences in the shapes indicate different processes. For the CSR distribution (Figure 3.10 b), the spectrum of $f(\alpha_q)$ shows again the influence of the edge effects which slightly affects the value of $f(\alpha_q)$ at higher moments. Nevertheless, the curve exhibits a very low degree of multifractality than that of the clustered distribution given by the width of the spectrum. Similarly, it is the case of the regular distribution (Figure 3.10 a) where the ranges of α and $f(\alpha_q)$ suggest a behaviour closer to a single fractal pattern. For the clustered distribution, $f(\alpha_q)$ describes a multifractal behaviour (Figure 3.10 c) designated by the skewness to the left compared to the other two distributions. This indicates a significant degree of clustering of the pattern.

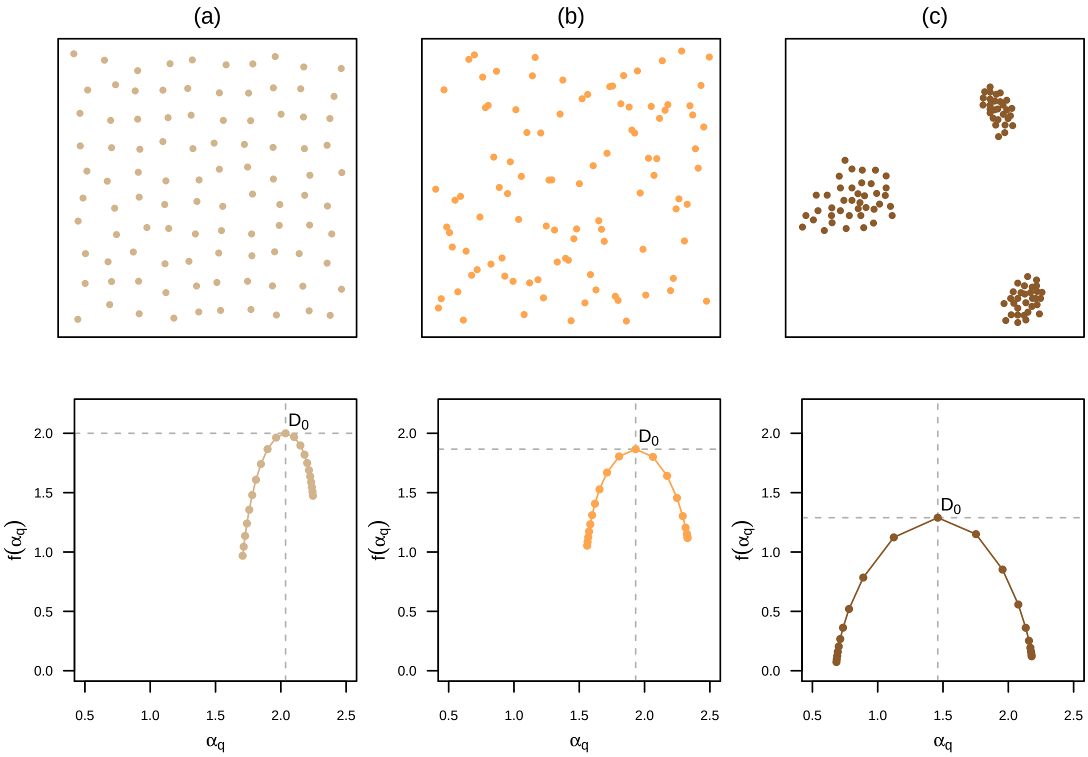

FIGURE 3.10: The multifractal singularity spectrum analysis for the three simulated patterns: (a) regular, (b) CSR and (c) clustered.

3.4.4 Ripley's *K*-function

Ripley's K-function or the "reduced second-order moment measure" describes how the interaction or dependence between events varies through the space (or time) [114]. For a spatial point process, the classical function is defined as [239, 240]:

$$K(r) = \frac{1}{\lambda} E[N(X \cap b(x,r) \setminus \{x\} \mid x \in X] \tag{3.19}$$

where $E[.]$ denotes the expected number of further events within a circle of radius r of an arbitrary event, r is positive and λ is the intensity of the point process (the mean number of events per unit area). For a detailed mathematical discussion see Cressie [74], Daley and Vere-Jones [80] and Illian et al. [143]. In other words, the K-function describes the expected number of events, relative to λ, in a circle (or sphere for 3D events) of radius r centred at an arbitrary event [114].

The K-function is a local neighbourhood measure computed by drawing circles of radius r_1 centred on each event of the pattern (see Figure 3.11). Then, the mathematical expectation is calculated for that radius r_1 and divided by the intensity of the process.

Next, the size of the circle is increased, where $r_1 < r_2 < ... < r_n$, until a maximum size r_n is reached and the mathematical expectation is calculated again for each r and divided by the intensity. The K-function values are plotted against the corresponding radius size r (see Figure 3.11 red line). Finally, the function is compared with a known theoretical curve, i.e. a CSR process (Figure 3.11 gray dashed line), for which the $K(r) = \pi r^2$, that is, the K-function is equal to the area of a circle of radius r [89, 91, 114, 143].

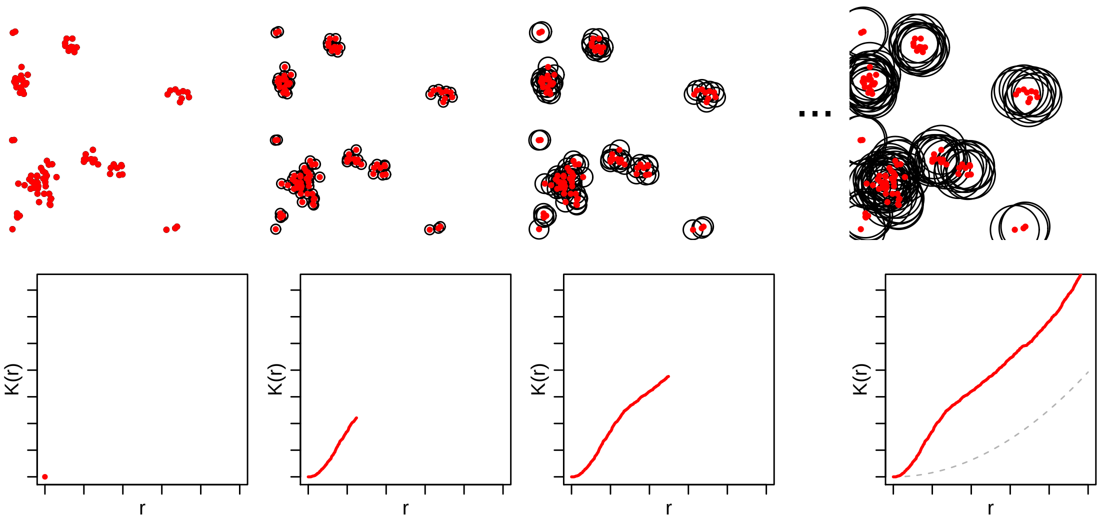

FIGURE 3.11: Graphical principle of Ripley's K-function estimation.

To avoid effects imparted by events falling outside the study area, edge corrections can be introduced in the computation of the K-function. Essentially, if a circle partially falls inside the study area, only the overlapping surface is counted [287]. In literature, some estimators have been proposed in order to correct for edge effects. The most commonly applied edge-correction is the one proposed by Ripley [242], where the estimate of $K(r)$ is expressed as:

$$K(r) = \frac{1}{\lambda} \sum_{i} \sum_{i \neq j, j=1} \frac{I(\parallel x_i - x_j \parallel \leq r)}{N \cdot w(x_i, x_j)} \tag{3.20}$$

where N is the total number of events, the weight function $w(x_i, x_j)$ is the proportion of the circumference of a circle centred at x_i passing through x_j falling inside the study region, and $I(.)$ is the indicator that equals 1 when the distance $\parallel x_i - x_j \parallel$ is less than or equal to r [74].

Estimating the K-function for the three simulated patterns, Figure 3.12, we show that for the random point distribution the $K(r)$ values corresponds exactly to the theoretical values $= \pi r^2$ (Figure 3.12 b). For the clustered distribution, the K-function is greater than πr^2 (Figure 3.12 c) indicating an excess of events at short distances. And for the dispersed distribution (Figure 3.12 a), the $K(r)$ values are lower than the expected values under a Poisson process (πr^2).

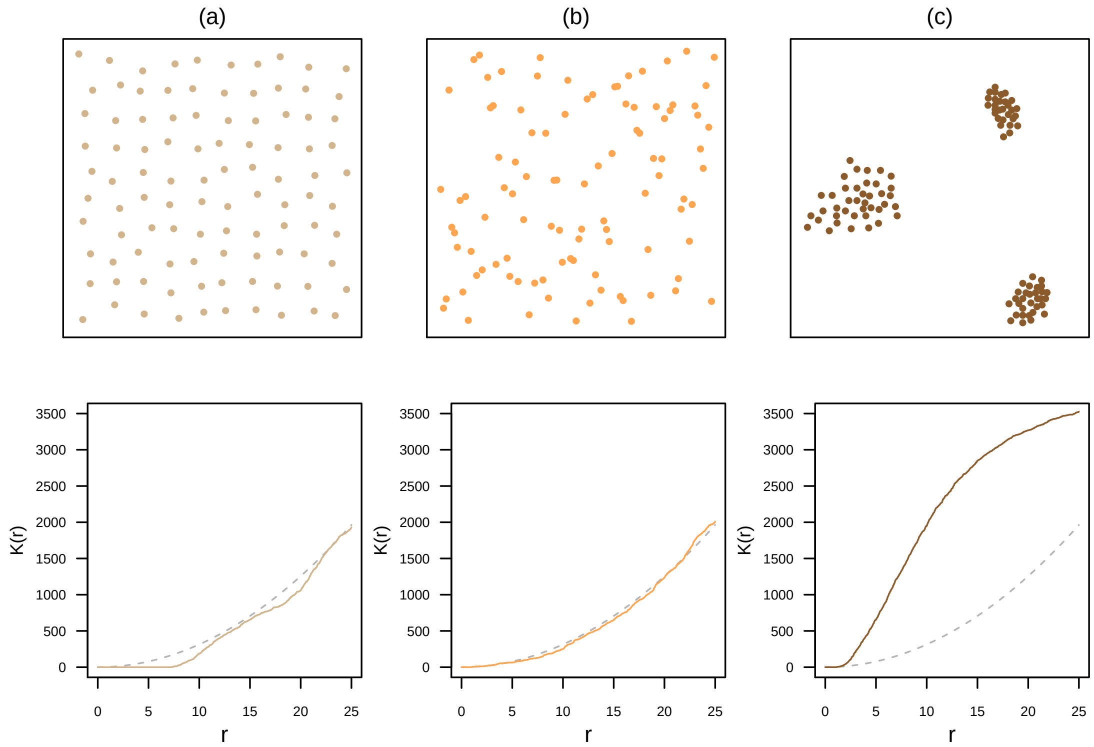

FIGURE 3.12: Ripley's K-function analysis for the three simulated patterns: (a) regular, (b) CSR and (c) clustered. The gray dashed line corresponds to the theoretical value.

In order to facilitate the visualisation of the departure from a CSR process, a transformation of the Ripley's K-function was proposed by Besag [24], the L-function, defined as:

$$L(r) = \sqrt{\frac{K(r)}{\pi}} \tag{3.21}$$

The square root also has the effect of stabilising the variance of the estimator [287] and for a CSR process the theoretical value of the L-function equals r. Plotting the L-function minus r, the theoretical value for a CSR process equals zero, as presented in Figure **??**. The interpretation of the L-function is the same as for the K-function. For a dispersed distribution, the $L(r)$ values are lower than theoretical curve (Figure 3.13 a) and for a clustered distribution, the $L(r)$ values are greater than the theoretical values (Figure 3.13 c).

3.4.4.1 The spatial inhomogeneous K-function

A particular situation encountered in many environmental studies is the non-homogeneity (heterogeneity) distribution of the intensity in the events given mainly by the spatial variability, for instance, of the underlying processes. This considers, then, that the intensity of a specific phenomenon varies across the space for which the assumption of a stationary process is not possible. Therefore, to take into account the local variability

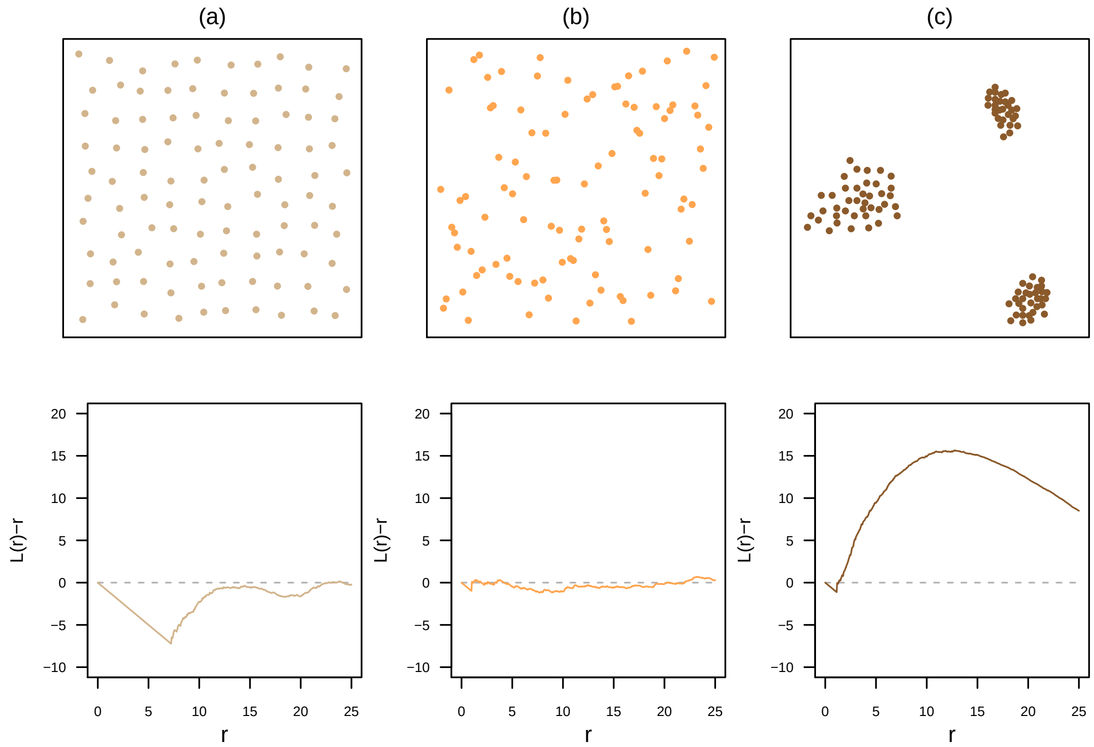

FIGURE 3.13: The L-function analysis for the three simulated patterns: (a) regular, (b) CSR and (c) clustered.

of the intensity characterising the events, one assumes that each point x_i is weighted by its local intensity $\lambda(x_i)$ [287].

A generalisation of the K-function for an inhomogeneous (non-stationary) point pattern was proposed by Baddeley et al. [14]. By using the inhomogeneous K-function, we remove the assumption of an underlying homogeneous point process while still assuming isotropy and stationarity [139]. It is defined as [14]:

$$K_{inhom}(r) = \frac{1}{\mid A \mid} \mathbb{E} \sum_{x_i \in X \cap A} \sum_{x_j \in X \setminus x_i} \frac{I(\| x_i - x_j \| \leq r)}{\lambda(x_i)\lambda(x_j)w(x_i, x_j)} \tag{3.22}$$

where $\mid A \mid$ is the area of the circle A of radius r, $I(.)$ denotes the indicator function, X is the set of all events x_i and x_j, and $w(x_i, x_j)$ is the Ripley's edge correction. The generalisation consists on considering the intensity as a function evaluated now at both locations x_i and x_j, so $\lambda(x_i)$ and $\lambda(x_j)$ represent the mean number of events occurring at locations x_i and x_j respectively.

Intuitively, the inhomogeneous K-function has the same interpretation as the homogeneous K-function [14]. Thus, for CSR events, the $K_{inhom}(r)$ is equal to the area of a circle of radius r ($= \pi r^2$). Similarly, the same considerations when comparing a pattern with the theoretical CSR distribution can be contemplated. For instance, as showed in Figure 3.12 of the K-function for the three simulated patterns, a clustered distribution gives K_{inhom} values greater than πr^2, while for the regular distribution, the values of the function are lower than the CSR process. On the other hand, the K_{inhom}-function can also be transformed to the L_{inhom}-function for an easier visual interpretation of the departure from a CSR process.

3.4.4.2 Temporal K-function

For the purely time analysis, the temporal K-function, $K(t)$, proposed by Diggle et al. [93] can be taught as an analogous formulation of the spatial version of Ripley's K-function. It is used to describe the temporal dependence of events at the timescales t, detecting whether or not the events are temporally clustered, regularly or randomly distributed. For its computation, the spatial parameters of the $K(r)$ (see Section 3.4.4) are replaced by time-based parameters as follows [93] :

$$K(t) = \frac{1}{\lambda_t} E[N(X \cap b(x,t) \setminus \{x\}) \mid x \in X] \tag{3.23}$$

where λ_t is the temporal intensity of the process, and $E[\cdot]$ is the expected number of further events within a time-interval of timescale t of an arbitrary event x.

The temporal K-function is a local neighbourhood measure computed by drawing intervals of equal length t_1 centred on each event of the pattern. Then, the mathematical expectation is calculated for that timescale t_1 and divided by the temporal intensity of the process (total number of events $N(T)$ divided by the total time period T). Next, the size of the timescale is increased, where $t_1 < t_2 < ... < t_n$, until a maximum size t_n is reached and the mathematical expectation is calculated again for each t and divided by their corresponding temporal intensity. The temporal K-function values are plotted against the corresponding timescales t. Finally, the function is compared with a temporal random process, for which the $K(t) = 2t$ [93].

As for the spatial case, the same interpretations for the patterns are considered, where for a clustered time process, the $K(t)$ values are greater than $2t$. While, for the dispersed temporal distribution, the $K(t)$ values are lower than $2t$. Likewise, for an easier visual interpretation of the patterns, the temporal K-function $K(t)$ can be transformed to the L-function $L(t)$.

3.4.5 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 coined in health science by Naus [202, 203], and more recently, still in the health domain, Martin Kulldorff developed the spatial [156] and spatio-temporal extensions [159, 160]. 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 centred on each event (Figure 3.14). Each window moves across the entire area, varying its radius continuously from zero (z_0) up to a fixed upper limit (z_{max}). Each circle takes the nearest neighbour events falling inside (yellow circles) and compares them with those lying outside (black circles). Under the null hypothesis of CSR, these events are inferred to be distributed accordingly to a Poisson distribution which parameters can be estimated. Given this assumption, it is possible to test whether or not these events are randomly and independently distributed in a specific area.

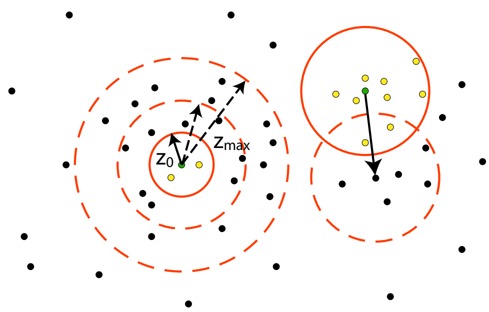

FIGURE 3.14: Principle of the spatial scan statistics: the scanning windows are centred on single events (green dots), increasing their radius (e.i. circles on the left), and moving across the study region (e.i. circles on the right).

For each possible scanning window, a likelihood function is computed as follows [157]:

$$\frac{\mathcal{L}(z)}{\mathcal{L}_0} = \frac{\left(\frac{c(z)}{\mu(z)}\right)^{c(z)} \left(\frac{c(A\backslash z)}{\mu(A\backslash z)}\right)^{c(A\backslash z)}}{\left(\frac{c(A)}{\mu(A)}\right)^{(c(A))}} \tag{3.24}$$

where $c(z)$ and $c(A\backslash z)$ are the number of observed events within and outside the zone z respectively, $\mu(z)$ and $\mu(A \setminus z)$ the expected number of events within and outside the zone z respectively, and $c(A)$ and $\mu(A)$ the number of events and the expected number of events in the total study area. Thus, $\mathcal{L}(z)$ is the likelihood function for zone z, expressing how likely the observed data are given a differential rate of events within and outside the zone, and $\mathcal{L}_0$ is the likelihood function under the null hypothesis (Poisson process).

The probability that a specific zone contains a cluster (or the most likely cluster) is defined by the zone that maximises the likelihood ratio over all possible zones [157]:

$$S_z = max_{z \in A} \frac{\mathcal{L}(z)}{\mathcal{L}_0} \tag{3.25}$$

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, the Monte Carlo hypothesis testing is performed with a large number of random replications generated under the null hypothesis [158]. The rank of the maximum likelihood from the real events is compared with the rank of the maximum likelihood from the CSR replications. 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) [160]. In this way, a cluster can be rejected when its corresponding p-value is greater than the fixed threshold. Based on the desired threshold, the number of the Monte Carlo simulations is established [305].

3.4.5.1 The space-time permutation scan statistics model

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. The cylinder sizes can be increased from zero up to a maximum value in both space (radius) and time (height), see Figure 3.15. As in purely spatial scan statistics, each cylinder visits each event geographical location and, additionally, it visits each possible time period.

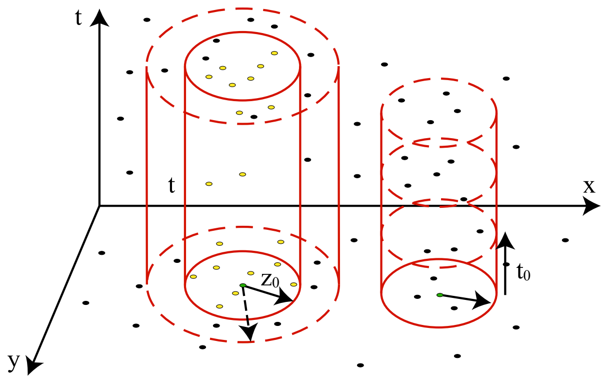

FIGURE 3.15: Principle of the space-time scan statistics: the scanning cylinders are centred on single events (green dots), increasing their radius (e.i. cylinders on the left) and their height based on a time-frame (e.i. cylinders on the right).

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 inside each scanning window. However, when the control population data is not available or not known, the problem is overcome by using the space-time permutation scan statistic model (hereafter STPSS). This model only requires case data which corresponds to the observed events. Thus, the expected number of cases is estimated on the base of the observed events under the assumption of no space-time interaction, meaning that the spatial and temporal locations of all events are independent of each other.

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 [160, 286, 293]:

$$\mu_A = \sum_{z,d \in A} \mu_{zd} \tag{3.26}$$

where

$$\mu_{zd} = \frac{1}{C}(\sum_{z} c_{zd})(\sum_{d} c_{zd}) \tag{3.27}$$

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 [286, 293]. Thus, a Poisson Generalised Likelihood Ratio (GLR) can be computed as follows:

$$GLR = (\frac{c_A}{\mu_A})^{c_A}(\frac{C - c_A}{C - \mu_A})^{(C - c_A)} \tag{3.28}$$

This ratio is calculated and maximised for every possible cylinder and Monte Carlo simulations are performed to test the statistical significance of the detected clusters. For further explanations of the permutation model refer to Kulldorff et al. [160].

3.5 Validity domains and simulated data for the forest fire occurrences in Canton of Ticino

3.5.1 Validity domains

In the previous Section 3.4, we presented some methods used to evaluate and detect the presence of clustering in point patterns: the Morisita index, the Box-counting method, multifractality and the Ripley's K-function. Although many of them are simple and computationally light to apply, the estimation and the interpretation of the real clustering of observed data in terms of statistical significance is often difficult. The correct estimation of the degree of clustering of real data deals with both the finite number of events and the complexity of the geographical space where events are observed [290]; for instance, geographical constraining factors such as lakes, topography, political and socio-economic factors such as administrative limits, build-up areas, road-network, etc. All these aspects define the so called "Validity Domain" (hereafter VD) [290]. Thus, it can be defined as the space of interest which constraints the studied space reducing the dimensionality of the analysed process.

An illustrative example of the VD is given by Tuia and Kanevski [290] and it is exposed here in Figure 3.16.

This figure shows one hypothetical survey of a point pattern distributed in a bounding-box space (Figure 3.16 left), and the same point distribution considering a VD describing the forest cover (shaded area in Figure 3.16 on the right). When considering the bounding-box space (left), the point pattern seems to be clustered; however, when taking into account a more constrained space where the phenomenon is embedded (e.g. the forest cover on the right), the pattern distribution is considered more homogeneous. Thus, the degree of the spatial aggregation of the point pattern is also affected by the limitations of the space where underlying processes take place.

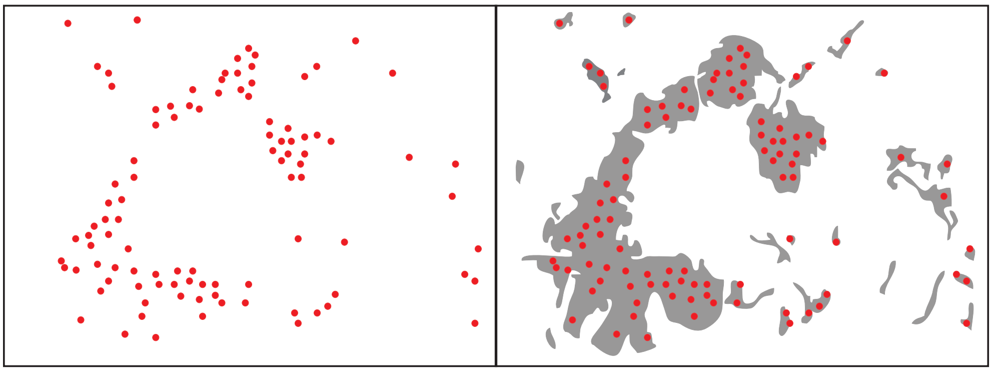

FIGURE 3.16: Example of the importance of validity domains: sampling in a square study region (left) and sampling within a VD of the forest cover (right). Source: image modified from Tuia and Kanevski [290] p. 35.

In this sense, the concept of VD is of great significance in the interpretation of the clustering measures. The location of events outside of the VD regions is not pertinent for the analysis and they can bias the estimation of the real degree of clustering of the studied distribution. Still, the integration of the VD concepts into cluster analysis is not a trivial task. Notwithstanding, one straightforward way is through a comparative analysis as following:

1. Definition and selection of one or several VDs.

2. Generation of many Monte-Carlo simulations (random patterns generated from a Poisson distribution) within the VDs. Each simulation must be composed of the same number of events as the phenomenon under study. Thus, a reference degree of clustering along with a confidence level can be further obtained [120, 143].

3. And finally, the clustering measure is applied to both the observed events (or raw data) and the simulated patterns inside each VD. The different results are analysed and compared. If need, statistical tests can be also applied [25, 120].

The concept of the VDs was applied to the case of forest fires in Ticino in two cases. One case is the two-dimensional space where three geographical spaces are defined as presented in Figure 3.17: (a) the forest cover in the Canton, (b) the political administrative boundaries of Canton of Ticino and (c) a bounding box covering the entire geographical space of the Canton.

The second case is the one-dimensional space for the time analysis, i.e., the time axis. Based on the same principle of the 2-D space, we create two time-VDs as follows:

1. **tVD 1** comprising all the days of the entire studied period; that is, 14,463 days from January 3^{rd} 1969 to August 8^{th} 2008 (Figure 3.18 gray line).

2. **tVD 2** comprising only the days where fires had been detected; that is, 1,653 days out of 14,463 days (Figure 3.18 green lines).

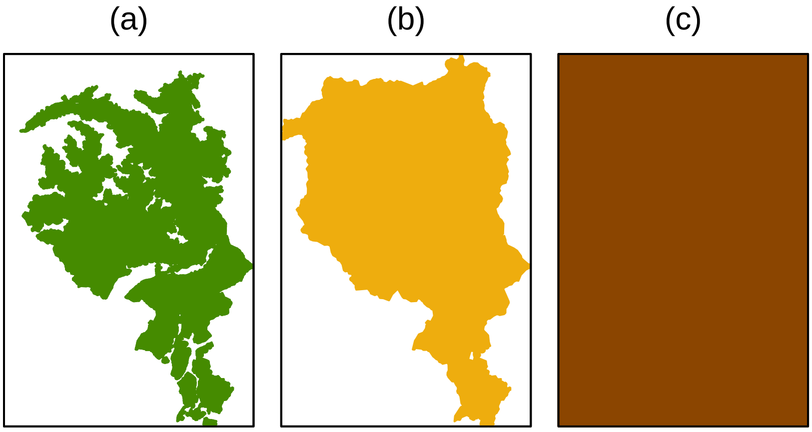

FIGURE 3.17: Validity domains for the forest fires in Ticino: (a) Forest vegetation cover, (b) Political administrative boundaries of Canton of Ticino and (c) Bounding box covering the canton region.

FIGURE 3.18: Time validity domains for the forest fires in Ticino: (gray line) All days in the entire studied period and (green line) days with fire.

3.5.2 Simulated data in VD

Let us note that, contrary to the classic statistical applications, we only dispose of a single realisation (a single measure) of the studied process, i.e the set of points representing the forest fire events. This important limitation obliges us to make assumptions of homogeneity, isotropy and ergodicity in order to carry out the analysis of the structure of the real observed events.

A process is said to be homogeneous if it is invariant by translation, i.e., if its properties do not vary from a point to other one of the space. A process is said isotropic if it is invariant by rotation, i.e., if its properties do not vary with the orientation of the space. And a point process is ergodic if the mean in the probability (concerning various realisations) can be replaced by the spatial mean of the domain of study [74, 151]. The hypotheses of homogeneity and isotropy assure that the set of points presents the same properties everywhere in the field of study. However, with a unique realisation, the hypothesis of ergodicity can not be tested. For this reason, the application of simulations is a crucial tool in the point process statistics.

In the case of the spatial VDs, based on the same number of the observed forest fire events (raw data), we generated 99 CSR patterns (Poisson distributions) by applying Monte Carlo simulations within each of the three VD. These simulated data were used as reference patterns enabling comparisons with the real case. Figure 3.19 shows one sample of the 99 simulations in each VD.

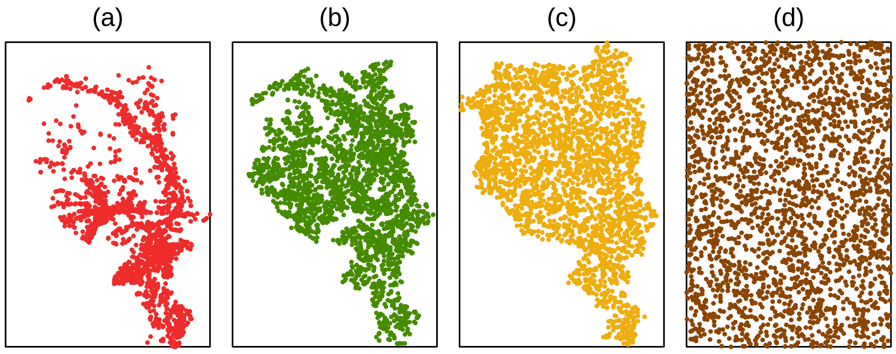

FIGURE 3.19: Simulated patterns in the three spatial VDs for the case of forest fires in Canton of Ticino: a) the observed forest fire pattern (raw data), b) one sample of the 99 CSR patterns simulated inside the forest cover, c) one sample of the 99 CSR patterns simulated inside the political boundaries of Canton of Ticino, and d) one sample of the 99 CSR patterns simulated inside the bounding-box.

For the case of the temporal patterns, based on the same principle of the random simulations in the 2-D space VDs, we have also generated 99 temporal random (TR) distributions in each of the two time VDs as presented in Figure 3.20.

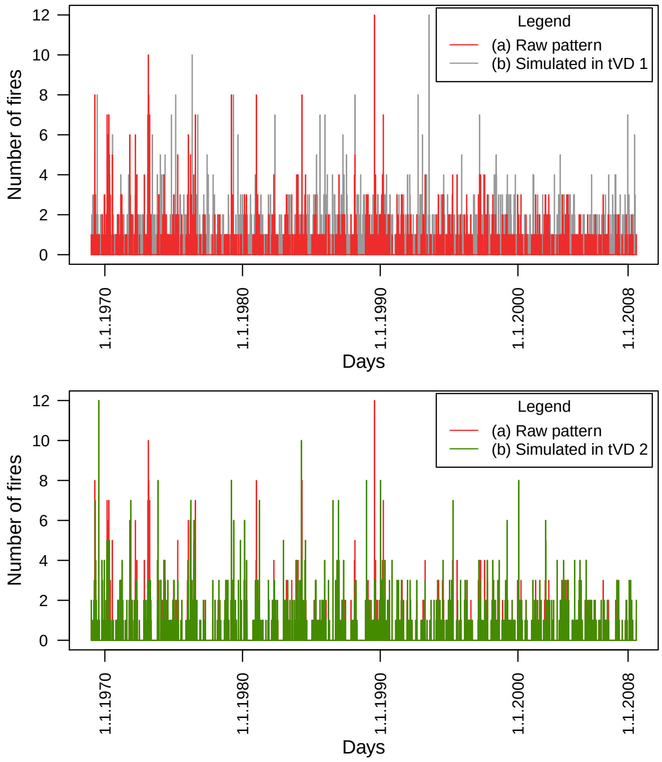

FIGURE 3.20: Simulated patterns in two temporal VDs for the case of forest fires in Canton of Ticino: (in red) the observed forest fire sequences (raw data), (in grey) one sample of the 99 TR patterns simulated inside the tVD 1, and (in green) one sample of the 99 TR patterns simulated inside the tVD 2.

The reference TR patterns were generated by creating 99 random permutations of the frequency of fires per day in order to destroy the dependence existing between the number of events and the temporal location (days) but ensuring that the marginals of the frequency of events remain unchanged. For the tVD 1, the frequency of events per day were shuffled and then assigned randomly to the 14,463 days of the entire studied period (Figure 3.20 upper panel).

On the other hand, for the tVD 2, the frequency of events per day were shuffled and then assigned randomly to the original set of days containing fires (Figure 3.20 bottom panel). Then, the clustering methods are also applied to these simulated samples and their results are compared with the original structure of the forest fire events in Ticino to evaluate the departure of the real pattern from a random structure.

Chapter 4

Applications and results

4.1 Introduction

This chapter presents the adaptation and the application of all clustering methods, presented in Chapter 3, on the case of the forest fire occurrences in Canton of Ticino which is representative of a complex region in the Swiss Alps, and the forest fire case in Portugal. In this Thesis we only concentrate on the issues concerning the occurrence of fire ignition. For that, we are interested in both assessing and understanding the dynamic of forest fire regimes and, additionally, in assessing the potential causes or underlying processes producing and/or controlling the distribution of the observed patterns in space and/or in time.

As presented in the previous Chapter, the fundamental assumption of the adopted stochastic methods taken in this Thesis is that the events are generated by some underlying random mechanisms (i.e. Poisson process). Under this context, the forest fire phenomenon can be modelled as a stochastic point process, for which the main characteristic of the data consists on representing the events in the form of a set of points distributed in the space and/or in time. To represent these events as a point pattern, we considered the ignition point of each fire event. These points describe the geographical location of the fire ignitions (X and Y coordinates) and the date of fire break-out (start date: day/month/year). As additional attributes (marks) we count with information about their ignition causes, burnt area and topographic characteristics such as altitude, slope and aspect.

Each applied clustering method is illustrated with the case of the 2,401 forest fire occurrences registered in Canton of Ticino in the last 4 decades (1969–2008). Moreover, an application of the space-time permutation scan statistic method is also carried out on the case of the forest fire occurrences in Portugal in the last 3 decades (1980–2007) comprising more than 300,000 fire events (Subsection 4.4.2).

Section 4.2 presents the results of the spatial clustering analysis carried out for the case of forest fires in Ticino by applying the following global methods: the Morisita index (subsection 4.2.1), the Box-counting method (subsection 4.2.2), the multifractality formalism constituted by the Rényi generalised dimensions and the multifractal singularity spectrum (subsection 4.2.3) and the Ripley's K-function (subsection 4.2.4). Section 4.3 discuses the results of the time analysis carried out on the same data by adapting the above mentioned spatial clustering measures to the temporal case: the temporal Morisita index (subsection 4.3.1), the temporal Box-counting method (subsection 4.3.2), the temporal multifractality (subsection 4.3.3), and the temporal K-function (subsection 4.3.4). And finally, section 4.4 discuses two applications of the spatio-temporal analysis carried out using the space-time permutation scan statistics (STPSS): 1) The case of the forest fires in Ticino (subsection 4.4.1) to detect and to identify space-time fire-ignition hot spots during three periods (1969–1978, 1979–1990, 1991–2008) using single fire event locations. And 2) the case of forest fires in Portugal (subsection 4.4.2) to assess the robustness of the STPSS model to correctly detect clusters on aggregated datasets and to identify space-time fire hot spots in the Portugal database where the fire occurrences are georeferenced and aggregated at the Parish level (the smallest administrative unit).

4.2 Spatial analysis

Different measures of spatial clustering (fractal, statistical) are considered and applied in the spatial analysis of forest fires in Canton of Ticino. The spatial variability of the forest fires is a very complex process which is conditioned by an intermixture of human, topographic, meteorological and vegetation factors. To compute measures of clustering in complex shaped regions the concept of validity domain is applied to restrict the spatial dimensionality of the phenomenon on the mapping space. Within the validity domain, it is possible to generate spatially randomly distributed events which structure properties are well known. These properties can be compared to the real phenomenon and the deviation between these measures quantifies the real clustering. Each measure is then described and executed for the raw data (observed forest fire events in Canton of Ticino) and its results are compared to the ones obtained from the reference patterns generated under the null hypothesis of spatial randomness embedded in each validity domain. These comparisons enable estimating the real degree of clustering of the raw data.

In this regard, taking into account the spatial characteristics of the forest fires in Canton of Ticino during the period 1969–2008, the spatial clustering measures are used to discover different structures of the analysed phenomenon and the 99 CSR patterns generated in each of the three VDs (see Section 3.5) are used to assess the significance of the spatial clustering. The analysis is also carried out for the phenomenon presented as a marked point process, that is, taking into account additional information (marks) such as the ignition causes. Considering this, we divided the data into natural-caused fires (Raw natural fires) given by the fires ignited by lightning, and anthropogenic-caused fires in order to detect differences in the patterns of these two general causes. For the latest causes, we go further by carrying analysis out for particular anthropogenic-originated fires: the arson-caused fires (criminal actions) and the negligence-caused fires (non-intentional actions). These analyses can reveal information about the underlying processes and their relationship with the phenomenon under study. This section presents the clustering measures developed and completely implemented in R software. For the cases of the Morisita index, the Box-counting method and the multifractal formalism, the used functions were completely created and developed by the author of this Thesis in the R environment (see Appendix B). For the Ripley's K-function, the measure is estimated using the Kinhom and the Linhom functions from the package "spatstat" in R.

4.2.1 The Morisita index

The Morisita index analysis for the observed and the marked forest fire distributions in Canton of Ticino for the period 1969–2008 and the simulated random distributions in the three VDs is presented in Figure 4.1.

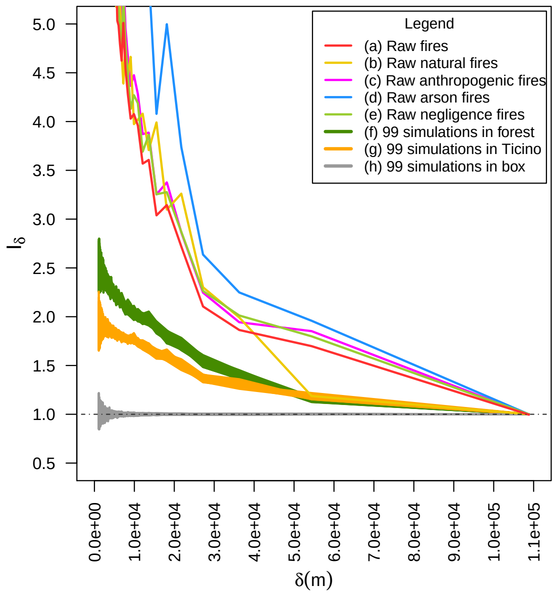

FIGURE 4.1: The Morisita index analysis for the forest fire occurrences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 CSR simulations in the forest cover, (g) 99 CSR simulations in the political boundary limits of Ticino, and (h) 99 CSR simulations in a bounding box located in the geographical region of Ticino.

Looking at the Morisita curves of the CSR patterns in the three different VDs (Figure 4.1 f, g, h), it is possible to depict the constraints imposed by the geographical space defined for both the boundaries of Ticino and the forest cover. The Morisita index for the 99 CSR patterns in the bounding box follows almost the same theoretical behaviour as for an infinite CSR pattern. However, at very lower scales, great fluctuations are detected due to the finite number of events of the forest fire phenomenon.

Comparing the CSR structures with the raw patterns, except for the natural fires, the Morisita index shows significant degree of clustering at every scale (Figure 4.1 a, c, d, e). Natural fires (Figure 4.1 b) exhibit also a high level of clustering, though at scales greater than 52 km it exhibits a random behaviour. One interesting characteristic is that after this scale down forward lower scales, the raw, natural and anthropogenic fires show similar clustering behaviours (Figure 4.1 a, b, c respectively).

Arson and negligence fires also present significant levels of clustering at all scales, however, differences in their patterns are also visible. According to the Morisita index, arson fires are more clustered than negligence fires. These latest fires evidence a clustering behaviour similar to that of all anthropogenic fires; a behaviour that can be explained by the fact that negligence fires account for the 41.69% of all human-caused fires while arson fires only account for the 14.20%.

4.2.2 The Box-counting fractal dimension

The analysis of the Box-counting fractal dimension on the raw and marked forest fires structures is presented in Figure 4.2.

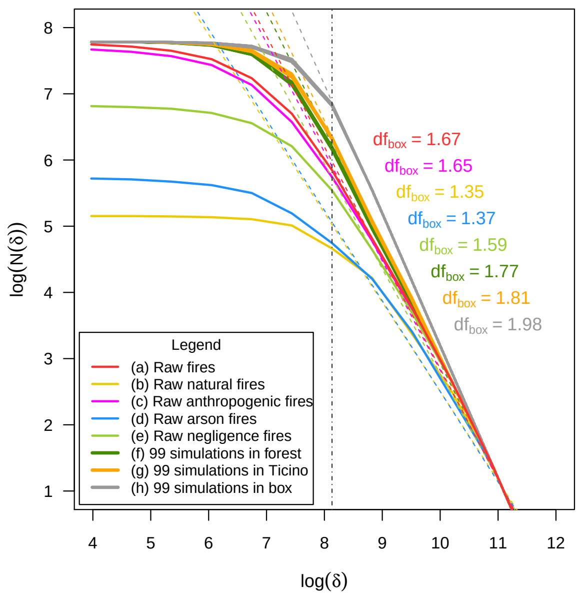

FIGURE 4.2: The Box-counting fractal dimension analysis for the forest fire occurrences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 CSR simulations in the forest cover, (g) 99 CSR simulations in the political boundary limits of Ticino, and (h) 99 CSR simulations in a bounding box located in the geographical region of Ticino. The log() refers to the *natural logarithm* and δ is measured in metres.

As mentioned in Subsection 3.4.2, the fractal dimension of a random 2-D point distribution fluctuates around 2. This theoretical behaviour can be seen in the Box-counting curve of the 99 CSR patterns in the bounding-box, $df_{box} = 1.98$ (Figure 4.2 h), for scales higher than 3 km ($> e^8$). It is not exactly 2 due to the finite number of events (2,401 points in each pattern). The 99 CSR patterns in Ticino boundaries and that on the forest cover manifest a decrease in the fractal dimensions, $df_{box} = 1.81$

and $df_{box} = 1.77$ respectively (Figure 4.2 f, g), because of the geographical constraints of imposed in each VD. Both distributions show similar fractal behaviours; although, differences become visible at scales lower than 3 km ($< e^8$) when their boundaries effects appear. With this, we demonstrate that forest areas are clustered and impose directly a degree of clustering to the forest fire structures. However, can we consider forest fire clustered? Analogous to the Morisita index, the Box-counting fractal dimension of the different forest fire patterns (Figure 4.2 a) exhibits significant levels of clustering which can be depicted by their values notably lower than the 99 CSR patterns in the forest cover, $< df_{box} = 1.77$ (Figure 4.2 f).

The Box-counting fractal dimension for all raw fires is 1.67 (Figure 4.2 a) which shows a departure from a random distribution at scales lower than 14 km ($e^{9.5}$). Similar is the clustering behaviour of the anthropogenic fires with a Box-counting fractal dimension of 1.65 (Figure 4.2 c). Therefore, it is difficult to differentiate from these two distributions. Natural fires evince the highest clustering, $df_{box} = 1.35$ (Figure 4.2 b), at almost all scales comparing to the other structures. This behaviour is followed by the arson fires with a Box-counting fractal dimension of 1.37 (Figure 4.2 d) and with a departure of a CSR behaviour at scales lower than 49 km ($e^{10.8}$). Negligence fires are also highly clustered (Figure 4.2 e) with departure from the CSR process at scales lower than 22 km (e^{10}).

4.2.3 Multifractality

The multifractal analyses are carried out by means of the Rényi generalised dimensions D_q and the multifractal singularity spectrum $f(\alpha(q))$. This multifractal formalism is limited to only positive values of q because, when working with finite datasets, the boxes with very small events will dominate in the entropy. This generates instability in the statistics computation of the box and will dramatically affect the interpretation of the real scaling behaviour of the areas with low number of events. This is the reason why we focus on the characterisation of the areas with relative high number of events (clustering).

These analyses are performed by inspecting the raw and marked forest fire patterns at different levels by changing the values of q. Consequently, the probability distribution, p_i, was calculated as the fraction of the number of events in the i_{th} box of size δ ($p_i = n_i/N$, where n_i is the number of events falling in the box i and N is the total number of fires in Ticino).

4.2.3.1 The Rényi generalised dimensions

Figure 4.3 portrays the Rényi generalised dimensions, for $0 \leq q \leq 10$, used to characterise the multifractal behaviour of the raw and marked forest fire patterns in Ticino. The non-linearity of the D_q curves for all the considered fire structures discloses their multifractal nature.

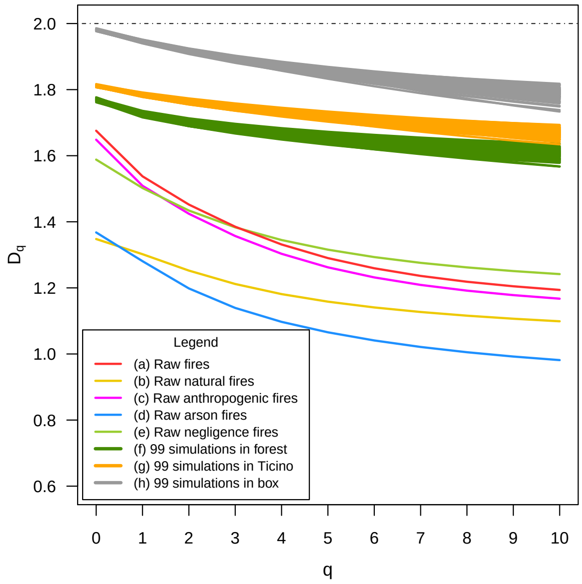

FIGURE 4.3: The Rényi generalised dimension analysis for the forest fire occurrences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 CSR simulations in the forest cover, (g) 99 CSR simulations in the political boundary limits of Ticino, and (h) 99 CSR simulations in a bounding box located in the geographical region of Ticino.

D_q values denote the degree of clustering of the distribution. The D_q curves of the raw and marked forest fire patterns in Ticino do not fall inside the curves of the CSR simulated pattern in all VDs, indicating an important departure from a random process. Furthermore, their curves decline faster than those of the CSR simulations. These substantial departures from randomness stipulate the significance of the clustering of all raw and marked fire patterns. Although, natural fires (Figure 4.3 b) exhibit higher clustering than the anthropogenic fires and the entire fire dataset (given by the lower values of D_q), its multifractal behaviour is less strong than other structures ($D_{q=0} - D_{q=10}$). This means that the areas with lower and higher number of events are not too different.

Regarding the width of the D_q spectrum ($D_{q=0} - D_{q=10}$), all raw, anthropogenic, arson and negligence fires (Figure 4.3 a, c, d, e) exhibit strong irregularities between the areas of high and low number of events. That is, there are great differences between sparser and denser areas. This can also be depicted by comparing the differences between $D_{q=0}$ and $D_{q=1}$. The $D_{q=0}$ values provide information of how the support of the distributions (the points where the events are counted) occupies the geographical space and it is equivalent to the Box-counting fractal dimension presented in the former Subsection 4.2.2. The values of $D_{q>0}$ determine how the events are distributed on the support; with $D_{q=1}$ measuring the pure state of the events distribution.

One interesting point is that looking at arson fires (Figure 4.3 d), its D_q values not only distinguish from a highly clustered distribution (low values of D_q) but it also falls down to a value of 1. As explained in Section 3.4.2, the fractal dimension of a pattern embedded in a 2-D space can range from 0 to 2, where 0 is the topological dimension of a point, 1 the topological dimension of a line and 2 the dimension of a geographical space. Thus, the D_q values of the arson fires evince a behaviour that is closer to fill up a linear geographical space. This makes us associating the distribution of arson fires along linear structures, for instance, the road network. Meanwhile, negligence fires (Figure 4.3 e) are less clustered than the arson fires and follow a clustering behaviour more similar to the raw fires (Figure 4.3 a).

4.2.3.2 The multifractal singularity spectrum

The multifractal singularity spectrum analysis for all fire patterns in Ticino is illustrated in Figure 4.4.

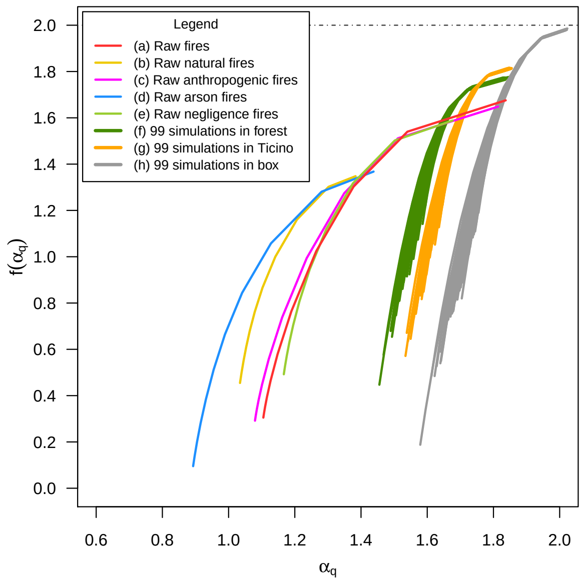

FIGURE 4.4: The multifractal singularity spectrum analysis for the forest fire occurrences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 CSR simulations in the forest cover, (g) 99 CSR simulations in the political boundary limits of Ticino, and (h) 99 CSR simulations in a bounding box located in the geographical region of Ticino.

As in the previous analysis of the D_q values, all fire structures exhibit significant departures from the random patterns in the three VDs. The asymmetry of the raw and marked distributions is more skewed to the left than for the CSR patterns confirming that these structures present a significant degree of spatial clustering. This reflects the uneven distribution of the irregularities of the raw and marked patterns (differences between areas of high and low number of events).

The computed multifractal singularity spectrum is in agreement with the estimations of the Rényi generalised dimensions. The anthropogenic fires (Figure 4.4 c) have a multifractal behaviour similar to that of the raw fires (Figure 4.4 a), but its heterogeneity is just slightly higher as shown by the width of the $\alpha_{max} - \alpha_{min}$ values. This similarity can be interpreted as the strong contribution of the anthropogenic fires to the whole structure. We must remember that these represent the 92.71% of all fires.

Additionally, the multifractal singularity spectrum for natural and arson fires (Figure 4.4 b, d respectively) are more skewed to the left than the other structures, indicating stronger clustering. This is also depicted by their $f(\alpha_q)$ curves which are lower and wider than those of the raw, anthropogenic and negligence fires (Figure 4.4 a, c, e respectively). Nevertheless, the irregularities of the arson fires are more heterogeneous and unevenly distributed than the natural and negligence fires.

4.2.4 Ripley's *K*-function

The Ripley's K-function analysis for the raw, marked and VDs' simulated forest fire patterns is presented in Figure 4.5.

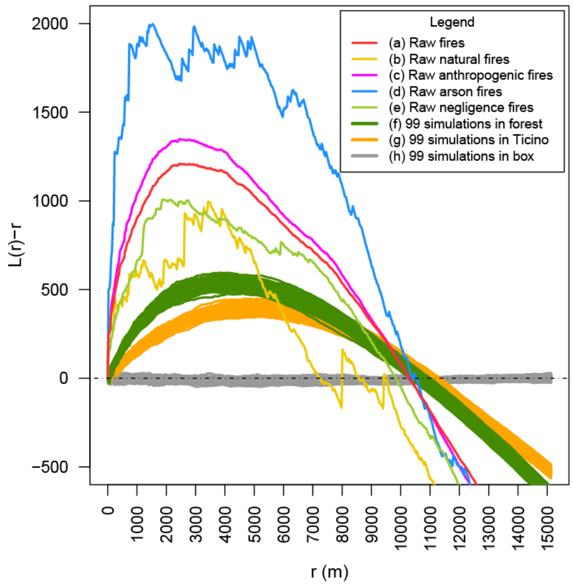

FIGURE 4.5: Ripley's K-function analysis for the forest fire occurrences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 CSR simulations in the forest cover, (g) 99 CSR simulations in the political boundary limits of Ticino, and (h) 99 CSR in a bounding box located in the geographical region of Ticino.

This analysis exhibit the significant clustering of all raw and marked forest fire patterns at scales up to 9 km. That is, the spatial distances at which these patterns are above the curves of the 99 CSR in forest. This K-function analysis also depicts the spatial constraints imposed by the forest cover and Ticino limits shapes, confirming once again the importance of the introduction of the VD concept in the analysis of real point patterns.

It is interesting noticing that raw, anthropogenic, arson and negligence fires (Figure 4.5 a, c, d, e respectively) exhibit almost the same spatial scales of clustering (up to 9–10 km) and maximum of clustering at scales around 3 km. Meanwhile, natural fires (Figure 4.5 b) exhibits clustering behaviour at scales lower than 5 km and a maximum clustering at around 4 km. Arson fires (Figure 4.5 d) show higher clustering behaviour regarding the other patterns, while negligence fires (Figure 4.5 e) are less clustered compared to the patterns of anthropogenic-caused fires.

4.2.5 Conclusions

Statistical, fractal and second-order measures were applied to characterise the clustering behaviour of the raw, marked and CSR simulated structures of the forest fire occurrences in Ticino during the period of 1969–2008. The concept of the validity domain was implemented.

These methodologies allowed quantifying the homogeneity and/or irregularity of the spatial distributions of the forest fire events (raw data, natural and anthropogenic fires, from which arson and negligence fires were also studied). Each clustering measure detected significant degree of clustering of the forest fire structures at every scale. By applying the multifractal formalism we were able to both quantify the degree of clustering of each pattern and the dissimilarities of their clustering properties. We also demonstrated how different constraints of the geographical space, where events are distributed, decrease the dimensionality of the phenomenon on the mapping space. Furthermore, comparisons between the clustering measures of the raw structures and the CSR samples in the three VDs enabled estimating the significance of clustering of the real data. It was shown that the results of the clustering analysis are very different from measures computed in a regular geometrical space (usually considered in the literature) because it includes empty spaces.

The discrepancy between the theoretical values for the random patterns and the obtained values of the CSR patterns in the different VDs can be explained from the finite size of the data (2,401 points of the forest fire database). The generation of simulated data sets (CSR patterns) with similar characteristics of the real data (i.e. the same number of points) allowed reducing the effects of the finite size of the data when applying the clustering measures, and allowed evaluating the significance of the degree of clustering of the real dataset. In this context, the results presented in this Thesis suggest attention when interpreting clustering estimated from experimental datasets.

Thus, these methodologies allowed detecting interaction between events whether the studied point pattern is clustered, randomly or regularly distributed and identifying ranges of spatial scales where events exhibit different behaviours. Moreover, it allowed us identifying underlying processes influencing the distribution of the different fire structures. Other several distributions were compared to infer dependence between variables (marks), but these analyses are not presented in this document, for instance:

- Human fires occurring in the winter season (Nov–Apr) are more clustered than those fires occurring during the summer season (May–Oct).

- Winter fires are more clustered than summer fires.
- Arson and negligence fires are distributed dependently of each other.

Understanding these structures can assist fire managers and policy-makers to conduct fines distribution of fire-fighting resources. This knowledge can also support nearest-neighbour analyses to identify areas with higher fire incidences.

4.3 Time analysis

The temporal fluctuations of events can be characterised by means of temporal point process models which describe the clustering patterns of observed events distributed in time (one dimension). In this context, the forest fire occurrences in Ticino can be modelled as stochastic time point processes where the events are characterised by their occurrence in time.

The theoretical development of the applied measures was established in Chapter 3 for the adaptation of the spatial measures to time analyses. Here we show the application of those clustering measures reformulated to analyse the time fluctuation of point patterns. The basic principle is that the points are considered occurring along the time axis instead of a 2-D space.

Figure 4.6 shows the daily time sequences of forest fires in Canton of Ticino in the period 1969–2008. This visualisation displays a forest fire activity inhomogeneously spread along the time. This fire regime periodicity can be explained by meteorological conditions [324, 325]. But, are there other mechanisms influencing these sequences?

In this context, clustering analyses can help at describing different dynamics that are not easily detected by a simple scatter plot. Thus, purely time analyses on the forest fire event sequences in Canton of Ticino in the period 1969–2008 are carried out considering the entire data (raw fires), marked point processes taking into account the ignition causes: natural and anthropogenic-caused fires, which in turn, are also analysed for the particular cases of arson and negligence causes (as presented in the previous Section 4.2), and the 99 random patterns generated in each of the two time VDs (Section 3.5), used to asses the significance of the temporal clustering of the real data.

This section presents the spatial clustering measures adapted to time analysis. They were developed and completely implemented in R software. For the Morisita index, the Box-counting method and the multifractal formalism, the used functions were completely created and developed by the author of this Thesis (see Appendix B). For the temporal K-function, the measure is estimated using the function `stkhat` from the package "splancs" in R environment.

4.3.1 The Morisita index

The estimation of the Morisita index for time analysis of the raw and marked forest fire sequences in Ticino and the random simulated patterns in the two tVDs is presented in

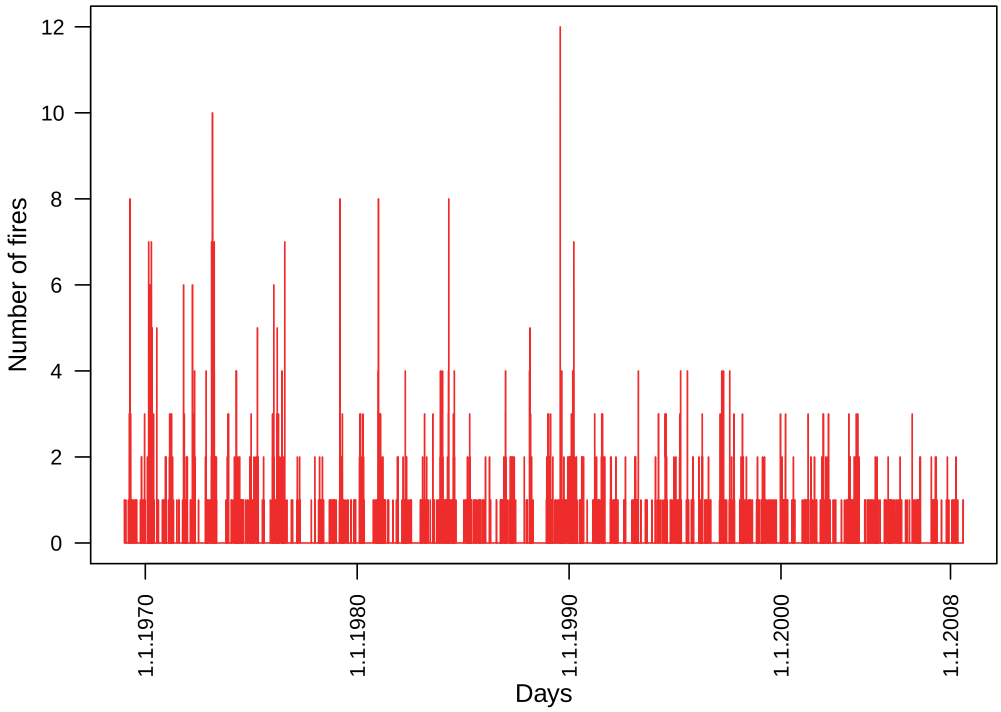

FIGURE 4.6: Daily time series sequences for forest fires in Ticino from 1969 to 2008.

Figure 4.7.

The Morisita index exhibits high degree of temporal clustering of all fire sequences in Ticino at almost all timescales, given by the visible departure of the raw fires (Figure 4.7 a) from the random patterns in the two tVDs (Figure 4.7 f, g).

The marked natural and anthropogenic fires are also highly clustered (Figure 4.7 b, c respectively). Greater time fluctuations are exhibited in the case of natural fires due to the small number of events (174 fires) during the entire period. Likewise, the time sequences of arson and negligence induced fires present also an important degree of temporal clustering (Figure 4.7 d, e respectively). However, fires caused by arson actions show a temporal clustering behaviour only at scales lower than 20 years.

It is also evident a slightly degree of temporal clustering of the random patterns in the tVD 1 at smaller timescales (Figure 4.7 f), while for the random patterns in tVD 2 the degree of clustering is greater (Figure 4.7 g). This can be explained by the fact that the number of days presenting fire occurrences is less than the number of days in the entire studied period. This condition imposes a certain level of temporal constraint of the events. Therefore, the departure from temporal randomness is not evaluated by considering the theoretical value of 1 as it is defined in the case of an infinite point process, but instead, it is evaluated by looking at the departure from the Morisita values of the simulated random patterns in the two tVDs.

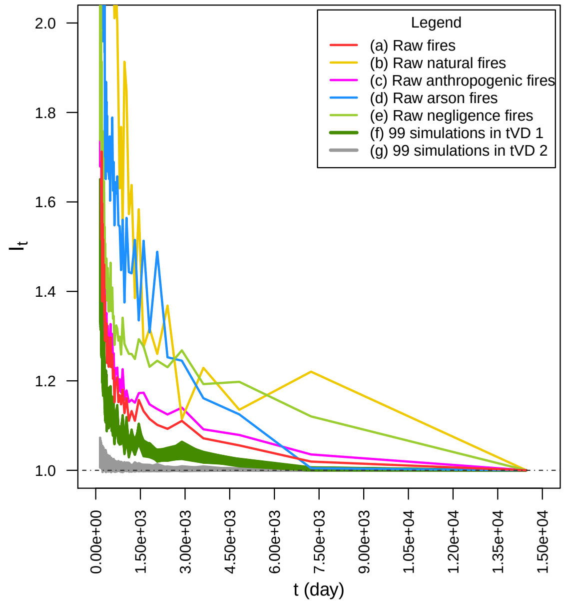

FIGURE 4.7: The Morisita Index analysis for the forest fire sequences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 TR simulations in tVD 1, and (g) 99 TR simulations in tVD 2.

4.3.2 The Box-counting fractal dimension

The Box-counting fractal dimension analysis for the raw and marked forest fire sequences in Ticno during the period 1969–2008 and the temporal random simulated patterns in the two tVDs is showed in Figure 4.8.

This analysis for the temporal fire structures in Ticino were defined from timescales of 40 years ($e^{9.6}$) down to 20 days (e^{3}). This interval corresponds to the timescales where the simulated patterns in tVD 1 exhibit temporal randomness ($df_{box} = 0.99$). In this interval, the fractal dimension of the raw fires is 0.93 being slightly lower than the random patterns in tVD 1 (Figure 4.8 a, f respectively). This indicates a feeble temporal clustering of the raw data. Nevertheless, down to a timescale of 90 days ($> e^{4.5}$) forest fires do not exhibit temporal clustering because its fractal behaviour falls inside the envelope of the random patterns in the tVD 1.

Moreover, the Box-counting fractal dimension of the raw data is exactly the same as the fractal dimension of the random patterns simulated in tVD 2. This fractal behaviour is completely expected as the Box-counting method does not distinguished from the number of events inside the time intervals, instead, it only considers whether the time interval contains events or not (a binary count). It is necessary to remember that tVD 2 consists on the same days of fire as the raw data. Thus, the intervals with fires (without considering the number of fires) are the same for the two distributions: the raw data and the TR patterns of the tVD 2. Consequently, in this case, it is impossible differentiating the real pattern from the TR structures in tVD 2.

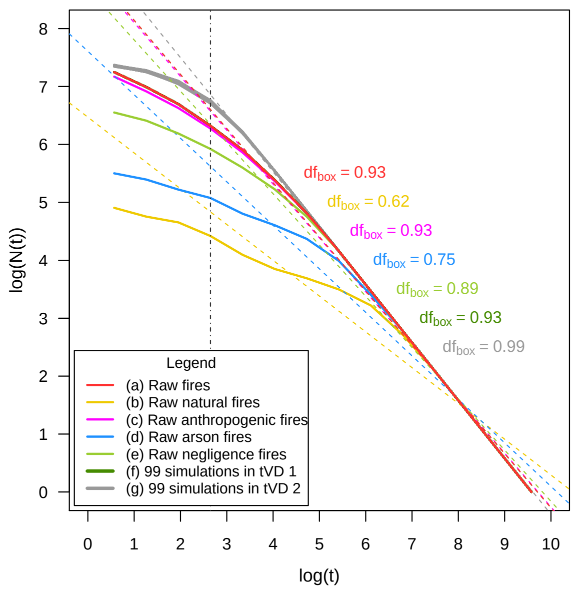

FIGURE 4.8: The Box-counting fractal dimension analysis for the forest fire sequences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 TR simulations in tVD 1, and (g) 99 TR simulations in tVD 2. The log() refers to the *natural logarithm* and t is measured in days.

The marked natural and anthropogenic fires present different temporal structures (Figure 4.8 b, c respectively). Greater time fluctuations are exhibited in the case of natural fires showing a high level of temporal clustering ($df_{box} = 0.62$) with a departure from temporal randomness at timescales lower than 3 years (e^7). The temporal anthropogenic pattern behaves equals to the temporal pattern of all fires together. Again, the method is not able to distinguish between these two temporal structures and the random patterns in tVD 2 ($df_{box} = 0.93$). This can be explained by the fact that anthropogenic fires are present along the entire studied period, thus, the time-intervals containing anthropogenic fires are almost the same as for the raw fires and the random patterns in tVD 2.

On the other hand, the time sequences of arson and negligence induced fires exhibit temporal clustering with fractal dimensions of 0.75 and 0.89 respectively (Figure 4.8 d, e). This reveals that arson fires are more clustered than negligence fires. Furthermore, arson fires show a departure from temporal randomness at timescales lower than 1.5 years ($e^{6.3}$) while negligence fires exhibit a departure at timescales lower than 6 months ($e^{5.2}$).

4.3.3 Multifractality: the Rényi generalised dimensions

The time multifractal analysis for the raw and marked forest fire sequences in Ticino during the period 1969–2008 and the temporal random patterns simulated in the two tVDs was carried out by means of the Rényi generalised dimensions (for $q \leq 0$) and presented in Figure 4.9.

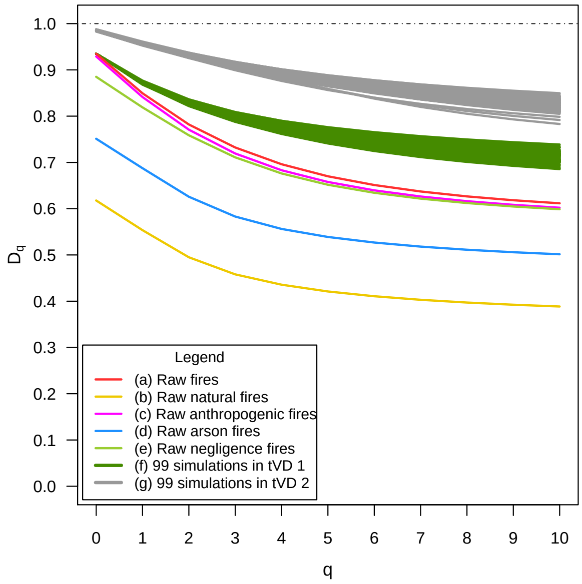

FIGURE 4.9: The time multifractal analysis based on the Rényi generalised dimensions for the forest fire sequences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 TR simulations in tVD 1, and (g) 99 TR simulations in tVD 2.

All of the observed fire patterns exhibit a multifractal behaviour which is depicted by the non-linearity of the dependence between the $D_q(t)$ values and their corresponding q order-moments. Their substantial departure from the unstructured temporal distributions (Figure 4.9 f, g) indicates the significance of the temporal clustering of the observed fire structures. Additionally, natural fires (Figure 4.9 b) are more clustered than the other fire structures given by the lower $D_q(t)$ values.

The multifractal analysis also reveals that anthropogenic and all raw fires present higher complexities than the natural, arson and negligence fires ($D_{q=0}$–$D_{q=10}$). This complexity indicates that the intervals with high and lower number of events are more heterogeneous and unevenly distributed along different timescales. Regarding the width of the $D_q(t)$ spectrum, the five observed fire structures (Figure 4.9 a - e) reveal a certain irregularity between the time-intervals with high and low number of events; though, it is more accentuated in the anthropogenic induced fires. This can also be depicted by comparing the differences of $D_{q=1}$–$D_{q=0}$.

Unlike the Box-counting method, the multifractal analysis is able to differentiate the structure of the raw fire sequences from the random patterns in tVD 2. Similarly, the method allows differentiating them from the structure of the anthropogenic fires. Moreover, the reason why the random patterns show themselves multifractal behaviours is due to the fact they were built under the same frequency of fires of the raw data.

4.3.4 The temporal K-function

The temporal K-function analysis for the raw and marked forest fire sequences in Ticino during the period 1969–2008 and the temporal random patterns simulated in the two tVDs is carried out by means of the L-function and presented in Figure 4.10. This function is a transformation of the K-function facilitating the graphical interpretation of the curve.

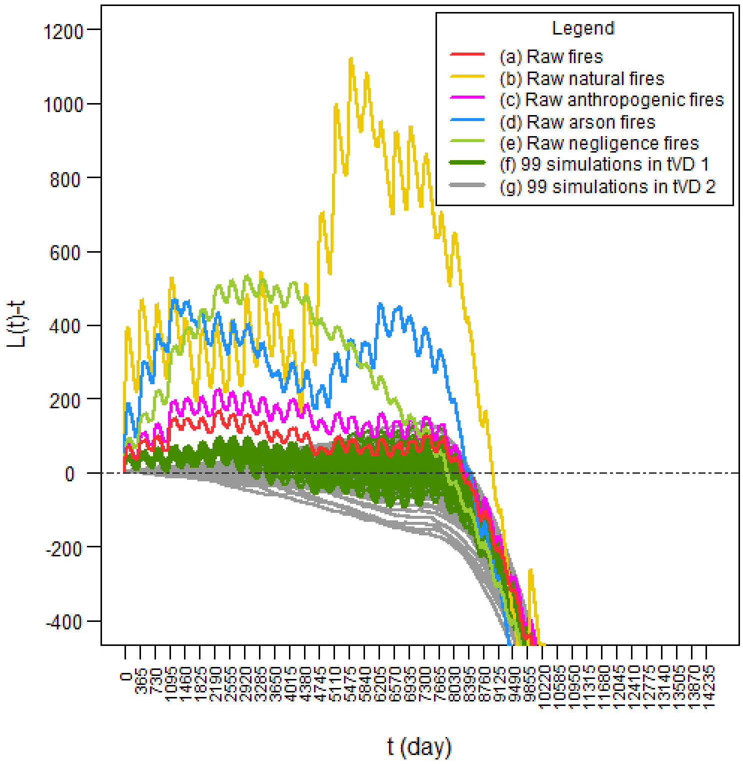

FIGURE 4.10: The time K-function analysis by means of the L-function for the forest fire sequences in Ticino in the period 1969–2008: (a) Raw pattern (all fires), (b) Raw natural fires, (c) Raw anthropogenic fires, (d) Raw arson fires, (e) Raw negligence fires, (f) 99 TR simulations in tVD1, and (g) 99 TR simulations in tVD 2.

The curves of the natural, arson and negligence fires (Figure 4.10 b, d, e respectively) exhibit significant temporal clustering for time lengths of 24, 23 and 19 years respectively (where the curves intersect the random patterns in tVD 1 and tVD2). Natural fires (Figure 4.10 b) present a temporal clustering peak at a timescale of 15 years. Negligences fires (Figure 4.10 e) exhibit a temporal clustering peak at 8 years, while arson fires (Figure 4.10 d) show two temporal clustering peaks at 3 and 17 years respectively.

For the case of the raw forest fires and the anthropogenic fires, the level of temporal clustering is less significant than the above mentioned structures. These patterns portray temporal clustering for timescales up to 13 and 16 years respectively (Figure 4.10 a, c). However, they both present a clustering peak at a timescale of 7 years. On the other hand, all fires show temporal aggregations at the season level ($\sim$ 6 months, the small peaks of the curves) with yearly fluctuations. These fluctuations are more pronounced on the natural fires because these events only occur during the summer season and not along the entire year.

4.3.5 Conclusions

We have studied the properties of the temporal structure of forest fires sequences in Ticino by means of the stochastic point processes formalism. The time dynamics of fire sequences in Ticino were characterised and quantified by means of temporal analogous forms of the spatial statistical measures explained in Chapter 3. All these clustering measures were reformulated to perform purely temporal analysis. Computations were programmed and implemented in the R project environment. The functions for the Morisita index, the Box-counting method and the multifractal formalism were programmed by the author of the Thesis, while the function for the temporal K-function, we used *splancs* package.

These statistical measures are closely linked through the fractal theory offering an interesting interpretation in terms of departure from temporal randomness. Each measure was executed for the raw data (the forest fires) and the marked point processes such as fires caused by human, natural, arson and negligence actions. They were also compared to patterns of reference generated under the null hypothesis of temporal randomness embedded in the same time period of the raw data and with the same number of fire events as the original data. The comparison enabled estimating the significance of the deviation of the real data from temporal random processes. The results revealed the presence of significant temporal clustering of all analysed structures at different timescales: days, months and years. According to Telesca et al. [279], these temporal clustering behaviours can be related to a mix of both meteorological (natural) conditions and anthropogenic factors, conducting to a highly concentrated pattern of the forest fire dynamics.

An important contribution of this research deal with the adaptation, analysis and estimation of these global clustering measures for time sequences. This helps understanding the temporal structure of point events. Furthermore, the information resulting from these analyses are useful for the definition of the hyperparameters of the local clustering detection and mapping methods such as Scan Statistics [305] or space-time kernel density applications.

4.4 Space-time analysis

The space-time analysis in this Thesis is carried out by the implementation of the space-time permutation scan statistics model (STPSS) developed by Kulldorff et al. [160]. The analyses presented in this section were published in the articles of Vega Orozco et al. [305] and Pereira et al. [214]. Here, we present only the discussion of the analyses carried out in each article, thereby, for further details one can refers to the concerned article.

The spatio-temporal analysis is carried out by means of the space-time permutation scan statistics (STPSS) applied to the phenomenon of forest fires in two different cases: 1) First case is presented in Subsection 4.4.1 developed for the forest fire occurrences in Canton of Ticino during three periods 1969–1978, 1979–1990 and 1991–2008. This study intended at detecting and identifying space-time fire hot-spots applying the scan statistics method in a data with single event locations. 2) The second case is presented in Subsection 4.4.2 using the forest fire occurrence database of Portugal during the period 1980–2007. This investigation aimed at both assessing the robustness of the model to correctly detect space-time clusters on aggregated data and identifying space-time clusters in the Portugal database with the fire events aggregated at the Parish level (the smallest administrative unit).

Although these two phenomena represent the same environmental problem, they present completely different characteristics either from the temporal/geographical contexts where events occurred as for the construction of the database where events are organised. Thereby, different challenges and scientific interests are encountered in the application of the mentioned methodology for the spatio-temporal cluster detection in the two cases.

The STPSS model has proved to be very useful for the analysis of the distribution of environmental data. The main advantage of this statistical tool is that it only uses the observed cases instead of requesting both the case and the population-at-risk data. 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 localise 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 using the SaTScan^{TM} software developed by Martin Kulldorff [158]. This program allows the user indicating all the requirements (hyperparameters) to perform the analysis, such as input data, coordinates system, study period, number of Monte Carlo replications, etc.

4.4.1 The space-time permutation scan statistics model using single events locations

This analysis aims at identifying hot spots in forest fire sequences by means of the STPSS model and a geographical information system (GIS) for data and results visualisation, as well as to analyse the ignition causes of the resulting clusters. The proposed scan statistical methodology (see also Subsection 3.4.5) uses a scanning cylindrical window, which moves across the 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 fire occurrences in Ticino from 1969 to 2008. This dataset consists of 2,401 georeferenced single fire events including the location of the ignition points and additional information (marks). 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 Thesis: one analysis using all the events enabling an overall investigation of all fires, and the second analysis considering only the fire events caused by lightning (natural fires). Because the STPSS method does not distinguish between clusters generated by an increase risk of fire or by a different geographical event distribution at different times [158], 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, thus:

- **Dataset I** contains 833 fires occurring in the period 1969 to 1978,
- **Dataset II** holds 762 events happening between 1979 and 1990, and
- **Dataset III** comprises 806 fires burning between 1991 and 2008

The definition of these three datasets is based on different fire-preventative dispositions that conditioned the distribution and frequency of anthropogenic fires in Ticino in the last 40 years, such as: the major fire brigades reorganisation implemented in 1978, the systematic use of helicopters for both transport of the fire fighters and aerial fire-fighting since 1980; and the implementation of two preventive legal acts (1989 and 1991) aiming at prohibiting burning activities in the open spaces [220]. The lightning fire dataset was separately analysed over the entire study period (1969–2008) comprising a total of 175 events, because their fire regime is quite different from the anthropogenic-caused fires, and their ignition occurrences are not affected by the measures mentioned above. Concerning the input data and the study period, cluster analyses are conducted for four datasets as mentioned above and for each dataset a simulation is performed using the SaTScan software [158]. All fire events were specified as individual locations with each point representing one fire occurrence (case) and the related date of fire-ignition.

The results are non-overlapping clusters identified using the retrospective space-time permutation model. For the definition of the scanning space-time cylindrical window parameters in SaTScan, multiple simulations were executed using different maximum-size values. These hyperparameters were also compared with different comprehensive spatial and temporal structural analyses that were completed in the previous Sections 4.3 and 4.2. Only the two most representative results, consistent with qualitative analyses from the forest fire experts in Ticino, are presented in this paper. The scanning space-time cylindrical 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.

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

Results of the STPSS model of the analysis completed for datasets I, II and III are presented in Table 4.1. By means of GIS techniques, the resulted clusters are graphically mapped in Figure 4.11.

TABLE 4.1: Significant clusters of the STPSS analysis for datasets I (1969–1978), II (1979–1990) and III (1991–2008) of the forest fire occurrences in Ticino.

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	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

Source: Vega Orozco et al. [305]

Ten significant clusters are distinguished: 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 are predominantly localised in the hilly deciduous forest belt of the southern part of the study area (Sottoceneri), where the highest population density is encountered. These clusters are mainly defined by 74 fires 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).

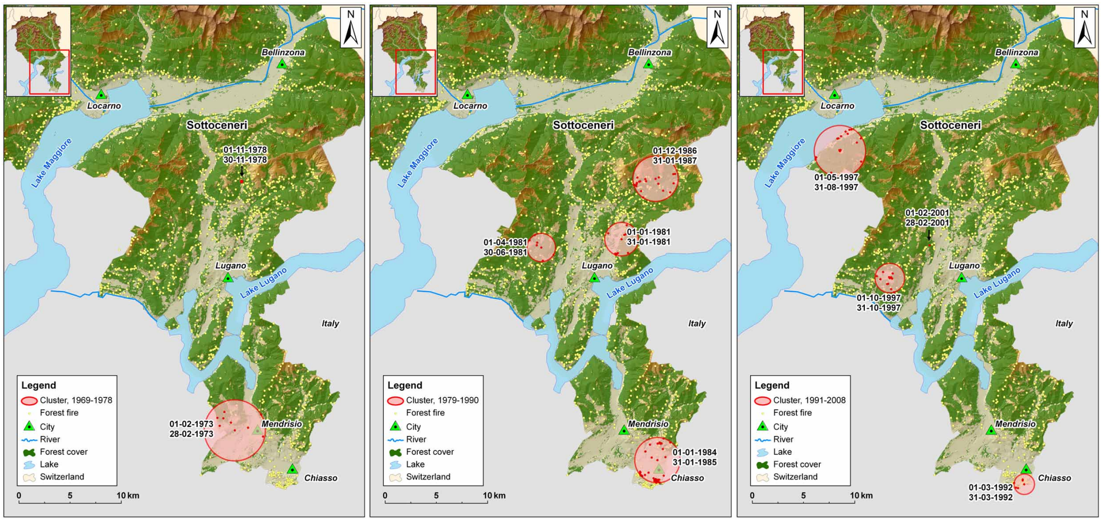

FIGURE 4.11: Visualisation of the significant clusters resulting from the STPSS analysis for datasets I - 1969–1978 (left), II - 1979–1990 (centre) and III - 1991–2008 (right) of the forest fire occurrences in Ticino.

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 are 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 and all started by anthropogenic actions. In this region, cluster 1 from the sub-period 1969–1978 (Figure 4.11 left) has the greatest radius range of all ten identified clusters with a value of 2,987 m spotted in February of 1973. The second cluster in this area is cluster 2 from sub-period 1979–1990 (Figure 4.11 centre) presenting the longest time frame (1 year).

The second zone is the central part of the Sottoceneri region near Lugano city. Four clusters are recognised: two from dataset II (Figure 4.11 centre) and two from dataset III (Figure 4.11 right). These clusters are largely ignited by criminal actions (arson). 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 broadleaves forests and chestnut stands. In this zone a particular cluster (cluster 2 from dataset III, Figure 4.11 right) consisting of 4 fires, all detonated by arson actions at the same geographical location.

Forest fires caused by lightning display very different spatio-temporal patterns; consequently, it is indispensable to make a distinction between forest fires due to natural causes and to anthropogenic activities [224, 298, 318]. The outcome clusters are reported in table 4.2 and displayed in Figure 4.12.

As presented in Figure 4.12, 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 human development is settled. It is

TABLE 4.2: Significant clusters of the STPSS analysis for natural fires in Ticino in the period of 1969–2008.

Study period	Cluster	Radius (m)	Time frame	Observed cases	Expected cases	p-value
1969–1978	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

Source: Vega Orozco et al. [305]

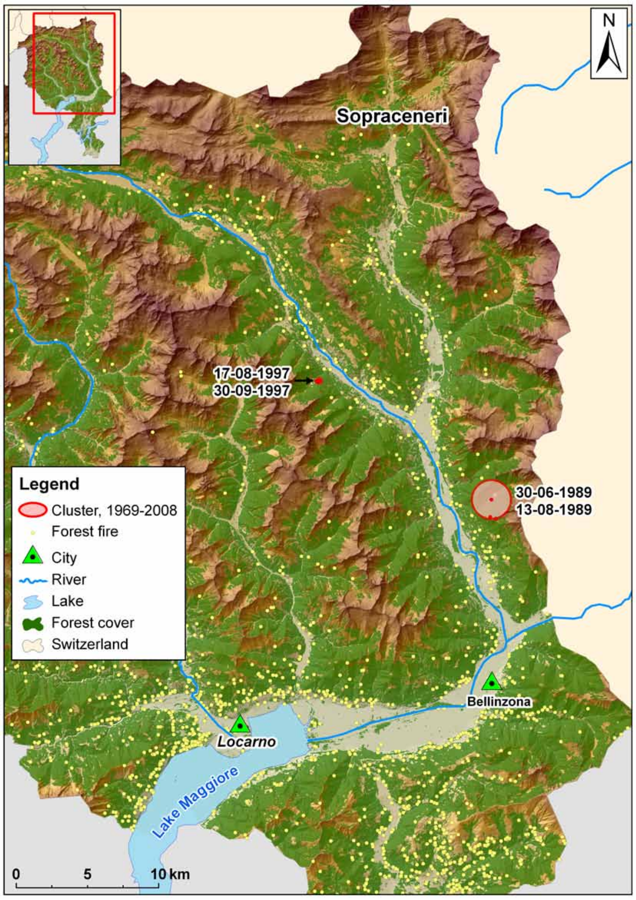

FIGURE 4.12: Visualisation of the significant clusters resulting from the STPSS analysis for natural caused fires (lightning) in Ticino in the period of 1969–2008.

important to call to mind that the original geodatabase reports only lightning-fires 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 1,200 to 1,720 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 1,017–1,120 m.a.s.l., extreme slopes between 85 and 112 % and all burned in spruce stands.

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 is compared with respect to the whole area during the same time- frame of the cluster. Results of these analyses are shown in table 4.3; 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.

TABLE 4.3: Relative frequency (RF%) of fire causes inside the clusters and the ratio (R) compared with the fire causes outside the clusters for the same period.

Study period	**Cluster**	**Arson**		**Negligence**		**Unknown**		**Natural**	
		RF (%)	**R**	**RF (%)**	**R**	**RF (%)**	**R**	**RF (%)**	**R**
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	0.0	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	50.0	1.6	0.0	0.0	0.0	-

a - no arson outside the cluster is detected.

Source: Vega Orozco et al. [305]

Clusters from dataset I (1969–1978) exhibit mainly frequencies of fires generated by negligent human actions. Dataset II (1979–1990) 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 (1991–2008) 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.

4.4.2 The space-time permutation scan statistics model using aggregated data

In this subsection, the STPSS model is applied to the case of forest fire occurrences in Portugal. We have decided to add this analysis in this Thesis because this dataset is different from that of the forest fires in Ticino imposing other scientific challenges for the implementation of the method. This work was achieved under the collaboration established with the *CITAB*[1] of the University of Trás-os-Montes and Alto Douro (Portugal).

This study focuses on the use of the STPSS model to assess both the existence and the statistical significance of clusters on aggregated datasets. Here, only the most relevant results are exposed, for further details refer to Pereira et al. [214]. The investigated

[1]Centre for Research and Technology of Agro-Environment and Biological Sciences

case study is represented by the Portuguese Rural Fire Database (PRFD) containing the forest fire occurrences during the period 1980–2008 georeferenced to the Parsih level (the smallest administrative unit) where they were detected. One important characteristic of this database is that the fires are aggregated to geographical regions instead of using single fire event locations. Considering this, the main goals are as follows:

1. Assessing the robustness of the STPSS to correctly detect clusters on aggregated datasets;

2. Testing the existence of space-time clusters in the PRFD; and

3. Characterising the detected clusters.

In order to assess the potential bias introduced by the aggregation of the data on the performance of the STPSS method, a synthetic database was designed (this part is not developed in this Thesis, for details, refer to the article directly [214]).

Portugal has two official observed fire databases, both provided by the Institute for the Conservation of Nature and Forests (ICNF, 2013), namely the national mapping burnt areas (NMBA) for the years 1990–2011 and the previously mentioned PRFD [213]. This latest database is constructed from ground measurements and provides detailed temporal information (date and time of both fire ignition and extinction); yet, the spatial information is restricted to the name of the smallest detectable administrative regions (i.e. the parish) where the fire was ignited. The history and characteristics of the PRFD was recently assessed by Pereira et al. [213]. This database comprises more than 300,000 forest fires during the period of 1980 to 2007. In addition, it contains comprehensive information on the amount of burned area of each fire (≥ 0.1 ha), the affected land cover type, the fire ignition and extinction dates and time. The geographical coordinates of the parishes' centroids are assigned to each fire in the PRFD (Figure 4.13). This process naturally leads to an artificial aggregation of fires with potential impact on the cluster analysis; this is enhanced by the inhomogeneity of the parishes' sizes, ranging between 0.06 km^2 and 442 km^2.

According to the results of Pereira et al. [213], only fires with burnt area equal and greater to 0.1 ha are included in the analysis. For the STPSS analyses, the maximum spatial and temporal window sizes were set up equal to the 50% of the observed cases and the 50% of the study period respectively. Furthermore, the database was split into four subdatasets considering different burnt area thresholds to identify different patterns for small, medium, large and very large fires. Therefore, we carried out STPSS analyses separately for the subdataset containing only all fires (burnt area (BA) $\geq$ 0.1 ha), the subdataset of fires with BA $\geq$ 1 ha, the subdataset of fires of BA $\geq$ 10 ha and the subdataset of fires of BA $\geq$ 100 ha.

Figure 4.14 shows the most significant clusters obtained by the STPSS model. These analyses reveal the existence of statistically significant space–time clusters in all cases, with different distribution patterns. The temporal dimension of the identified clusters

Figure 4.13: Parish units of Portugal and their centroids assigned to each fire event. Source: this image was taken from the article [214].

ranged between 1 to 13 years. Likewise, the spatial size ranges from few kilometres up to the maximum size allowed by the analyses.

The obtained results for fires with BA > 0.1 ha and BA > 1.0 ha (Figure 4.14 a, b respectively) are very similar regarding their duration, spatial size and location. In fact, the two main detected clusters (1 and 2) have approximately the same size and reveal an aggregation of fires during the same periods (1982–1988 for cluster 1 and 1995–2007 for cluster 2). These clusters are geographically located in the centre (cluster 1) and north (cluster 2) of the country in the mountainous region where the greatest number of fires is recorded every year. Forest management held during the 80s in the central region should have favoured the decrease in the fire occurrences after this decade, whereas the rural abandonment in the inner northern region should have contributed to the increase of fires in the most recent period (1995–2007). A third small cluster is identified when considering all fires in the database (BA > 0.1 ha).

For larger fires BA > 10 ha (Figure 4.14 c), four clusters are spotted, all of them in the northern half of the country at locations slightly centred eastwards. Finally, for very large fires (BA > 100 ha) the characteristics of the detected clusters are radically different (Figure 4.14 d). The most significant cluster is centred in Guarda district during the year 1994. The large cluster in the south of the country is due to the large fires in 2003 and 2004 in the districts of Faro, Portalegre and Castelo Branco, formerly

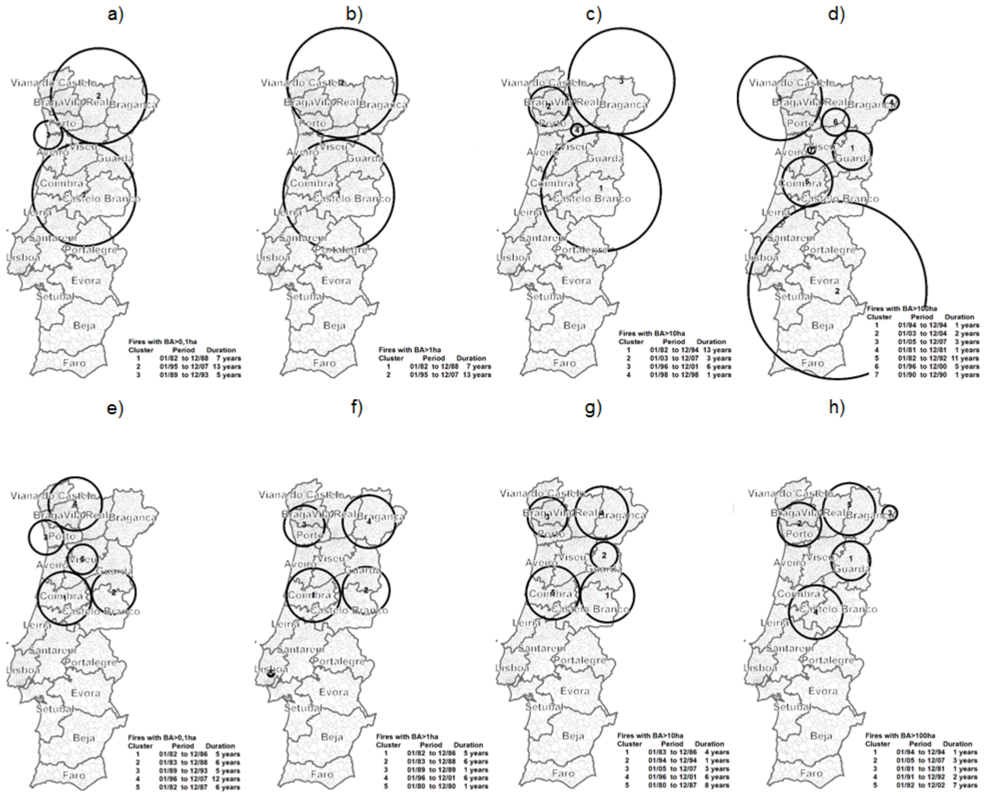

FIGURE 4.14: Visualisation of the most significant clusters resulting from the STPSS analysis for the forest fire occurrences in Portugal. Clusters of fires with burned area $\geq$ 0.1 ha, 1.0 ha, 10.0 ha and 100.0 ha (from the left to the right, respectively). In the top panel, clusters are detected with a maximum scanning window size equal to the 50% of the observed events and 50% of the total temporal dimension. On the bottom panel, clusters are detected with a maximum scanning window size equal to 50 km radius and 50% of the total temporal dimension. Source: this image was taken from the article [214].

identified by Trigo et al. [289] which explain them as a consequence of extreme weather conditions. A third cluster was detected in the NW region of the country with a duration of three years (2005–2007). Another cluster is uncovered in the central area spanning 11 years (1982–1992). Three other clusters are also revealed in the north-central regions characterised by smaller spatial scales of less than 50 km and lasting between 1 and 5 years.

The vast spatial distribution of fires in Portugal makes the definition of the maximum spatial window size a difficult task. Therefore, the analyses are repeated again reducing the spatial scanning window at a maximum radius of 50 km ($z_{max} = 50$ km). Considering the quasi-rectangular shape of Portugal, this chosen value roughly corresponds to the half of the diameter of the shortest West-East distance of the country ($\sim$ 200 km). The reduction of the maximum size of the spatial window (z_{max}) basically led to an increment of the number of statistically significant clusters and to a decomposition or reduction in the size of the larger clusters (Figure 4.14 e–h). It is also worth noting that the decomposition/reduction of the larger clusters appears to comply with the order of the statistical significance of the detected clusters. This becomes evident by comparing the upper and the bottom panels of Figure 4.14, which can be summarised as follows:

1. Clusters in the central regions, which completely fall into the country boundaries, are detected as neighbouring smallest clusters with the only exception of the case of BA $\geq$ 100 ha, as a consequence of the great distance between the locations of the fires implicated inside the cluster area.

2. Big clusters detected in the north and exceeding the country boundaries were identified again within a smaller area.

3. When allowing for very large scanning windows, the detected clusters can be the result of a lower density of events outside the cluster region instead of a true over-density of events inside the cluster area. This is especially true when the number of observed cases is low, as it can be seen when considering BA $\geq$ 100 ha: for instance, cluster 2 in Figure 4.14 d is not statistically significant when imposing a lower z_{max}.

Finally we can conclude that, in the context of the present study, clusters detected by imposing an z_{max} of 50 km are more representative of the reality.

4.4.3 Conclusions

This subsection proposed the application of the STPSS model for forest fire events represented as point patterns, and allowed both assessing and characterising the spatio-temporal clusters of forest fires in Ticino and in Portugal. With this analyses we were able to evaluate the capability of the STPSS method to correctly identify real clusters in two different datasets presenting different characteristics: one dataset (forest fire occurrences in Ticino) representing single fire events, while the second dataset (forest fire occurrences in Portugal) represents fire events aggregated at the Parish level.

The detection of spatio-temporal fire hot spots, which was achieved with the STPSS, is a very important task for performing a better management and implementation of fire-fighting measures. Furthermore, the analysis of the predominant ignition-cause inside the detected clusters can be of great utility when mitigating different fire related problems.

The results of the evaluation of aggregated events also provides useful analyses not only for further studies in the field of fire management, where data are often organised in aggregated formats, but also in other fields where the data are structured in the same format, as it is often the case.

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 [158] 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. Nevertheless, one must keep in mind that the number and size of the detected clusters are strongly dependent of the spatial and temporal sizes of the scanning windows. The results of the analyses suggest that a good practice is to allow for larger maximum scanning windows, compared to the overall size of the study area, in a first stage of the detection analysis.

It is important to underline some of the specificities and limitations of this study. The scanning windows used in this research had cylindrical shapes, yet other geometrical shapes for the scanning window are also possible. Nonetheless, this study was not intended to determine the most appropriate form of windows to detect clusters of fires. On the other hand, we must be aware that clusters resulting from the STPSS analyses could be the product of either an increase in the fire-ignition risk (e.g. arson activities, pasture-fire practices during non-fire-weather conditions, negligence behaviour) or from changes in structural-ignition conditions for specific zones in a particular time-period, 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 [305].

Moreover, 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 to identify fire-prone zones. Outcomes of the STPSS model simulations were integrated into a GIS environment allowing the exposition of the detected clusters.

Chapter 5

Geospatial data mining and mapping

5.1 Introduction

With the statistical analysis developed in Chapter 4, we demonstrated that the forest fires in Ticino are significantly clustered and that most of them break out close to urban developments and are strongly related to human-environmental factors. In general, we can consider that the temporal characterisation of fire regimes can mainly be explained by the weather and the climate conditions favouring the aggregation of fire ignitions in certain periods of time (e.i. the seasonal regimes exposed in the temporal analysis in Section 4.3). On the other hand, in space, anthropogenic and ground (topographic) conditions play an imperative role in the geographical distribution of the fire activity. Indeed, in Section 4.2, we showed that most of the forest fire hotspots caused by anthropogenic sources (e.i. negligence or arson causes) are located near important urban areas such as Lugano, Mendrisio and Chiasso (see Figure 4.11 for the scan statistics permutation model analysis). Furthermore, the multifractal analysis of forest fires in Ticino suggested that the spatial distribution of arson fires (fires caused by criminal activities) follow linear structures that can be related to the road network (see Figure 4.3 d).

In this regard, and concerning only the spatial distribution of the human-caused forest fire occurrences in Canton of Ticino, anthropogenic and ground factors related to forest fire activity can be used for both characterising and mapping vulnerable zones exclusively related to fire management problems such as the wildland urban interface (WUI). Thanks to the collaboration with the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), the knowledge acquired in Chapter 4 was used for defining and characterising the WUI in the Swiss Alpine region. This research was developed under the framework of the project "Wildland Urban Interface (WUI) and forest fire ignition in Alpine conditions - (WUI-CH)", co-directed by the Swiss Federal Office for the Environment and the WSL research program "Forest and Climate Change - Phase I". The main objective of this research was to propose a systematic, statistical and flexible approach for assessing and mapping the WUI in Ticino. This methodology consists on the following three main steps: 1) carrying out a random forests (RF) analysis (a machine learning algorithm) to select the anthropogenic and ground variables most relevant to fire distribution; 2) performing a statistical analysis to define a buffer distance between the anthropogenic variables selected with the RF analysis and the wildland vegetation (according to the user's needs); and 3) building a GIS-based procedure (ModelBuilder) to automatically mapping the resulting WUI.

Furthermore, the implementation of the RF analysis not only allowed selecting the variables that most contribute to the fire ignition distribution and which are used to define the WUI, but also led us to the estimation of the prediction of the fire occurrence. This prediction conveyed to the elaboration of a susceptibility map of fire ignition occurrences in Canton of Ticino during the period 1990–2008. The results of this study were published in the article Conedera et al. [68] and presented in several national and international conferences as well (see Appendix A. Publications & Proceedings).

Accordingly, this Chapter is based on the mentioned study; however, only the methods and the results developed and obtained by the author are shown such as the data preparation and preprocessing, the implementation of the RF algorithm for variable selection, the development of a GIS ModelBuilder routine and the elaboration of the final maps. This analysis is also restricted for the case of Canton of Ticino in order to remain in the study area presented along this Thesis. Additionally, a second study is also developed in this Chapter addressing the mapping of susceptible areas for forest fire events. These two by-products are the first two attempts for forest fires activity in a complex mountainous region as it is the case of Canton of Ticino and for which fire managers currently focus their attention for fire management, monitoring, prevention and allocation measures.

In this sense, both the WUI characterisation and the fire occurrence prediction (susceptibility mapping) together with the previous spatio-temporal statistical analysis in Chapter 4 provided a comprehensive inside and useful information for understanding predominant fire regimes and for identifying fire vulnerable regions in a mountainous context. This inclusive assessment is fundamental to conduct fire management goals and strategies in different spatio-temporal conditions, and to support policy and decision-making for fire mitigation and risk monitoring.

Section 5.2 explains the data elaborated and used for this research concerning the anthropogenic and topographical variables that are related to the human-caused forest fire occurrences during the period of 1990–2008. Section 5.3 imparts the contributions of the author in the methodology for WUI definition and mapping for Canton of Ticino. Section 5.4 presents the methodology developed for the elaboration of the susceptibility map of forest fire ignition in Canton of Ticino based on the variable importance measure provided by the RF algorithm which allows estimating the probability of ignition of fire occurrence. And finally, Section 5.5 presents the conclusions.

5.2 Data preparation and pre-processing

As mentioned before, this Chapter deals only with human-caused forest fires, anthropogenic infrastructure and topographic data of Canton of Ticino during the period 1990–2008. The reason of restricting the analysis to this frame period is due to the need of using a forest fire dataset consistent with the last updated features information of the Swiss Topographic Landscape Model (TLM3D). That is, we use the infrastructure data existing at the time of the fire ignition. This model, updated since 2008 by the Swiss Federal Office of Topography Swisstopo, is a big 3D geodatabase covering the entire country. It includes natural and artificial landscape features such as roads and tracks, public transportation, buildings, areas, land cover, hydrography, administrative boundaries, toponomy, and the digital terrain model (DTM).

Concerning the forest fire ignition points, only proven direct anthropogenic fire ignitions were considered. Specifically, we did not considered fires that were originated from indirect anthropogenic activities such as high voltage power lines, natural causes such as lightning-induced fires and summer fires of unknown origins. Since winter fires are all of anthropogenic origins [219], we retained those fires of unknown origins breaking out at least 50 metres away from railways or high voltage power lines. Taking into consideration all these conditions, we finally count with a total number of 672 fires (see Table 5.1 and Figure 5.1).

Table 5.1: Anthropogenic forest fire ignitions in Canton of Ticino during the period 1990–2008.

HUMAN CAUSE	**No. fires**
Agriculture	6
Arson (criminal)	151
Forestry	1
Negligence fires (tourists, hikers, thrown cigarettes)	271
Not extinguished camping fires	7
Other known causes	28
Other winter fires of direct human origins	208
TOTAL	672

The variables representing the human infrastructure considered in the present study are: railways, highways, roads (down to 2 metres wide unpaved drivable roads including those for 4WD tracks) and pathways considered as the transport facilities (see Figure 5.2) and the buildings and vineyards representing human activity. This information is presented at 0.2 to 1.5 metres resolution.

In order to use these anthropogenic variables, we transformed these shapefile data into raster layers (objects are composed by cells) of 25 metres cell-size (see Figure 5.3). Then, raster maps giving the Euclidean distances from each cell to the closest source were calculated for each infrastructure feature (see Figure 5.4). This procedure was carried out by using the “Euclidean Distance” tool from the “Spatial Analyst” toolbox

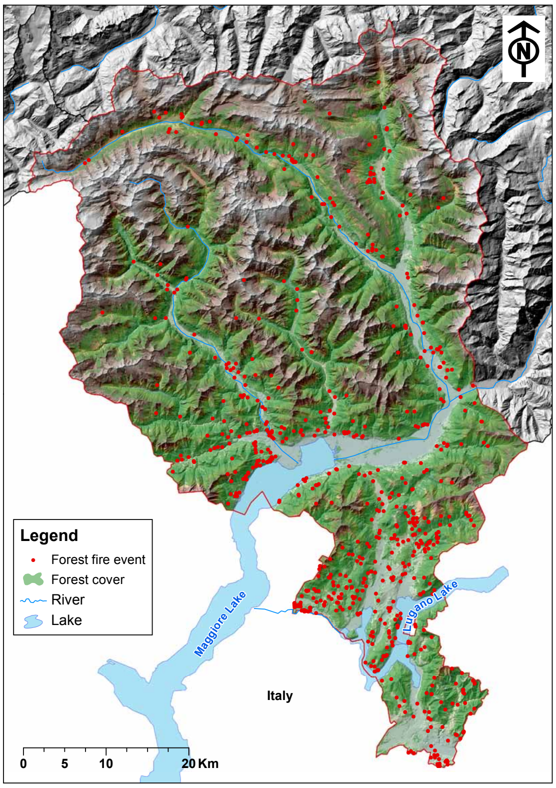

FIGURE 5.1: Anthropogenic forest fire distribution in Canton of Ticino in the period 1990–2008 (red dots) and the forest cover (green areas).

in ArcGIS. The maps of distances to the human infrastructures are presented in Figure 5.5.

Regarding the topographic components, we considered the altitude, the slope, and the aspect which in turn it was divided into the North-South direction and West-East direction (see Figure 5.6). These features were extracted from the digital height model (DHM) of 25 metres resolution provided by Swisstopo. And finally, for the wildland component where fires take place and from where the susceptibility mapping is predicted, we considered the forest cover which comprises the following categories: "forest stands", "open forests" and "shrubland forests", extracted from the Swiss TLM3D layer (see Figure 5.1) of 1 to 3 metres resolution.

5.2.1 Data for WUI definition

For the definition of the WUI, the used dataset comprised only 672 direct anthropogenic forest fire ignition points below 2500 m.a.s.l. (assumed as the upper limit of regularly undertaken human activities in the Alps) [68], and their related anthropogenic infrastructure information which represent the urban component of the WUI (see Subsection 1.4.2). In order to relate these variables to the forest fire ignition points, we extracted their corresponding values of the distance maps from each ignition point.

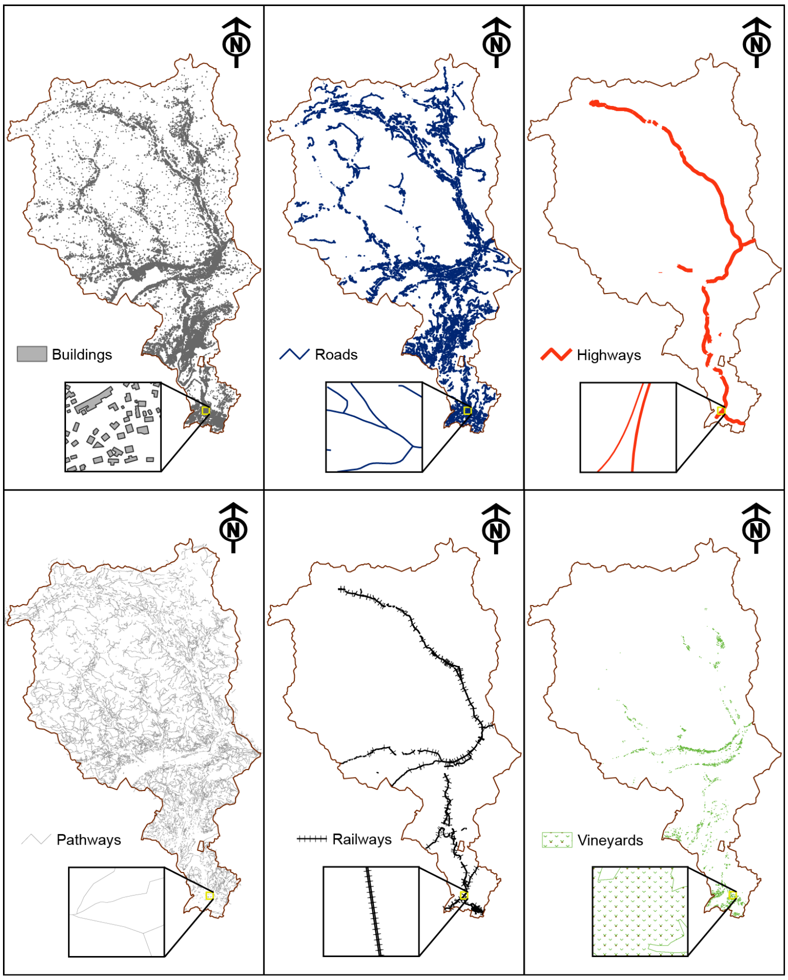

FIGURE 5.2: The considered anthropogenic infrastructures (left to right): buildings, roads, highways, pathways, railways and vineyards.

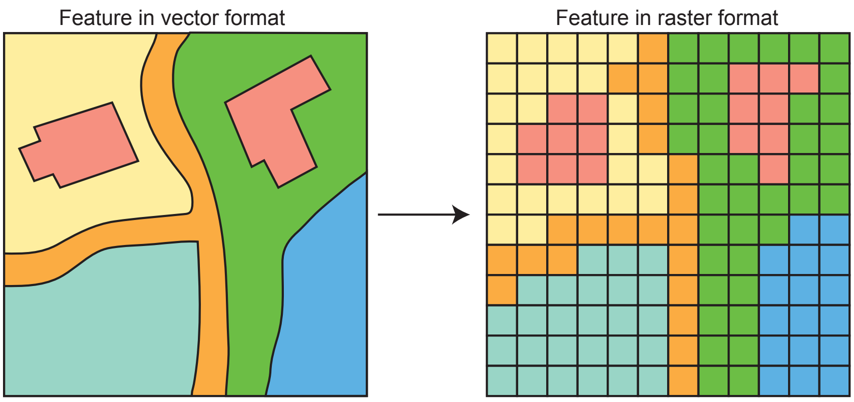

FIGURE 5.3: Transformation of features in vector format to raster. Source: image modified from the ArcGIS Help on line.

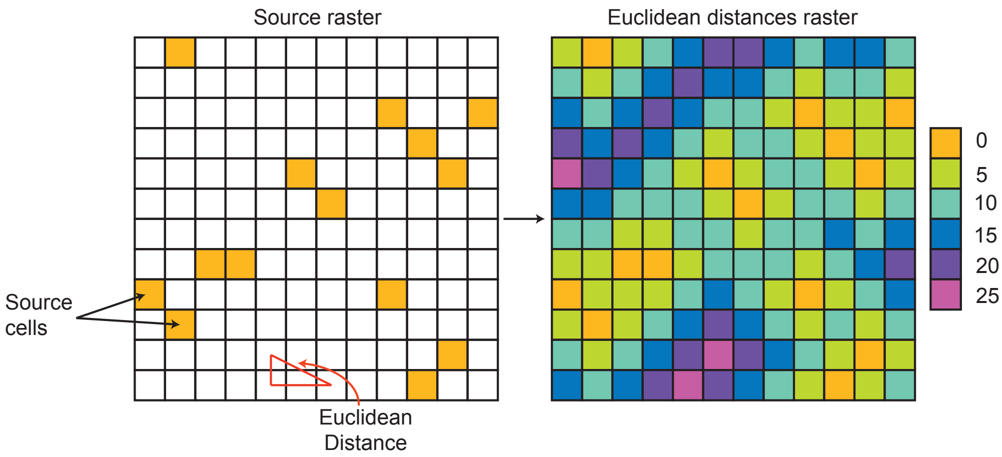

FIGURE 5.4: Euclidean distance map calculation. Source: image modified from the ArcGIS Help on line.

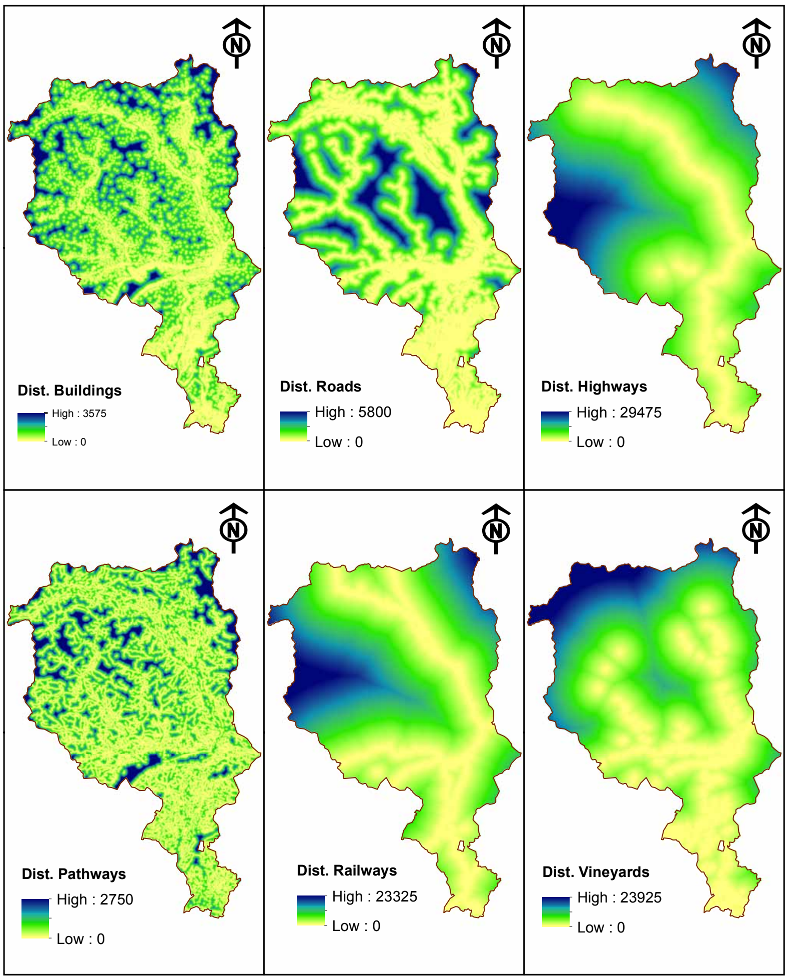

FIGURE 5.5: Maps of distances to the human infrastructures: a) Distance to buildings, b) distance to roads, c) distance to highways, d) distance to pathways, e) distance to railways and f) distance to vineyards.

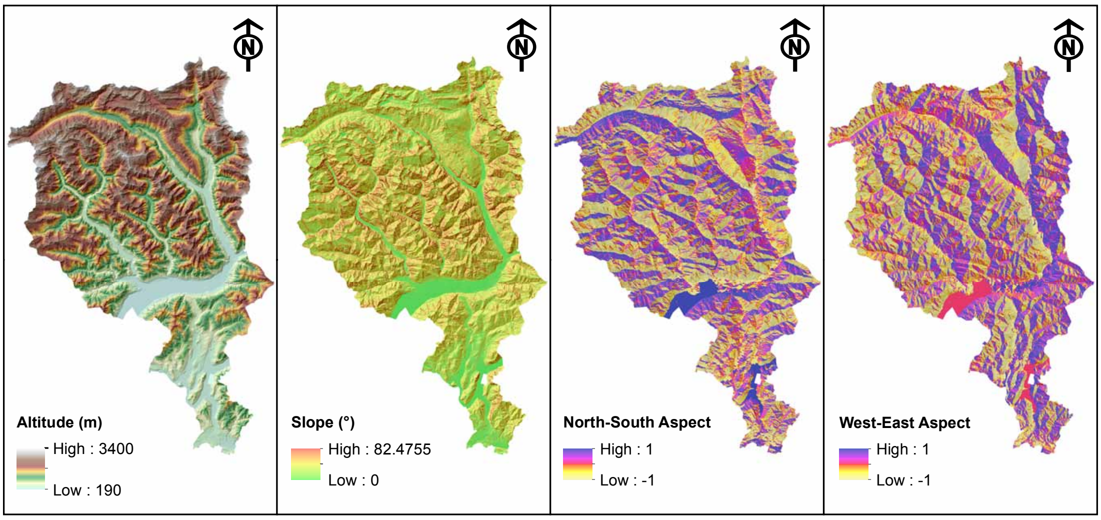

FIGURE 5.6: Topographic characteristics in Canton of Ticino: a) altitude, b) slope, c) aspect North-South direction and d) aspect West-East direction.

On the other hand, the RF algorithm requires information about "positive" and "negative" labels, that is, information regarding where the events take place (positive labels) as well as the information where events do not take place (negative labels). This is not a straightforward task due to the complexity of the forest fire phenomenon. We only know the points where fire has broke out, wherefore we do not know the group of variables conditioning the absence of fire. To overcome this difficulty, and in collaboration with Michael Leuenberger, we developed a procedure that defines zones where no fires take place by means of the k-means clustering method. This machine learning technique aims at partitioning n observations into k clusters in which each event belongs to the cluster with the nearest mean. This method allowed us identifying zones across the studied area with similar characteristics of presence of fires. For the other zones, we assume that the presence of fires is weak, though, we used these areas to randomly distribute the same number of points as the anthropogenic forest fire ignition dataset (i.e. 672). These points were classified as the negative class, meaning an absence of fire. For this negative labelled points, we also extracted all the information related to the euclidean distances to the human infrastructures, and then, we assembled them with the real anthropogenic forest fire data in the same database. Thus, this data, holding a total number of 1,344 points (672 indicating the presence of fire and the other 672 points indicating the absence of fire), was used to evaluate the predictor anthropogenic components most influencing the forest fire occurrences in Canton of Ticino (see Table 5.2 and Figure 5.7).

5.2.2 Data for fire occurrence susceptibility mapping

As presented in the introduction, one objective of this Chapter is the elaboration of a fire ignition susceptibility map. Susceptibility refers to the estimation of the probability that an event occurs in a specific area without considering an absolute temporal scale.

TABLE 5.2: Dataset 1 comprising the direct anthropogenic forest fire ignition points (positive labels) and the negative class points holding information regarding their anthropogenic features.

Variable name	Variable type
1. Distance to Buildings	Human infrastructure
2. Distance to Highways	
3. Distance to Roads	
4. Distance to Pathways	
5. Distance to Railways	
6. Distance to Vineyards	Land cover
7. Fire	Positive label (Forest fire) Negative label (No fire)

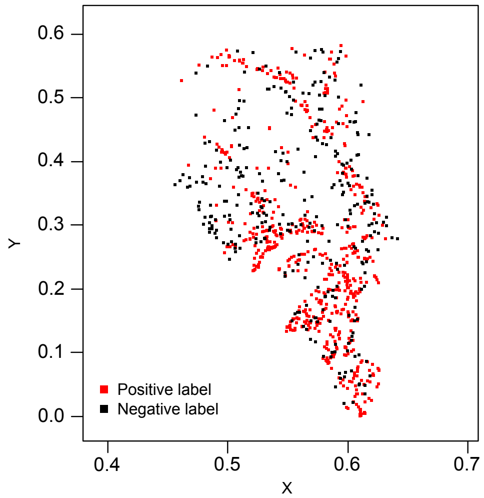

FIGURE 5.7: Total point database used in the Random Forests for WUI definition. The positive labelled points (in red) correspond to the observed anthropogenic forest fire ignition points and the negative labelled points (in black) correspond to the absence of fire.

For this type of map, two datasets were generated. The first dataset (1,344), similar to dataset 1 for the WUI definition, comprises all points of both anthropogenic-induced fires and negative class (no fire), with additional information regarding the topographic features (altitude, slope, North-South aspect and West-East aspect) (see Table 5.3).

The second dataset corresponds to the grid where the probability of fire occurrence will be estimated. This dataset consists on a regular grid of 25 meters resolution covering the forest cover of Ticino. This database contains only the information regarding each topographic feature at the location of each grid point and the minimum distances of each grid point to the anthropogenic features (see Table 5.4). It does not hold information about fire occurrence.

TABLE 5.3: Dataset 2 comprising the direct anthropogenic forest fire ignition points (positive label) and the negative class points holding information regarding their anthropogenic and topographic features.

Variable name	**Variable type**
1. Altitude 2. Slope 3. North-South Aspect 4. West-East Aspect	Topographic features
5. Distance to Buildings 6. Distance to Highways 7. Distance to Roads 8. Distance to Pathways 9. Distance to Railways	Human infrastructure
10. Distance to Vineyards	Land cover
11. Fire	Positive label (Forest fire) Negative label (No fire)

TABLE 5.4: Dataset 3 comprising the grid points where the prediction of the probability of fire occurrence is estimated.

Variable name	**Variable type**
1. Altitude 2. Slope 3. North-South Aspect 4. West-East Aspect	Topographic features
5. Distance to Buildings 6. Distance to Highways 7. Distance to Roads 8. Distance to Pathways 9. Distance to Railways	Human infrastructure
10. Distance to Vineyards	Land cover

5.3 Wildland-urban interface (WUI)

As explained in Chapter 1, in the wild vegetative areas with strong anthropogenic pressure emerges a critical zone known as the wildland urban interface (WUI). This term is broadly used to indicating areas where human infrastructures interact or intermingle with wildland/forest areas [3, 59, 130, 231, 232, 263, 280]. Many environmental/ecological problems are associated to the WUI; but the most relevant issues, particularly in densely populated areas, are those related to anthropogenic-induced fire hazard and management [162, 233]. This coexistence enhances both anthropogenic ignition sources and flammable fuels. Furthermore, the growing trend of the WUI and global climate change effects may even worsening the situation in the near future [68, 154, 188, 280]. Therefore, lately, many researches are engaged to the WUI problem, and it is currently an important subject exclusively related to the forest fires phenomenon.

In fire prone regions around the world, i.e. the Mediterranean areas, the WUI designates principally zones of significant infrastructure threats due to the exposition to burning vegetation [68, 161] placing human lives and property at a high level the risk. Conversely, in the Alpine region, the WUI represents susceptible areas where fire is essentially triggered. In this region, most human settlement and infrastructures are located on the valleys or on slope terraces, and houses are usually surrounded by cultivated and open areas with important exposure to sunshine [68]. In such conditions, although fire ignitions are highly clustered in areas adjacent to urban infrastructures [67, 305, 324], areas with significant fuel loads (i.e. forests) hardly conduct fires toward human infrastructures [68].

Few methods exist for mapping WUI and they mostly differ from their operational approaches and implementations [221, 264]. Nevertheless, they all coincide with using the same WUI components: the urban element represented by human infrastructures and activities, the wildland element constituted by the burnable vegetation (forest) and the interaction of the former two elements which represents the interface generally expressed as a buffer distance.

Thus, in order to characterise the WUI in Canton of Ticino, we designed a systematic and flexible methodology based on a machine learning algorithm, a statistical analysis and a GIS routine (see Figure 5.8) considering the three main WUI components, as follows:

1. First we performed a RF analysis to select the anthropogenic variables that most influence the occurrence of fire ignition. This analysis allowed defining the elements characterising the urban component of the WUI.

2. Then, with the forest fire ignition points and the anthropogenic variables selected from the RF analysis, we carried out a statistical analysis to define the buffer distance that designates the interface area between the forest and the human activities. Whereas the wildland component was defined as the forest cover.

3. And finally, using a GIS routine created in ModelBuilder (ArcGIS), we automatically mapped the WUI by intersecting the forest cover with the buffer areas around the selected anthropogenic variables.

5.3.1 Anthropogenic variable selection by means of the Random Forests method

The first step consists on defining the anthropogenic features that most contribute to the ignition of human-caused forest fires in this Swiss Alpine domain, Canton of Ticino. As revealed in the clustering analysis in Chapter 4, these underlying factors may act differently at different scales. Thus, it is necessary evaluating the relative importance (contribution) of each variable to the fire ignition at local scales.

Different approaches exist to evaluate the importance of a high number of variables in the prediction of environmental events. Some examples are the Multiple Linear Regression [105, 255, 270], the classification and regression trees [35], the geographically

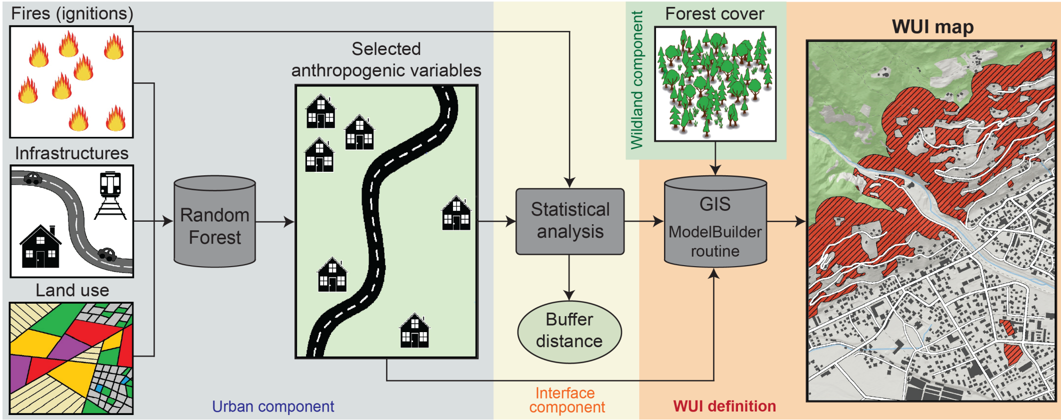

FIGURE 5.8: Methodology for WUI characterisation and mapping. The urban component (blue box) is defined by using the Random Forests algorithm to select the anthropogenic variables most influencing the distribution of human-caused fire ignitions. The wildland component (green box) is defined as the forest cover. The interface component (yellow box) is defined by buffer distances estimated with the statistical analysis. And, the WUI definition (orange box) is carried out through the application of a GIS ModelBuilder routine. Source: schema modified from Conedera et al. [68].

weighted regression [102], the multivariate adaptive regression splines [106], the maximum entropy [146], among others. Although, the advantages of these models and techniques to deal with high dimensionality data and to consider multivariate relationships between the predictor variables, they also present gaps and drawbacks that hinder the interpretation or reliability of the possible results, particularly in the case of complex phenomena such as the forest fires. For instance, some of these models do not deal with the nonstationary relationships between the variables, or do not provide exhaustive outputs indicating the fire probability [7].

Among the machine learning methods, the non-parametric technique named the Random Forests (RF) has markedly proved to be highly capable to deal with correlated variables in multi-dimensional data (high number of variables), particularly, in complex regions as the Alpine environment [68, 191]. This method is a classification and regression technique based on an ensemble of decision trees that provides a measure of the importance of variables based on a permutation test [68]. It overcomes the problem of instability in using single classification trees, resulting in higher prediction accuracy [182, 269].

For the assessment of the importance and the related prediction power of each single studied variable, the basic idea of this method relies on the fact that if a variable is important to the prediction of the studied events and if such variable is removed from the analysis, then the accuracy of the classification is degraded; otherwise, if the variable is not important, the classification accuracy does not decay when the variable is dismissed from the estimation [68]. RF model works as follows:

1. We first separate the data into two datasets: the training set holding about 70% of the observed events (see Table 5.2) and the testing set with the 30% remained

events (see Figure 5.9). The testing set is used to evaluate the quality of the prediction, even if RF also provides an assessment of the prediction error.

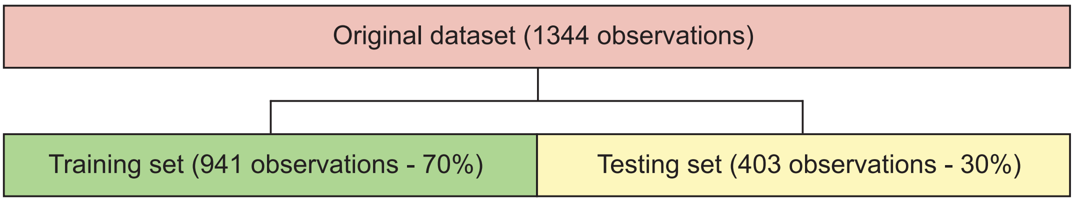

FIGURE 5.9: The dataset splitting into training (70 % of the original dataset) and testing datasets (30 % of the original dataset).

2. Now, using only the training set, RF generates a great number of subsets (inBag and out-of-bag subsets). The inBag subsets, also named bootstrap subsets, are generated through the bootstrap technique where the observed events in the training dataset is iteratively resampled several times. The out-of-bag (OOB) samples are obtained by keeping the observations left out during the resampling procedure (about one-third of the overall dataset, see Figure 5.10). This allows improving the classifier's performance because the OOB observations are not used in the fitting of the trees, thus the OOB's estimates are essentially cross-validated accuracy estimates [75]. For the study case, we chose to generate 1000 inBag sets and, hence, 1000 OOB sets were also created.

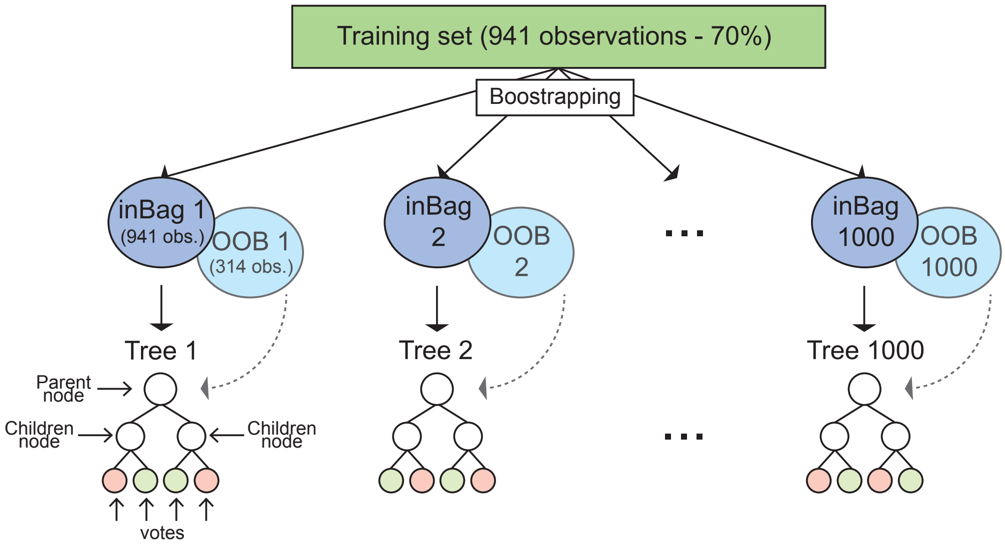

FIGURE 5.10: Random Forests procedure for boostrapping, tree generation and classification: the training dataset is iteratively resampled several times with replacement (boostrapping technique) and creates the inBag subsets. The data that have been left out in the resampling procedure are retained and named the out-of-bag (OOB) subsets. Then for each inBag subset a decision tree is grown by splitting each parent node into two children nodes by means of using the best eligible splitter from a random selection of the predictor variables. Each tree cats a vote at its terminal nodes.

3. Then, for each inBag subset a decision tree is created (see Figure 5.10). The decision tree is grown-down by splitting a parent node into two children nodes. The splitting is carried out by applying the Gini index which allows choosing the best splitter among some randomly selected predictor variables at each parent node. For each training dataset, this index measures the inequality (or impurity, diversity) of the data D as [135]:

$$Gini(D) = 1 - \sum_{i=1}^{m} p_i^2 \tag{5.1}$$

where p_i is the probability that a sample in D belongs to class C_i (estimated by $\mid C_{i,D} \mid / \mid D \mid$) [135]. The sum is computed over m classes; which in the case of this study, it corresponds to two classes ($m = 2$): Positive label (Forest fire) or negative label (no fire). The Gini index considers a binary split for each attribute. A weighted sum of the impurity of each resulting partition is computed. For example, if a binary split on attribute A partitions D into D_1 and D_2, the Gini index of D given that partitioning is [135]:

$$Gini_A(D) = \frac{\mid D_1 \mid}{\mid D \mid} Gini(D_1) + \frac{\mid D_2 \mid}{\mid D \mid} Gini(D_2) \tag{5.2}$$

This can be generalised as:

$$Gini_A(D) = \sum_{i=1}^{k} \frac{\mid D_i \mid}{\mid D \mid} Gini(D_i) \tag{5.3}$$

where k is the number of children nodes (in our case $k = 2$). This provides the quality of the split at the node.

For attributes with more than two values, subsets of values are considered. For a discrete-valued attribute, the subset that gives the minimum Gini index for that attribute is selected as its splitting subset. For continuous-valued attributes, each possible split-point must be considered: the midpoint between two adjacent sorted values is taken as a possible split-point. The point giving the minimum Gini index for a given continuous-valued attribute is taken as the split-point of that attribute. The reduction in impurity (diversity) that would be incurred by a binary split on a discrete- or continuous-valued attribute A is:

$$\Delta Gini(A) = Gini(D) - Gini_A(D) \tag{5.4}$$

The attribute that has the minimum Gini index (that is, the one maximizing the reduction in impurity) is selected as the splitting criterion. Every time a node is split, the Gini value for the two descendant nodes is less than the parent node. This binary partitioning is performed until further splitting no longer decreases the Gini index [35, 75]. A great variety of classification trees are grown returning each different results.

4. Then, the OOB subsets are passed down the trees. This allows RF estimating the prediction error rate (ER), also known as misclassification rate; where the predicted class estimated with the OOB observations (majority vote among all trees) is compared to the true class of the original set. That is, the ER is the proportion of times that an OOB prediction is not accurate (see Equation 5.5).

$$ER(\%) = 100 * \frac{1}{N} \sum_{i=1}^{N} E_i \tag{5.5}$$

   where N is the total number of observed points and E_i the error of the prediction of the ith point. For classification, E_i corresponds to 1 if the class of the ith point is voted (among all trees) different to the true class. This error rate has proven to be quite accurate and unbiased given the great number of trees grown [169], and it is useful for assessing the performance of the prediction.

5. Finally, RF provides a measure of the importance of the predictor variables. This is done by looking at how much prediction error increases when the OOB data for one variable is permuted while all others are left unchanged. That is, for each variable, the values in the OOB observations are randomly permuted, and then the modified OOB subsets are passed down the decision trees to get new predictions [68, 75, 169, 182] (See Figure 5.11). The assessment of the importance of each variable is carried out by estimating the difference between the misclassification rate for the OOB observations and the misclassification rate of the modified OOB subset and averaging over all trees (see Equation 5.6).

$$I_m = \frac{1}{T_k} \sum_{k=1}^{T_k} (E_k^m - E_k) \tag{5.6}$$

   where T_k is the total number of trees, m the analysed variable, E_k the misclassification rate of the OOB subset in the kth tree, and E_k^m the misclassification rate after permuting values of the mth variable in the OOB subset in the kth tree. This measure, I_m, is given as the percent increase of the ER (% IncMSE), where the higher the %IncMSE, the higher the importance of the corresponding variable [68].

The RF analysis was carried out, in collaboration with our colleague Michael Leuenberger, using the "randomForest" package in R statistical software [230]. The participation of the author of this Thesis in this analysis consisted in the preparation of the input data and the interpretation of the results provided by the algorithm. Nevertheless, we considered important to show the development and functioning the RF procedure.

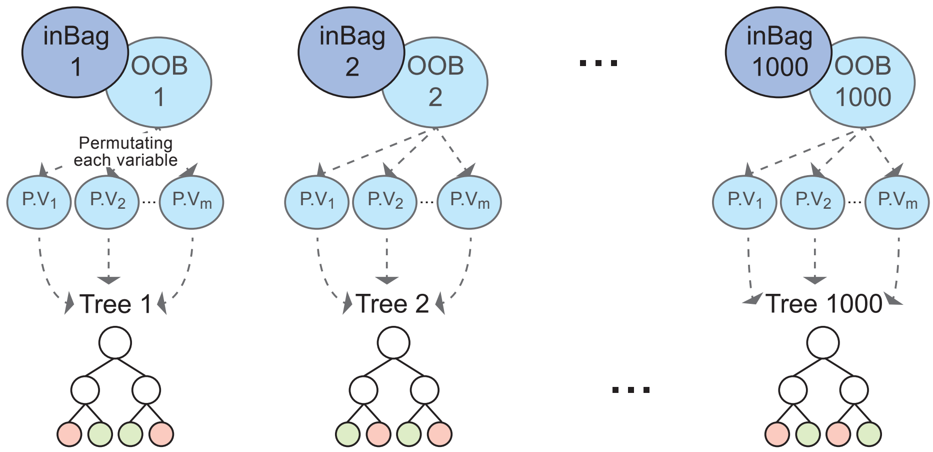

FIGURE 5.11: Random Forests procedure for variable importance estimation: one by one of the variables in the OOB observations are randomly permuted. The modified OOB data are passed down the trees to get new predictions.

The results of the RF analysis are presented in Figure 5.12. Based on the anthropogenic variables used in the analysis, RF revealed that buildings are substantially the most influencing features in the ignition occurrence of anthropogenic-caused fires. This indicates that the fire occurrence in Ticino is highly shaped by the local permanent presence of human activities. Following the importance variable rank, roads are the second most relevant factor manifesting the importance of the infrastructures providing human accessibility to wildland areas. This increases the probability of anthropogenic-caused fires. The third important factor exposed by the RF is the vineyard crops. This can be due to the close proximity of these cultivated lands to wild-vegetation areas. Nevertheless, their lower importance regarding roads and buildings can be explained as a response of the recent improvement in fire prevention measures implemented in the Canton [220] which has contributed to reduce the impact of agricultural activities as a fire-ignition source [68]. Likewise is the situation with the railways where the construction of protection infrastructures in the steepest sectors of the Gotthard line has been implemented since 1975 [61]. These measures have essentially reduced the number of fires caused by this activity, nevertheless, in rare cases, fire can break out by sparks of the rail transport.

The low importance of the highways can be explained by the fact that these infrastructures do not provide accessibility to wildland areas; moreover, the preventive measures against fire in such infrastructures contribute avoiding both fire ignition and fire propagation of existing fires [61]. Regarding the pathways, they are the less important human infrastructures in the distribution of fire ignition. Pathways are composed by very small non-car-passable roads of less than 2 metres wide. Probably, in the forest areas, these structures are less frequented by arsoners than e.g. the roads.

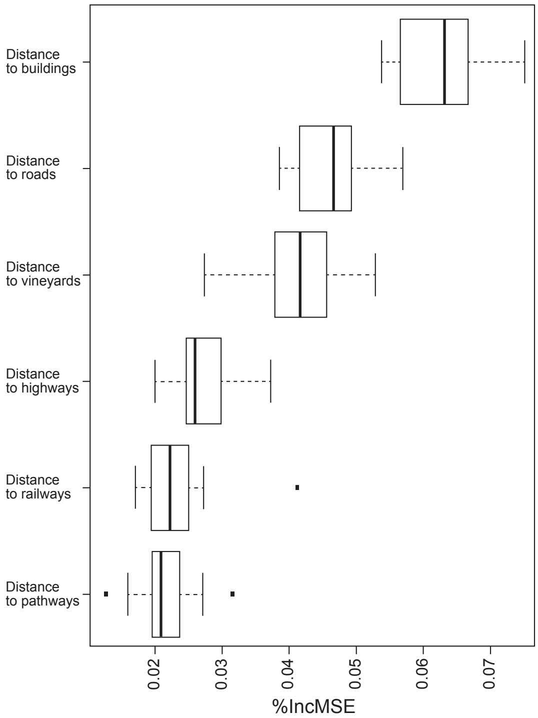

FIGURE 5.12: Boxplots of the variable importance obtained with the Random Forests algorithm. The percent increase of the mean squared error (%IncMSE) indicates the contribution of each anthropogenic variable to the occurrence of anthropogenic-ignited forest fires. Source: image taken from Conedera et al. [68].

5.3.2 Statistical analysis for buffer distance definition

The second step for WUI definition is the assessment of the size of the interface component, that is, the areas where the interaction between human activities (infrastructures) and wildland vegetation is importantly active. We characterise this component as a buffer distance from the anthropogenic infrastructures (in this case the two most relevant anthropogenic infrastructures detected by the RF algorithm: buildings and roads) to the forest cover.

Thus, we needed an analysis providing information about the local neighbourhood of the forest fires around the human activities. To this end, we performed the nearest neighbour distance function analysis in order to estimate the Euclidean distances at which fire ignition points arise from the two selected anthropogenic infrastructures. For this, we decided to use a classical tool in point process statistics which is a generalisation of the nearest neighbour distance function, $G_{I,J}(r)$, allowing estimating the (cumulative) distribution of the distances r from the typical point of pattern I to the nearest neighbour point of pattern J [13, 74, 89, 143, 316]. The values of this function range from zero (scale at which no nearest neighbours have been encountered) to unity (scale at which all points have a nearest neighbour) [316].

For the estimation of this function, we used the marked nearest neighbour distance function, Gmulti included in the "spatstat" package [13] of the R software environment

[230]. The results are presented in Figure 5.13.

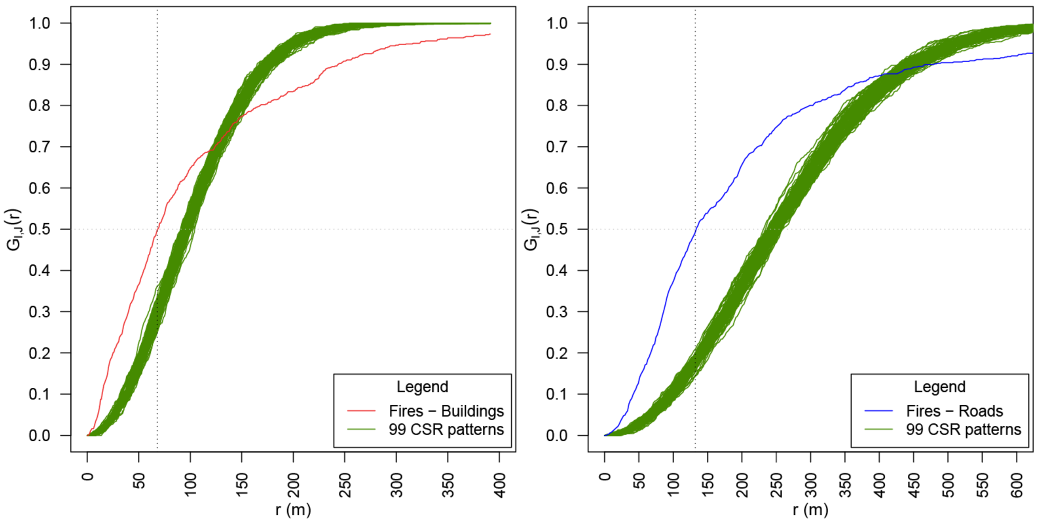

FIGURE 5.13: The nearest neighbour distance distribution function for: the anthropogenic forest fire ignition points to the nearest Buildings (left, red line); and the anthropogenic forest fire ignition points to the nearest Roads (right, blue line). The dashed green lines correspond to the probability distance distribution of 99 simulated random patterns.

The two curves show the probability distribution of the distances from the forest fire ignition points to the nearest or closest neighbouring buildings (Figure 5.13 red line in the left) or roads (Figure 5.13 blue line in the right). The differences between the empirical curves (fires to buildings and fires to roads) and the 99 random simulations allow rejecting the null hypothesis of independence. It is also possible depicting that forest fires arise closer to buildings than to roads (at shortest distances forest fires encounter more buildings than roads) confirming the results from the RF analysis. Looking at the distances to buildings, 50% of the forest fire ignition points encountered buildings at a distance of 68 m, while the distance to roads is about 132 m. Depending on the fire management purposes, one can use both distances to define the WUI; however, for practical reasons, we decided to average these distances in order to use a buffer distance that fits to the two infrastructure distributions. Therefore, the buffer distance estimated for the WUI definition in Canton of Ticino where 50% of the fires take place is assumed equal to 100 m.

5.3.3 The ModelBuilder routine

Finally, the third step of the methodology for the WUI definition was the construction of a GIS routine that both automatically maps and characterises the WUI in Canton of Ticino. This model routine was developed in ModelBuilder from the ArcGIS software environment [96]. ModelBuilder is a visual programming language used to create, to edit and to manage models. A model can be thought of a work-flow that strings together sequences of geoprocessing tools by feeding the output of one tool into another tool as input [96]. Figure 5.14 shows the ModelBuilder routine created for mapping WUI in Canton of Ticino. It can also be used and adapted to map the WUI of other study areas and with other input features and parameters.

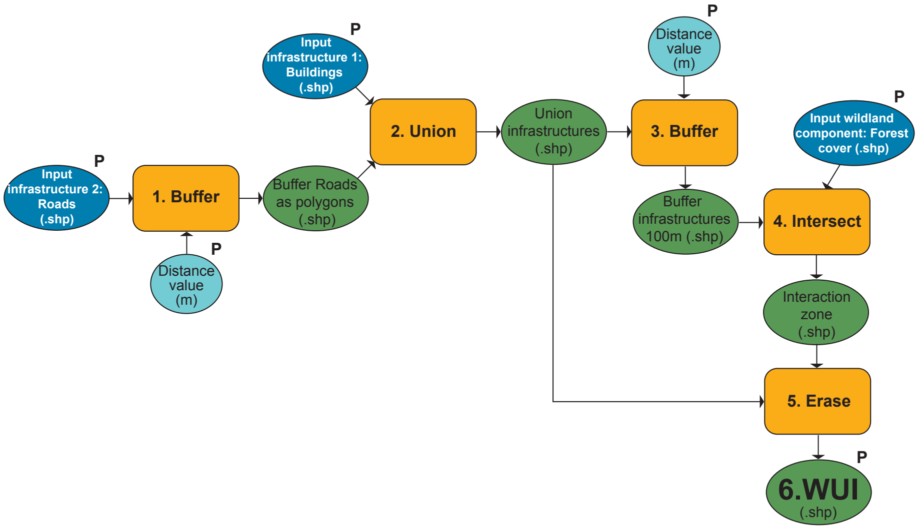

FIGURE 5.14: The ModelBuilder routine in ArcGIS 10.2 for the WUI definition in Canton of Ticino. Input data is presented in dark blue, ArcGIS geoprocessing tools in orange, tool's parameters in light blue, and output data in green. The input and the output data are in shapefile format (.shp). Source: schema modified from Conedera et al. [68].

An interesting option in ModelBuilder is the possibility to generate the model's dialog box for the final user, and thereby presenting the model as an ArcGIS tool. The model's dialog box can be personalised to facilitate final users to use the model exposing only the options and parameters that can/should be managed and modified. Figure 5.15 displays the WUI model's dialog box to be manipulated for the user. Here, it is possible to see that only few parameters of the model are displayed, and that the geoprocessing tools are left out as internal procedures beyond the reach of the model users. In this way, we can protect the model to be modified by undesirable manipulations.

Regarding the structure of the WUI model, it consists on integrating all the elements and parameters of the WUI components defined in the previous analyses to both build and map the WUI in the entire study region. The following describes the different steps of the model (see Figures 5.14 and 5.16):

1. **Buffer:** this procedure takes the half of the width value of the roads and creates polygon buffers around these elements (Figure 5.14 - 1.Buffer, result Figure 5.16 - 1). The output is a shapefile of all the roads in a polygon format. Since we are working at local scales, taking into account the real space occupied by the roads is very important for the WUI characterisation.

2. **Union:** in this step, the two selected human infrastructures (buildings and roads) are assemble together in the same shapefile (Figure 5.14 - 2.Union, result Figure

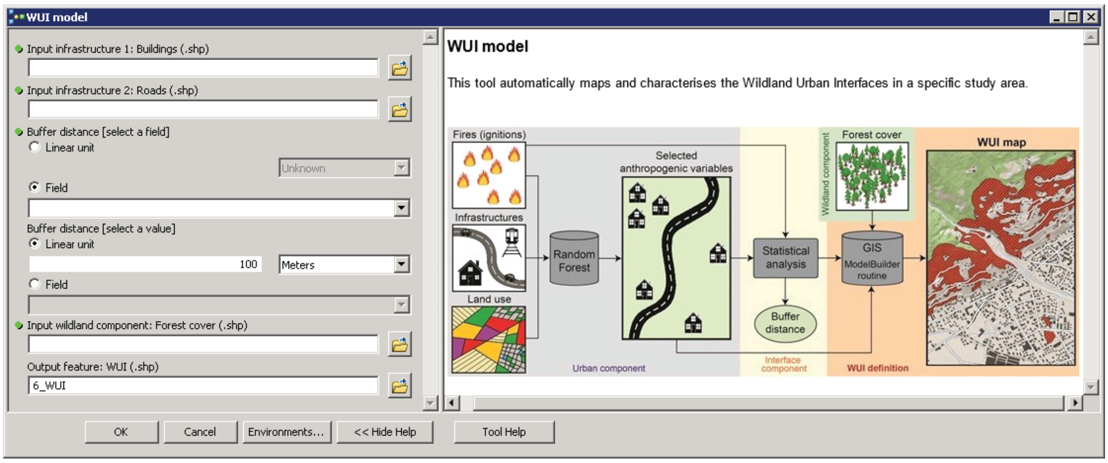

FIGURE 5.15: The model's dialog box created in ModelBuilder routine in ArcGIS 10.2 for the WUI definition in Canton of Ticino.

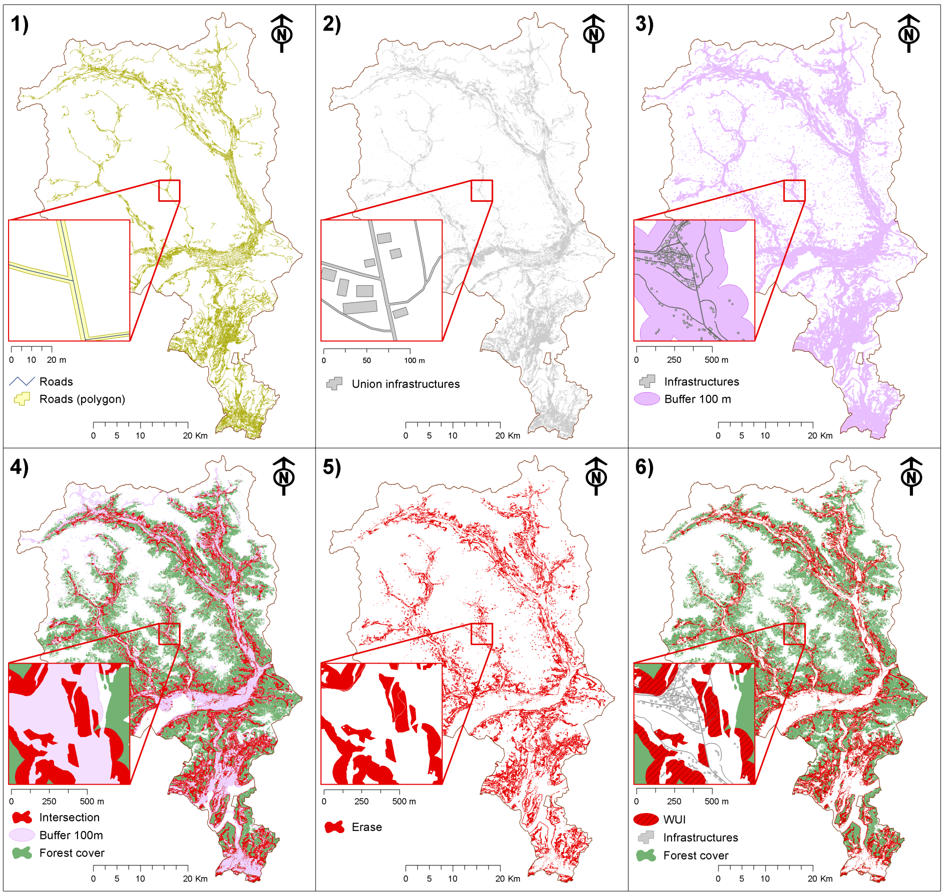

FIGURE 5.16: Procedure of the WUI model to estimate and characterise the WUI: 1) Converting roads into polygons. 2) Combining the buildings and roads infrastructures in the same shapefile. 3) Creating buffers of 100 m around infrastructures. 4) Intersecting the infrastructure-buffer areas with the forest cover. 5) Eliminating the infrastructures from the intersection areas. 6) Mapping WUI.

5.16 - 2). In order to carry out this task, both layers must be of the same format (e.g. polygon). This shapefile constitutes the infrastructure component of the WUI.

3. **Buffer:** polygon buffers are built around the two human infrastructures using the distance value defined in the statistical analysis (Figure 5.14 - 3.Buffer, result Figure 5.16 - 3). For the case of Canton of Ticino, this distance value was estimated at 100 metres. These polygons constitute the active interaction areas of humans activities and it can be considered as a human interface.

4. **Intersect:** a geometric intersection between the human interface and the forest cover is computed. The forest cover represents the wild-vegetation component of the WUI (Figure 5.14 - 4.Intersect, result Figure 5.16 - 4). Only the overlapping areas in both layers (human interface and forest cover) are retained. These areas correspond to the WUI.

5. **Erase:** since the human infrastructures themselves are not part of the WUI, it is necessary to eliminate them from the WUI polygons (Figure 5.14 - 5.Erase, result Figure 5.16 - 5). This is why we apply the erase geoprocessing tool.

6. **WUI:** finally, it is possible to both map and characterise the WUI surfaces for the study area (Figure 5.14 - 6.WUI, result Figure 5.16 - 6). These interfaces basically can be viewed as the interaction zones between the fire-inducing anthropogenic activities and the wildland fuel in this mountainous region [68].

Figure 5.17 shows the WUI map in Ticino created with our methodology (left) and a visualisation of the WUI areas in Google Earth (right). WUI areas in Ticino, estimated as the interface where the 50% of fires take place, count for about 31,000 ha, and are characterised by forest areas of 100 metres around drivable roads and accessible buildings. Therefore, these areas are mostly located in the low lands along the valleys. Keep in mind that the definition of the WUI may change depending on the parameters defined in the RF and the statistical analyses and on the existing conditions of the study area.

For forest management interest, the defined WUI allows reducing the intervention of forest areas for fire risk and fuel measures planning. That is, instead of considering the 130,000 ha of forested areas in Ticino, forest and fire managers can, in the first instance, reduce the operating zones to the WUI areas to concentrate financial and preventative technical measures. In this regard, WUI areas may depend on the forest management purposes, where interfaces including different percentage of fire ignitions, for instance, 30% or 70% of fire events, can be considered. Consequently, the size of the WUI areas may vary. Nevertheless, they will remain lower than the total forested areas. On the other hand, the WUI areas together with other fire-relevant information such as detailed fire risk maps [67], also facilitate the identification of fire ignition hot spots and other fire-related risk regions [68].

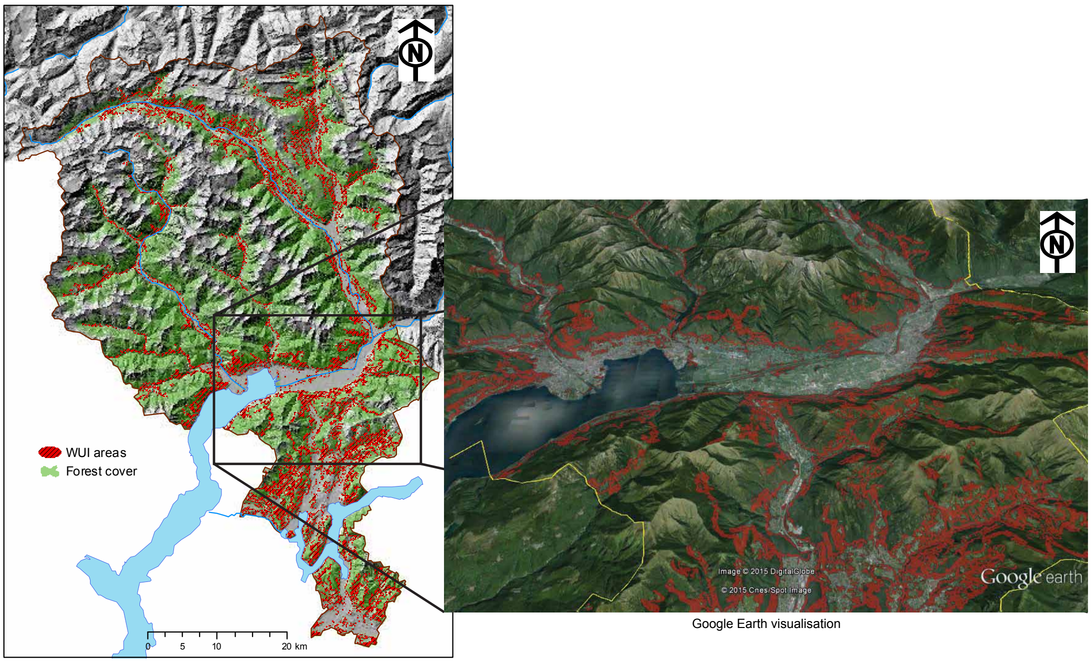

FIGURE 5.17: WUI final mapping (left). A visualisation of the WUI in Ticino in Google Earth: a zoom to the zone of Bellinzona city (right).

5.4 Susceptibility mapping using the Random Forests

The implementation of the RF algorithm not only led us to identifying and selecting the most relevant anthropogenic infrastructures for the human component of the WUI, but it also allowed us generating a map of fire susceptibility by estimating the probability of fire-ignition occurrence in specific areas without considering an absolute temporal scale. This type of mapping is called “susceptibiliy mapping”.

This study was entirely developed applying the random forest machine learning algorithm. It deals with the same data and study case as in the former Section 5.3 in Canton of Ticino, but, in addition, we also added information about the topographic features of fires. As for the random forests part in the former Section 5.3, we count with the collaboration of our colleague Michael Leuenberger. This thesis author’s participation on this study was the preparation and geoprocessing of data to run RF and the interpretation of the RF results.

The fundamental scientific problem conducting this research is the application of a machine learning method to analyse and to model forest fire patterns in a high dimensional input feature space (e.g. great number of predictor variables). Thereby, we intended to considering several number of ground environmental factors influencing the ignition distribution of forest fires in a complex region such as Canton of Ticino. For that two objectives were established: 1) estimating the importance of topographic and anthropogenic factors for predicting fire ignition occurrences by means of the RF algorithm, and 2) elaborating a susceptibility map.

The procedure for the RF algorithm is the same as explained in Subsection 5.3.1 for variable selection. Yet, this time we were not only interested in the variable selection but we also considered the classification resulted from the RF to make predictions of the fire occurrence on the forest cover in Ticino. The procedure is as following (see also Figure 5.18):

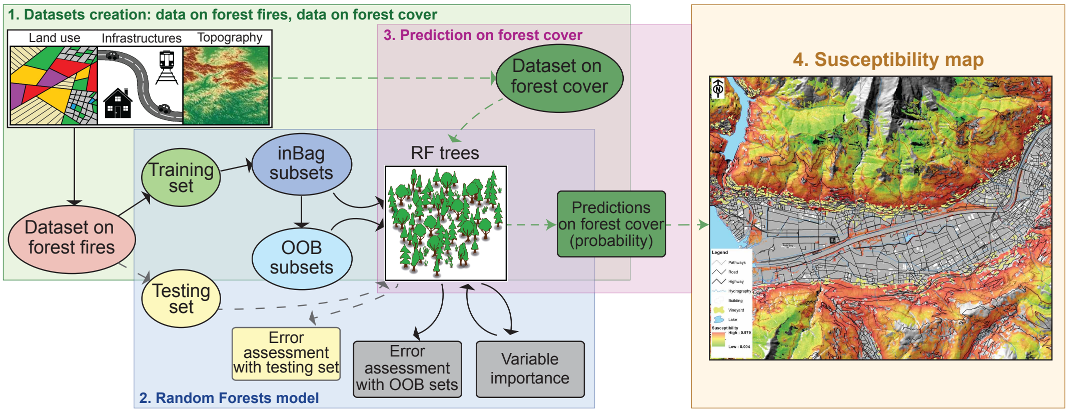

FIGURE 5.18: The methodology for the susceptibility mapping of fire occurrences in canton of Ticino: 1) Creation of the two datasets. 2) Generating the RF model by using dataset containing information on the forest fires (labelled points). 3) Running the RF model on the dataset containing the information on the forest cover (unlabelled points) in order to make predictions. And 4) Mapping the forest fire ignition susceptibility.

1. First we construct two datasets (see Figure 5.18 light green area): one dataset (Dataset 2 - Table 5.3) containing labelled points (we know the class for each point: Yes or No fires) used to build the RF model, and the second dataset (Dataset 3 - Table 5.4) of unlabelled points (e.g. the class for any point in the grid is unknown), where predictions are made based on the RF model that will be created with the first dataset.

2. With the training dataset, we performed a RF classification (see Figure 5.18, light blue area) as explained in Subsection 5.3.1. Then, we evaluated the precision of the RF model by analysing the misclassification provided by RF. We have also assessed the quality of the RF classification by running the model with the Testing dataset (see Figure 5.18 yellow ellipse). This allowed us to obtain a misclassification rate with data that was not used at all in the construction of the RF model. Furthermore, the algorithm also provides the variable importance measure which was used to display the predictor factors affecting the forest fire occurrences in Ticino.

3. An advantage of the RF is that we can also compute different indexes of accuracy such as the sensitivity which measures the proportion of true positive points that are correctly classified (see Equation 5.7). That is, it measures how much the RF model is able to classify positive points as positive instead of negative. This is

an important measure in environmental studies since it is crucial that a model classifies correctly, at least, the positive points, for instance, that a forest fire labelled point is not classified as "no fire".

$$Sensitivity = \frac{TP}{TP + FN} \tag{5.7}$$

where TP is the number of true positive points (positive points classified as positives) and FN the number of false negative points (the positive points classified as negative).

4. Next, using the RF model created with dataset 2, we passed dataset 3 (Table 5.4, dataset to make predictions on the forest cover grid) down each tree to estimate the prediction classes on the grid enveloping the forest cover in Ticino (see Figure 5.18 pink area). For prediction, a committee system, integrated in the RF algorithm, votes for the most popular class and assigns the predicted class (value) to the unlabelled points in the forest cover grid. The RF algorithm also provides a value of the probability of belonging to the positive class ("fire"), which we had retained as the probability of fire ignition (susceptibility estimation).

5. Finally, by means of GIS techniques, we mapped the resulting predictions which represent the susceptibility to fire ignition (see Figure 5.18 orange area). This is done by creating a table containing the geographical location of each point of the predicted grid and its corresponding probability value estimated by the RF. This table was then exported as a raster layer in ArcGIS software where the final fire susceptibility map was elaborated.

5.4.1 Results and mapping

The RF analysis was carried out using the "randomForest" package [169] in the R statistical software environment [230]. The RF algorithm was performed on the forest fire dataset of Canton of Ticino to make predictions on the probability of fire ignition. The assessment of the quality of the RF model is displayed in Table 5.5. This table presents the error rates obtained with the OOB observations (OOB ER), with the Testing dataset (Testing ER) and the Sensitivity measure.

The OOB ER directly estimated by the RF model is 21.28%. This is a good acceptable error value meaning that about 200 points out of the 941 points in the Training dataset were misclassified. That is, that for 200 points, the predicted class did not correspond to the true class (the label in the original dataset). Furthermore, when evaluating the accuracy of the RF model with the Testing data, that is, the 403 events that were not involved in creating the model, we obtained an error rate of 19.87%, quite similar to the OOB error rate. This is a good estimation because when labelled points were added to the model, the error did not increase. Additionally, the sensitivity measure indicates that 79% of the forest fires points were correctly predicted as "fire".

TABLE 5.5: Error assessment of the Random Forests model for the forest fire ignition occurrence in Canton of Ticino.

Index	Value
OOB ER (%)	21.281 (± 0.681)
Testing ER (%)	19.87 (± 0.821)
Sensitivity (%)	79.00 (± 0.01)

The variable importance provided by the RF algorithm for the forest fire occurrences in Ticino is presented in Figure 5.19

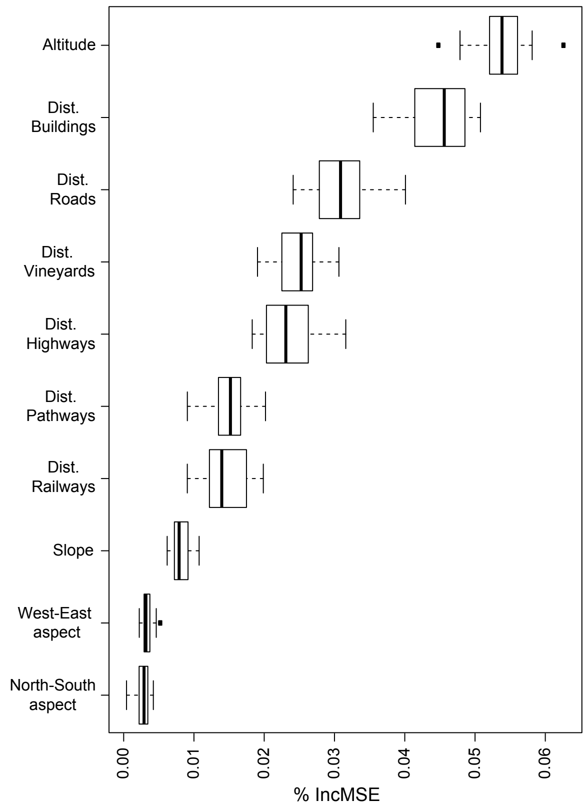

FIGURE 5.19: Variable importance provided by the Random Forests algorithm for the case of forest fire occurrences in Canton of Ticino using anthropogenic and topographic features.

Here we can see that the altitude is the most influencing predictor variable for fire ignition in Ticino. This is probably due to the fact that anthropogenic forest fires mostly occur at lower and mid-altitudes between 190–1000 m.a.s.l., which can be associated to the distribution of the urban areas in the Canton. Moreover, in the northern areas, the mountainous part of Ticino, the topography is characterised by steep slopes which limits the accessibility to higher forested areas. As a second most influencing predictor variable, we found the distance to buildings, followed by distance to roads, to vineyards, to highways, to pathways and to railways, slope, West-East aspect and North-South aspect.

As presented in the variable selection for WUI analysis, RF revealed once again that, among the anthropogenic infrastructures, buildings and roads are the most predisposing variables for forest fire ignition in Ticino. On the other hand, regarding the topographic features, RF reported that slope and aspect hardly contribute to the prediction of fire break-out.

To carry out predictions for the susceptibility mapping of fire ignition in Ticino, we did not select the most influencing predictor variables, instead, we used all the variables in the model and we proceeded to estimate the predictions. Figure 5.20 shows the fire ignition susceptibility map in Ticino (left) with a visualisation on Google Earth (right).

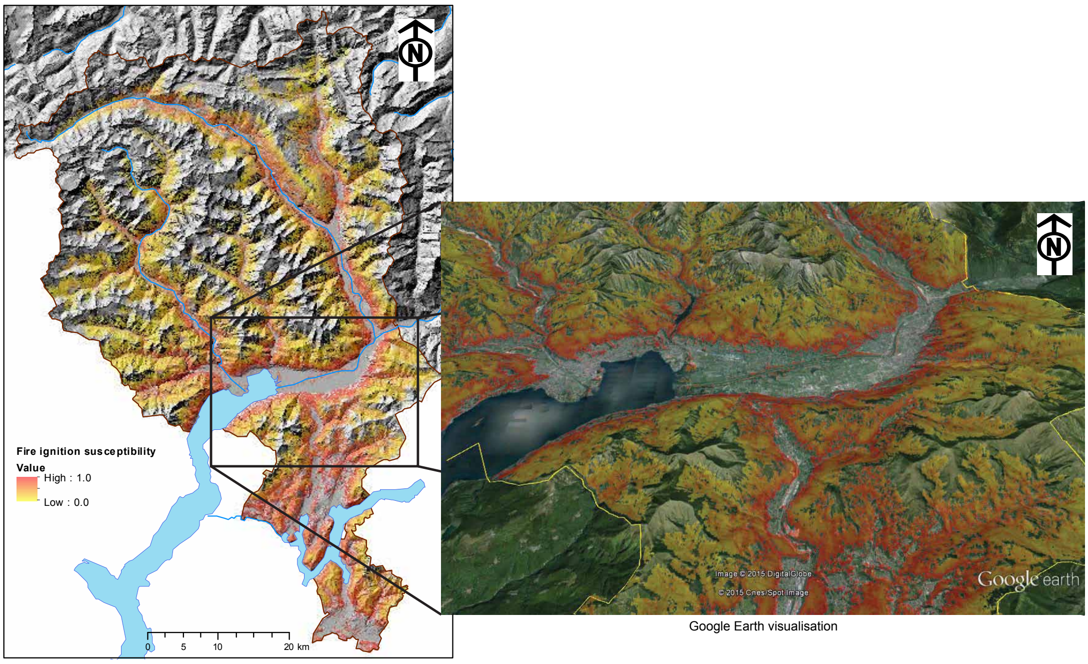

FIGURE 5.20: Susceptibility map of the forest fire ignition in Canton of Ticino.

As explained in Section 5.4, the prediction value corresponds to the probability of belonging to the fire class. This value is a percentage of the positive votes ("fire") over all trees. The higher the percentage of the positive votes, the higher the probability to have a fire event (Figure 5.20 red colour zones). Lower probabilities are given in yellow. This map can also be interpreted as the degree of predisposition to fire ignition. We can see that most of the higher susceptibility is located in the low areas along the valleys, close to the built-up areas and around the road system .

5.5 Conclusions

With the statistical analysis presented in Chapter 4, we provided support and knowledge about the direct impact of human activities in the distribution of anthropogenic fire ignitions which are characterised by a strong clustering behaviour around human settlements in a mountainous Swiss region such as the Canton of Ticino. Meanwhile, in this chapter, we confirmed that, among human activities, the main fire-ignition clustering attractors are motorised mobility and related easy access to settlements as shown by the Random Forest analysis [68]. This revealed that the major WUI issue in this mountainous region is the fire ignitions triggered mainly by human activities. Particularly, the role of drivable roads in providing access to wild-vegetation lands influences and enhances fire occurrence by negligence or criminal actions.

The two presented methodologies together with ArcGIS tools provide flexible, adaptable and reproducible decisional frameworks for WUI definition, WUI characterisation and susceptibility mapping in different environments. The two methodologies proved to be useful for planning and allocation of fire suppression resources and to support decision making.

Random forests evinced to be a powerful algorithm which can deal with the high complexity of the study area and with the high dimensionality of the data (a great variety of predictor variables). The variable importance measure provided valuable information to better understand the factors shaping the fire occurrence in the region. With this information, we can design different scenarios by selecting different number of variables that will be used to estimate predictions. As explained before in Subsection 5.4.1, we used all the predictor variables presented in the original database; nevertheless, we could have carried out predictions by only using, for instance, the two, three or five most influencing variables, and run the model to make predictions. This can be adapted according to the forest and fire managers purposes.

Moreover, the random forests method also showed to be a powerful and useful tool for fire ignition susceptibility mapping. We can improve this methodology by constructing larger input spaces and by carrying out analysis to optimise the negative class definition. Nevertheless, this technique can be used to extrapolate any predictor variable collected at sample locations across the landscape and to understand which predictors are driving the distribution with a high level of confidence. We, therefore, highly recommend this statistical modelling tool to be used in predictive environmental mapping.

All the tools presented in this Chapter should support managers for decision-making, to reduce the working geographical space by identifying the potential regions for fire susceptibility and WUI, and to select appropriate preventive plans, for instance, to allocate water points, to plan interventions for fuel management, and/or to simulate scenarios under different ground conditions to consider different fire-fighting strategies. It should also facilitate the revision of fire fighting concepts such as fire brigades organisation, infrastructure improvement, forestry measures and agricultural practices.

It is important to remind that both susceptibility maps and WUI maps need to be revised with a certain frame period accounting for the changes in urban growth and in land uses or any kind of ground conditions.

Chapter 6

Conclusions and future directions

6.1 Space-time pattern analysis

As discussed along Chapter 3, the statistical analyses are powerful techniques providing flexible solutions for both characterising and understanding the behaviour of the observed patterns as well as inferring about the underlying processes. The range of applications in which these methods are positively used is wide and will continue to expand. Nevertheless, their adaptations and practical applications in complex phenomena need to be further developed. In this context, this Thesis contributed in a better understanding of the methodological issues encountered when working and considering the complexity of the studied phenomenon, i.e. the heterogeneity of the embedded spatial/temporal space where the events take place, the variability of the clustering, and the multivariate nature of the phenomenon.

The application of this methodology also highlighted the importance of considering multiple scales and shows the sensitivity of the clustering measures for different point pattern distributions. From this perspective, the results emphasized the great power of this methodology as an advance tool towards real-world implementations for complex point processes. We should also point out the reformulation and adaptation of spatial clustering measures for temporal sequences analysis as it is the case of the Morisita index, the Box-counting method and the Rényi generalised dimensions. Besides, we must also say that all these methods remain useful in many other application domains (presented in subsection 1.3.5), though, this should not be adopted uncritically, i.e., hypotheses of underlying mechanisms relevant for a particular phenomenon, may not be relevant in other disciplines.

However, the classification of patterns as regular, random and clustered can be viewed as an over-simplification of the description of the event's structure, it is an useful approach in a early stage of the analysis [91], because patterns for which the hypothesis of randomness is not rejected, scarcely merit any further formal statistical analysis.

The methodology developed in this Thesis is a new comprehensive and powerful framework for the analysis of space/time point patterns that can be applied in different domains. It allows gaining deeper knowledge and understandings of a point pattern when applied to complex phenomena. As presented along this work, its practical use is a straightforward procedure allowing the characterisation of the degree of clustering of the phenomena structures. Nevertheless, although its intelligibility to be applicable to a wide variety of fields and, although, the efforts of the author of this Thesis in simplifying the calculations and performance of the clustering algorithms, the application of certain functions such as the Ripley's K-function or the space-time permutation scan statistics model can still result in time-computationally intensive tasks, taking up to several days or weeks, when facing huge databases. The computing time depends on a wide variety of variables such as the size of the input dataset, the number of spatial and temporal intervals, among others.

Furthermore, the introduction of the VD concept and its consideration into the clustering analysis empowered the interpretation of the real degree of clustering of the studied patterns. The results displayed and proved to be very distinct from the clustering values computed when considering only empty geometrical spaces (i.e. square) usually considered in the literature.

Simulations inside the VDs were generated using the Monte Carlo approach to generate random point patterns (Poisson distributions) to carry out "goodness-of-fit tests" for the point process model. The Poisson process is a very useful and simple point process, in particular for the estimation of intervals of confidence which are used to test the null hypothesis of CSR. The simulated pattens in each VD also allowed assessing the performance of the statistical methods and visualising the models providing a better understanding of the structures generated from the study event. We considered the practical aspects involved when working with finite datasets and their limitations imposed in the application of theoretical methods.

Consequently, the definition of the VDs in the practical implementation of point pattern analysis are of great importance and pertinent for the assessment and for the interpretation of the real clustering of the observed events. Indeed, theoretically, the values of the clustering measures for a randomly distributed patterns (i.e. Poisson distribution) are well known and defined (as explained in Section 3.4, i.e. for the Morisita index the value is 1, for the Box-counting and multifractal formalism (1 or 2 for 1- or 2-Dimension respectively), and 0 for the L-function). However, real-world distributions are subjected to a certain degree of variations related to both the finite number of points of the observed events and the constriction of the embedded space. And taking into account these constraints, one may likely find that the dimensionality of the point distribution is therefore decreased as well.

6.2 Application on forest fire occurrences

For the application of the methodology in the case of the forest fire occurrences in Canton of Ticino and in Portugal (presented in Chapter 4), theses analyses revealed valuable information of the structure of the spatio-temporal global and local clustering.

One of the main general findings is that anthropogenic and natural fire ignitions have distinctive spatio-temporal patterns and that they are essentially non-random processes. They depend on human and environmental variables that have also explicit patterns that vary in the space and in time. For anthropogenic-triggered fires this investigation provides very useful information concerning the characteristics of the clusters and their structural and temporal components, as well as for their importance in the definition of critical areas for fire mitigation such as the WUI areas and the susceptibility mapping. Similarly, the detection of lightning induced fire patterns (natural fires) 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 [66].

It is also important to distinguish between the two contexts of the clustering analyses on the case of forest fires in Canton of Ticino: one, is the exploratory analysis carried out through the application of the global clustering measures for testing hypotheses about the general distribution structure of the fire occurrences allowing inferring the role that different factors may play in causing the ignitions; and second, the predictive modelling developed with the local clustering measures (space-time permutation scan statistics model) to identify those areas that are most prone to ignitions. Both results can assist the Forest Service and policy makers in the identification of fires hotspots useful for a better definition of the infrastructural and organisational measures required to mitigate the fire risk in a specific region.

Another practical functionality of the proposed methodology to support decision-making is the incorporation of the model outputs into a geographical information system (GIS), to map and to identify fire-prone zones. Outcomes of the space-time permutation scan statistics model were integrated into a GIS environment allowing the exposition of the detected clusters. Likewise, the methodologies developed for the WUI characterisation and the susceptibility mapping benefited from these techniques for data preparation, pre-processing and results visualisation.

6.3 Contributions

The fundamental contributions of this Thesis can be summarised as follows:

- The development of a novel, flexible and comprehend methodology to both detect and analyse clusters in space, time and space-time point data.
- The adaptation of traditional statistical and fractal-based techniques for spatial and temporal analysis of point patterns.
- The introduction and the application of the validity domain concept in the point pattern analysis, and the development of corresponding clustering measures with relevant statistical tests.
- The application and the adaptation of the developed methodology in complex real phenomena as it is the case of the forest fire patterns.

- The performance of new analysis of forest fire occurrences using a generalised K-function for an inhomogeneous process, also within different validity domains.

- The application of the space-time permutation scan statistics to two different cases: 1) single event locations using the forest fire occurrences data in Canton of Ticino, and 2) aggregated data using the forest fire database from Portugal where the fire events are georeferenced at the smallest administrative unit (Parish).

- The application and the adaptation of the proposed point pattern analysis methodology in different environmental and socio-economic studies (for details refer to Subsection 1.3.5).

- The development of a novel, flexible and systematic methodology for the definition and mapping of the WUI in Canton of Ticino presented as a complex mountainous region.

- The development of a flexible methodology for the forest fire susceptibility mapping in Canton of Ticino.

- The development and customisation of the global clustering measures in R environment (an open source software for statistical computing and graphics).

6.4 Future perspectives

In global clustering measures. As presented along the Thesis, the proposed methodology offers a powerful and robust framework for the point pattern analysis of space-time data. Nonetheless, further developments can still be done taking into account other complexities of the real-world applications. In particular, further developments may consider the following significant points:

Functional measures

The global clustering measures can be adapted to detect the structure and dependences of the measures in a studied phenomena (i.e. burnt area for the case of forest fires) through the generalisation to functional measures. This may be conceptually interesting, since it forges a connexion between the support of the measures and the measures themselves: the degree of clustering of the support (point locations) is computed at different intensity thresholds of the studied phenomena (measurements) providing an insight into the spatial/temporal dependence of the measures [120].

Multivariate point processes

For instance, carrying out advance geostatistical analysis for multivariate analyses and anisotropic modelling and simulations of the point patterns.

In local clustering measures. Adaptation and application of the scan statistics for prospective analysis taking into account validity domains and irregular shapes of clusters. This task can follow the application of scan statistics for the retrospective analysis for the forest fires data as presented in Section 4.4.

In WUI definition and mapping. Crossing WUI maps with other fire-relevant maps to identify further fire ignition hot spots and/or other fire-related risk regions [68].

In susceptibility mapping. For fire ignition susceptibility we could further increase the input space by adding information about, for instance:

- Livestock units as a proxy for biomass removal amount of fuel.
- The proportion of livestock owners from permanent resident population as an indicator of the potential number of ignition sources.
- Population density maps.
- Vegetation type as a proxy for fuel distribution.
- Meteorological data.

Appendix A

Publications & Proceedings

Peer review articles:

1. [214] M.G. Pereira, L. Caramelo, **C. Vega Orozco**, R. Costa and M. Tonini. Space-time Clustering Analysis Performance of an Aggregated Dataset: The Case of Wildfires in Portugal. *Environmental Modelling & Software*, vol. 72: 239-249, **2015**. doi:10.1016/j.envsoft.2015.05.016.

2. [68] M. Conedera, M. Tonini, L. Oleggini, **C. Vega Orozco**, M. Leuenberger and G.B. Pezzatti. Geospatial Approach for Defining the Wildland-Urban Interface in the Alpine Environment. *Computers, Environment and Urban Systems*, vol. 52: 10-20, **2015**. doi:10.1016/j.compenvurbsys.2015.02.003.

3. [306] **C. Vega Orozco**, J. Golay and M. Kanevski. Multifractal Portrayal of the Swiss Population. *Cybergeo: European Journal of Geography [Online], Section: Systems, Modelling, Geostatistics*, vol. 714, **2015**. doi:10.4000/cybergeo.26829, url:http://cybergeo.revues.org/26829.

4. [120] J. Golay, M. Kanevski, **C. Vega Orozco** and M. Leuenberger. The Multipoint Morisita Index for the Analysis of Spatial Patterns. *Physica A: Statistical Mechanics and its Applications*, vol. 406: 191-202, **2014**. doi:10.1016/j.physa.2014.03.063.

5. [305] **C. Vega Orozco**, M. Tonini, M. Conedera and M. Kanevski. Cluster Recognition in Spatial- Temporal Sequences: The Case of Forest Fires. *Geoinformatica*, vol. 16(4): 653-673, **2012**. doi:10.1007/s10707-012-0161-z.

Articles in progress:

6. M. Tonini, M.G. Pereira, **C. Vega Orozco**. A Multivariate Cluster Analyses of Forest Fires: from Spatio-Temporal Point Events to Smoothed Density Maps.

Proceedings:

2015

1. M. Pereira, R. Costa, M. Tonini, **C. Vega Orozco** and J. Parente. Influence of the input database in detecting fire space-time clusters. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, **2015**.

2. M. Tonini, M. Pereira, **C. Vega Orozco** and J. Parente. Multivariate cluster analysis of forest fire events in Portugal. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, **2015**.

2013

3. R. Costa, M.G. Pereira, L. Caramelo, **C. Vega Orozco** and M. Kanevski. Assessing SaTScan Ability to Detect Space-Time Clusters in Wildfires. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 14055, **2013**.

4. J. Golay, M. Kanevski and **C. Vega Orozco**. The Multipoint Morisita Index for the Analysis of Geodemographic Data. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.

5. M. Leuenberger, M. Kanevski and **C. Vega Orozco**. Forest Fires in a Random Forest. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 3238, **2013**.

6. M. Leuenberger, **C. Vega Orozco**, M. Tonini and M. Kanevski. Random Forest for Susceptibility Mapping of Natural Hazards. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.

7. I. Penna, M. Tonini, **C. Vega Orozco**, C. Longchamp, M.H. Derron and M. Jaboyedoff. Rock Avalanche Occurrence in the San Juan Province (Argentina): an Analysis of Their Spatial Distribution and Main Forcing Factors. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 10515, **2013**.

8. M. Tonini, **C. Vega Orozco** and M. Conedera. A Robust GIS Approach to Map WUI in the Alpine Regions. In *ICFFRMM - International Conference on Forest Fire Risk Modelling and Mapping: Vulnerability to Forest Fire at Wildland Urban Interface, Aix-en-Provence (France)*, **2013**.

9. M. Tonini, **C. Vega Orozco** and M. Kanevski. Spatio-Temporal Aggregation of Wildfires: from Global Cluster to Local Mapping. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.

10. M. Tonini, **C. Vega Orozco**, M. Kanevski and M. Conedera. Spatio-Temporal Patterns of Forest Fires: a Comprehensive Application of the K-Function. In *European Geosciences Union General Assembly, Geophysical Research AbstractsCopernicus Publications, Vienna (Austria)*, volume 15, page 5461, **2013**.

11. **C. Vega Orozco**, M. Kanevski, M. Tonini, J. Golay and M.J. Pereira. Time Fluctuation Analysis of Forest Fire Sequences. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 15, page 5518, **2013**.

12. **C. Vega Orozco**, M. Leuenberger, M. Tonini and M. Kanevski. Anthropogenic Forest Fires Susceptibility Mapping using Random Forest Algorithm. In *ICF-FRMM - International Conference on Forest Fire Risk Modelling and Mapping: Vulnerability to Forest Fire at Wildland Urban Interface, Aix-en-Provence (France)*, **2013**.

13. **C. Vega Orozco**, M. Tonini and M. Kanevski. Analysis of the Dynamic of Urban Areas and of their Interaction with Forest Fires. In 11^{th} *Swiss Geoscience Meeting, Symposium 17: Computation GIScience, Lausanne (Switzerland)*, **2013**.

2012

14. R. Ceré, M. Conedera, G. Matasci, M. Kanevski, M. Tonini, **C. Vega Orozco** and M. Volpi. Wildland-Urban Interface Evolution Mapping Using Multi-Temporal Landsat Imagery. The Case of Forest Fires in Southern Swiss Alps. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page , **2012**.

15. A. Champendal, **C. Vega Orozco**, R. Ceré, M. Kanevski and M. Tonini. A Geomatic Approach to Wildland-Urban Interface Detection. In *Proceedings of Spatial Analysis and GEOmatics, Liège (Belgium)*, pages 73-76, **2012**.

16. R. Costa, M.J. Pereira, L. Caramelo, **C. Vega Orozco** and M. Kanevski. Spatio-Temporal Clustering of Wildfires in Portugal. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page 12208, **2012**.

17. J. Golay, M. Kanevski, **C. Vega Orozco**. Multipoint-Morisita Index for the Analysis of Spatial Patterns. In *Proceedings of geoENV2012, IX Conference on Geostatistics for Environmental Applications, Valencia (Spain)*, pages 129-130, **2012**.

18. J. Golay, **C. Vega Orozco**, M. Tonini and M. Kanevski. Spatial Point Pattern Analysis of Environmental Data Using R. In *Symposium Proceedings, Open Source Geospatial Research & Education Symposium, Yverdon-les-Bains (Switzerland)*, pages 245-252, **2012**.

19. M. Kanevski, A. Champendal, **C. Vega Orozco**, M. Tonini and M. Conedera. A New Concept of Wildland-Urban Interface Based on the City Clustering Algorithm. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page 2092, **2012**.

20. **C. Vega Orozco**, J. Golay and M. Kanevski. Multifractal Portrayal of the Distribution of the Swiss Population. In *Proceedings of Spatial Analysis and GEOmatics, Liège (Belgium)*, pages 392-407, **2012**.

21. **C. Vega Orozco**, M. Kanevski, M. Tonini, J. Golay and M. Conedera. Multifractal Analysis of Forest Fires in Complex Regions. In *European Geosciences Union General Assembly, Copernicus Publications, Vienna (Austria)*, volume 14, page 1162, **2012**.

2011

22. M. Conedera, M. Tonini, L. Oleggini, B. Pezzatti and **C. Vega Orozco**. Wildland-Urban Interface and Forest Fire Ignition in Alpine Conditions. In *ICFBR 2011 International Conference on Fire Behaviour and Risk Focus on Wildland Urban Interface, Sassari (Italy)*, pages 47-48, **2011**.

23. M. Kanevski, **C. Vega Orozco**, M. Tonini, V. Timonin and M. Conedera. Spatial Analysis of Forest Fires Clustering. Case Study: Canton Ticino, Switzerland. In *Proceedings of Spatial Statistics: Mapping Global Change, Enschede (The Netherlands)*, page P2.31, **2011**.

24. **C. Vega Orozco**, M. Kanevski, M. Tonini and M. Conedera. Patterns Mining in Spatial-Temporal Sequences: the Case of Forest Fires in Ticino (Switzerland). In *Proceedings of the International Symposium on Spatial-Temporal Analysis and Data Mining STDM, London (UK)*, **2011**.

25. **C. Vega Orozco**, M. Tonini, M. Kanevski and M. Conedera. Point Pattern Analysis of Forest Fire Occurrences in Canton Ticino (Switzerland). In *ICFBR 2011 International Conference on Fire Behaviour and Risk Focus on Wildland Urban Interface, pages 126-127, Sassari (Italy)*, **2011**.

Appendix B

Software and Modules Development

Version 3.2.1

Date 2015-05-01

Title Global clustering measures

Author Carmen D. Vega Orozco with contributions of Jean Golay for the Morisita index function.

Maintainer Carmen D. Vega Orozco

Depends R (>= 2.10.0), stats

Suggests maptools

Description This appendix portrays all the software and modules developed in this Thesis in the R environment.

R topics documented:

---

boxcounting *The Box-counting method*

Description

Estimation of the Box-counting fractal dimension for point processes. It can be used to perform analyses for point processes in 1D (time sequences) or 2D (spatial). Data must be in a data.frame format.

Usage

```
boxcounting <- function(x, y, minQ=NULL, maxQ=NULL, winobs=NULL
                        ts=FALSE, mQ=NULL)
```

Arguments

x	Vector of x coordinates of data points.
y	Vector of y coordinates of data points.
minQ	Minimum number of boxes.
maxQ	Maximum number of boxes.
winobs	Vector containing the x and y ranges of the observed window: (min x, max x, min y, max y).
ts	Logical value indicating whether the point process is a time series.
mQ	Matrix holding the counts of the events for each box at all box sizes. When "NULL", the algorithm estimates this matrix.

Details

The computation of this method consists on breaking the fractal pattern (point set) into pieces by superimposing a regular grid of boxes of size δ_1 on the region under study. Then the number of boxes, $N(\delta_1)$, necessary to cover the pattern is counted; that is, one counts all occupied boxes despite the number of points in each box. Next, the linear size of the boxes is reduced, $\delta_2(< \delta_1)$ and the number of boxes, $N(\delta_2)$, is counted again. The algorithm goes on until a minimum size δ_{min} is reached. For a fractal pattern, the scales (δ) and the number of boxes ($N(\delta)$) follow a power law:

$$N(\delta) \propto \delta^{-df_{box}} \tag{B.1}$$

where df_{box} is the fractal dimension measured with the Box–counting method and it is obtained by calculating the slope of the linear regression fitting the data of the plot which relates $\log(N(\delta))$ to $\log(\delta)$. Thus, this method quantifies the dimensionality of a pattern by estimating the degree of the pattern spatial coverage of the space.

Value

This function returns a data.frame containing:

Q	the box size.
logQ	natural logarithm of the box size.
nQ	the number of boxes for each box size Q.
lognQ	the natural logarithm of the number of boxes for each box size Q.

---

morisita *The Morisita index*

Description

Estimation of the Morisita Index for point processes. It can be used to perform a Morisita analysis for point processes in 1D (time sequences) or 2D (spatial), functional measures and multipoint m-Morisita.

Usage

```
morisita <- function(x, y, marks=NULL, minQ=0, maxQ=1, nPmin=2,
                           nPmax=2, winobs=NULL, ts=FALSE)
```

Arguments

x	Vector of x coordinates of data points.
y	Vector of y coordinates of data points.
marks	(Optional) For marked point processes. Values must be given in a vector or data frame.
minQ	Minimum number of boxes.
maxQ	Maximum number of boxes.
nPmin	Minimum n-Point Morisita (number of points to be taken for probability estimation). By default is defined as 2.
nPmax	Maximum n-Point Morisita (maximum number of points to be taken for probability estimation). By default is defined as 2 (when only the 2-Point Morisita (conventional) is desired).
winobs	Vector containing the x and y ranges of the observed window: (min x, max x, min y, max y).
ts	Logical value indicating whether the point process is a time series.

Details

This function was made with the contribution of Jean Golay. The Morisita index of dispersion is a statistical measure used to characterise quantitatively the clustering (non-homogeneity) of a point set. This classical index, I_δ, is obtained by dividing the study region into Q identical boxes of size δ (i.e. length of the diagonal); first starting with a relatively small box size which in turn is increased until a chosen maximum value is reached. The number of events n_i within each box i of size δ is counted and related to the box size. The Morisita index is then computed as:

$$I_\delta = Q \frac{\sum_{i=1}^{Q} n_i(n_i - 1)}{N(N-1)} \tag{B.2}$$

where N is the total number of points. It is then possible to draw a plot relating every I_δ to its corresponding δ.

Now, the generalization of the classical formulation of the Morisita index, called m-Morisita, is made by considering m points with $m \geq 2$. It refers to a family of indices computed as:

$$I_{m,\delta} = Q^{m-1} \frac{\sum_{i=1}^{Q} n_i(n_i-1)(n_i-2)\cdots(n_i-m+1)}{N(N-1)(N-2)\cdots(N-m+1)} \tag{B.3}$$

Value

This function returns a data.frame containing:

Q	the box size.
n-2	the resulting values for two points Morisita.
⋮	
n-m	the resulting values for m-points Morisita.

---

mQcount *Quadrats count*

Description

Divides the observation window into boxes (quadrats) and counts the number of points or the probability function in each box. These two operations are computed for a range of defined box sizes. It can be used for boxes in 1D (time sequences) or 2D (spatial). Data must be in a data.frame format.

Usage

```
mQcount <- function(x, y, marks=NULL, minQ=1, maxQ=1, winobs=NULL,
                         ts=FALSE)
```

Arguments

x	Vector of x coordinates of data points.
y	Vector of y coordinates of data points.
marks	(Optional) For marked point processes. Values must be given in a vector or data frame.
minQ	Minimum number of boxes.
maxQ	Maximum number of boxes.
winobs	Vector containing the x and y ranges of the observed window: (min x, max x, min y, max y).
ts	Logical value indicating whether the point process is a time series.

Details

Quadrats count is an elementary technique for analysing spatial point patterns. The window containing the point pattern is divided into an $Qx * Qy$ grid of rectangular boxes or "quadrats". The number of points (or the probability) in each box is counted. Then, Q goes from $minQ$ until a maximum $maxQ$ is reached.

Value

This function returns a matrix containing:

Rows	corresponds the size of the quadrats.
Columns	contains the number of points (or the probability) in each quadrat.

---

mQcountFrac *Quadrats count for fractal analysis*

Description

Divides the observation window into 2^Q boxes (quadrats) and counts the number of points or the probability function in each box. These two operations are computed for a range of defined box sizes. It can be used for boxes in 1D (time sequences) or 2D (spatial). Data must be in a data.frame format.

Usage

```
mQcountFrac <- function(x, y, marks=NULL, minQ=0, maxQ=1,
                        winobs=NULL, ts=FALSE)
```

Arguments

x	Vector of x coordinates of data points.
y	Vector of y coordinates of data points.
marks	(Optional) For marked point processes. Values must be given in a vector or data frame.
minQ	Minimum number of boxes.
maxQ	Maximum number of boxes.
winobs	Vector containing the x and y ranges of the observed window: (min x, max x, min y, max y).
ts	Logical value indicating whether the point process is a time series.

Details

It is the same procedure as for the "mQcount" function but for fractal process, where the observation window is divided by 2^Q boxes, instead of a sequence of $Q + 1$ divisions. (See Quadrats count). The number of points (or the probability) in each box is counted.

Value

This function returns a matrix containing:

Rows	corresponds the size of the quadrats.
Columns	contains the number of points (or the probability) in each quadrat.

multifractal *The multifractal formalism*

Description

Estimation of the Multifractal formalism for point processes. It estimates two functions: the Rényi generalised dimensions and the multifractal singularity spectrum. It can be used to perform a multifractal analysis for point processes in 1D (time sequences) or 2D (spatial). Data must be in a data.frame format.

Usage

```
multifractal <- function(x, y, marks=NULL, minQ=NULL, maxQ=NULL,
                         qmin=0, qmax=0, winobs=NULL, qseq=1,
                         ts=FALSE, drange=NULL)
```

Arguments

x	Vector of x coordinates of data points.
y	Vector of y coordinates of data points.
marks	(Optional) For marked point processes. Values must be given in a vector or data frame.
minQ	Minimum number of boxes.
maxQ	Maximum number of boxes.
qmin	Minimum q order moment. By default is defined as 0.
qmax	Maximum q order moment. By default is defined as 0.
winobs	Vector containing the x and y ranges of the observed window: (min x, max x, min y, max y).
ts	Logical value indicating whether the point process is a time series.
drange	A vector indicating the scale range at which the linear regression will be made, c(1,2).

Details

Two different approaches are used here to conduct a multifractal analysis: (1) the Rényi generalised dimensions (see function "renyidim") which can be viewed as a global parameter examining how different densities are distributed in the space; and (2) the multifractal singularity spectrum (see function "singspect") viewed as a local parameter which examines the regularity of the distribution of regions with similar scaling indices. Although, each of these two methods present a different approach, they both describe the same information. They are both based on the Box-counting method, where a regular grid of boxes of size δ is superimposed on the point set and, for both methods , a probability distribution is computed in each box.

Value

This formalism returns a data.frame containing:

q	the evaluated order moments.
Dq	D_q, the Rényi generalised dimensions.
Alpha	α_q, the singularity strengths at order q.
fAlpha	$f(\alpha_q)$, the spectrum of the fractal dimensions of boxes with singularity strength α_q.
Tq	τ_q, mass exponent function at order q.

---

renyidim *The Rényi generalised dimensions*

Description

Estimation of the Rényi information for point processes. It can be used to perform analyses for point processes in 1D (time sequences) or 2D (spatial). Data must be in a data.frame format.

Usage

```
renyidim <- function(x, y, marks=NULL, minQ=NULL, maxQ=NULL, qmin=0,
                           qmax=0, qseq=1, winobs=NULL, ts=FALSE, mQ=NULL)
```

Arguments

x	Vector of x coordinates of data points.
y	Vector of y coordinates of data points.
marks	(Optional) For marked point processes. Values must be given in a vector or a data.frame.
minQ	Minimum number of boxes.
maxQ	Maximum number of boxes.
qmin	Minimum q order moment. By default is defined as 0.
qmax	Maximum q order moment. By default is defined as 0 (when the fractal dimension analysis is desired).
qseq	Value at which the q moment will increase.
winobs	Vector containing the x and y ranges of the observed window: (min x, max x, min y, max y).
ts	Logical value indicating whether the point process is a time series.
mQ	Matrix holding the counts of events or the estimation of the mass function for each box at all box sizes. When "NULL", the algorithm estimates this matrix.

Details

The spectrum of generalised dimensions, D_q, is estimated by computing the Rényi information, $I_q(\delta)$, of qth order:

$$I_q(\delta) = \frac{1}{(1-q)} \log(\sum_{i=1}^{N(\delta)} p_i(\delta)^q) \tag{B.4}$$

where $p_i(\delta)$ is the probability distribution in the ith box of size δ, $q \in Z$, and the sum is taken for all non-empty boxes. When $q \to 1$, $I_q(\delta)$ is defined as:

$$I_q(\delta) = -\sum_{i=1}^{N(\delta)} p_i(\delta) \log(p_i(\delta)) \tag{B.5}$$

Value

This function returns a data.frame containing:

Q	the box size.
logQ	the natural logarithm of the box size.
Rq	the Rényi information for each q order moment.

---

singspect *The singularity spectrum*

Description

Estimation of the singularity strength (α) and singularity spectrum (multifractal spectrum: $f(\alpha)$) for point processes. It can be used to perform analyses for point processes in 1D (time sequences) or 2D (spatial). Data must be in a data.frame format. The code is based on the algorithms presented the book of: Seuront L., Fractals and multifractals in ecology and aquatic science, CRC Press, Boca Raton, 2010.

Usage

```
singspect <- function(x, y, marks=NULL, minQ=NULL, maxQ=NULL,
                      qmin=0, qmax=0, qseq=1, winobs=NULL,
                      ts=FALSE, mQ=NULL)
```

Arguments

x	Vector of x coordinates of data points.
y	Vector of y coordinates of data points.
marks	(Optional) For marked point processes. Values must be given in a vector or a `data.frame`.
minQ	Minimum number of boxes.
maxQ	Maximum number of boxes.
qmin	Minimum q order moment. By default is defined as 0.
qmax	Maximum q order moment. By default is defined as 0 (when the fractal dimension analysis is desired).
qseq	Value at which the q moment will increase.
winobs	Vector containing the x and y ranges of the observed window: (min x, max x, min y, max y).
ts	Logical value indicating whether the point process is a time series.
mQ	Matrix holding the counts of events or the estimation of the mass function for each box at all box sizes. When "NULL" the algorithm estimates this matrix.

Details

Chhabra and Jensen, 1989 proposed a method estimating the multifractal singularity spectrum directly from the data as a function of the qth order moments without the application of the Legendre transform. Let $\mu(q)$ be the normalised measure of the probabilities in the boxes of size δ, such as:

$$\mu_i(q,\delta) = \frac{[p_i(\delta)]^q}{\sum_i [p_i(\delta)]^q} \tag{B.6}$$

where, q provides a tool for exploring denser and rarer regions of the singular measure. Then, α_q and $f(\alpha_q)$ can be computed as:

$$\alpha_q = \lim_{\delta \to 0} \frac{\sum_i^{N(\delta)} \mu_i(q,\delta) \log p_i(\delta)}{\log \delta} \tag{B.7}$$

and

$$f(\alpha_q) = \lim_{\delta \to 0} \frac{\sum_i^{N(\delta)} \mu_i(q,\delta) \log \mu_i(q,\delta)}{\log \delta} \tag{B.8}$$

Value

This function returns a `data.frame` containing:

Q	the box size.
logQ	the natural logarithm of the box size.
Aq	α_q, the singularity strength for the order q.
fAq	$f(\alpha_q)$, the fractal dimension of the singularity strength for the order q.

Bibliography

[1] I. Adjali and S. Appleby. The Multifractal Structure of the Human Population Distribution. In N.J. Tate and P.M. Atkinson, editors, *Modelling Scale in Geographical Information Science*, chapter 4, pages 69–85. Wiley, Chichester, 2001. ISBN 978-0-471-98546-4. URL `http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0471985465.html`.

[2] A.A. Ager, N.M. Vaillant, and M.A. Finney. A Comparison of Landscape Fuel Treatment Strategies to Mitigate Wildland Fire Risk in the Urban Interface and Preserve Old Forest Structure. *Forest Ecology and Management*, 259(8):1556–1570, 2010. doi: 10.1016/j.foreco.2010.01.032. URL `http://dx.doi.org/10.1016/j.foreco.2010.01.032`.

[3] J.R. Alavalapati, D.R. Carter, and D.H. Newman. Wildland-Urban Interface: Challenges and Opportunities. *Forest Policy and Economics*, 7(5):705–708, 2005. doi: 10.1016/j.forpol.2005.03.001. URL `http://dx.doi.org/10.1016/j.forpol.2005.03.001`.

[4] A.A. Ali, C. Carcaillet, and Y. Bergeron. Long-Term Fire Frequency Variability in the Eastern Canadian Boreal Forest: The Influences of Climate vs. Local Factors. *Global Change Biology*, 15(5):1230–1241, 2009. ISSN 1365-2486. doi: 10.1111/j.1365-2486.2009.01842.x. URL `http://dx.doi.org/10.1111/j.1365-2486.2009.01842.x`.

[5] C. Allain and M. Cloitre. Characterizing the Lacunarity of Random and Deterministic Fractal Sets. *Physical Review A*, 44(6):3552–3558, 1991. doi: 10.1103/PhysRevA.44.3552. URL `http://link.aps.org/doi/10.1103/PhysRevA.44.3552`.

[6] M.A. Amaral-Turkman, K.F. Turkman, Y. Le Page, and J.M.C. Pereira. Hierarchical Space-Time Models for Fire Ignition and Percentage of Land Burned by Wildfires. *Environmental and Ecological Statistics*, 18(4):601–617, 2011. ISSN 1352-8505. doi: 10.1007/s10651-010-0153-9. URL `http://dx.doi.org/10.1007/s10651-010-0153-9`.

[7] G. Amatulli, M.J. Rodrigues, M. Trombetti, and R. Lovreglio. Assessing Long-Term Fire Risk at Local Scale by Means of Decision Tree Technique. *Journal of*

Geophysical Research, 111(G4):n/a–n/a, 2006. doi: 10.1029/2005JG000133. URL `http://dx.doi.org/10.1029/2005JG000133`.

[8] G. Amatulli, F. Peréz-Cabello, and J. de la Riva. Mapping Llightning/Hhuman-Caused Wildfires Occurrence Under Ignition Ppoint Location Uncertainty. *Ecological Modelling*, 200(3-4):321–333, 2007. doi: 10.1016/j.ecolmodel.2006.08.001. URL `http://dx.doi.org/10.1016/j.ecolmodel.2006.08.001`.

[9] K. Angayarkkani and N. Radhakrishnan. Efficient Forest Fire Detection System: A Spatial Data Mining and Image Processing Based Approach. *International Journal of Computer Science and Network Security*, 9(3):100–107, 2009. URL `http://paper.ijcsns.org/07_book/200903/20090313.pdf`.

[10] S. Appleby. Multifractal Characterization of the Distribution Pattern of the Human Population. *Geographical Analysis*, 28(2):147–160, 1996. doi: 10.1111/j.1538-4632.1996.tb00926.x. URL `http://dx.doi.org/10.1111/j.1538-4632.1996.tb00926.x`.

[11] S.L. Arlinghaus. Fractals Take a Central Place. *Geografiska Annaler. Series B, Human Geography*, 67(2):83–88, 1985. URL `http://www.jstor.org/stable/490419`.

[12] A. Baddeley. Analysing Spatial Point Patterns in R. Technical report, version 4.1, CSIRO, Australia, 2010.

[13] A. Baddeley and R. Turner. Spatstat: an R Package for Analyzing Spatial Point Patterns. *Journal of Statistical Software*, 12(6):1–42, 2005. ISSN 1548-7660. URL `http://www.jstatsoft.org`.

[14] A.J. Baddeley, J. Möller, and R. Waagepetersen. Non- and Semi-Parametric Estimation of Interaction in Inhomogeneous Point Patterns. *Statistica Neerlandica*, 54(3):329–350, 2000. doi: 10.1111/1467-9574.00144. URL `http://dx.doi.org/10.1111/1467-9574.00144`.

[15] A. Badia, P. Serra, and S. Modugno. Identifying Dynamics of Fire Ignition Probabilities in two Representative Mediterranean Wildland-Urban Interface Areas. *Applied Geography*, 31(3):930–940, 2011. doi: 10.1016/j.apgeog.2011.01.016. URL `http://www.sciencedirect.com/science/article/pii/S0143622811000178`.

[16] T.C. Bailey and A.C. Gatrell. *Interactive Spatial Data Analysis*. Longman Scientific & Technical, 1995. ISBN 9780582244931. URL `http://books.google.ch/books?id=WwbvAAAAMAAJ`.

[17] P. Bak, C. Tang, and K. Wiesenfeld. Self-Organized Criticality. *Physical Review A*, 38(1):364–374, 1988. doi: 10.1103/PhysRevA.38.364. URL `http://dx.doi.org/10.1103/PhysRevA.38.364`.

[18] A. Bar Massada, V.C. Radeloff, and S.I. Stewart. Allocating Fuel Breaks to Optimally Protect Structures in the Wildland-Urban Interface. *International Journal of Wildland Fire*, 20(1):59–68, 2011. doi: 10.1071/WF09041. URL `http://dx.doi.org/10.1071/WF09041`.

[19] M. Batty and P. Longley. The Fractal Simulation of Urban Structure. *Environment and Planning A*, 18(9):1143–1179, 1986. doi: 10.1068/a181143. URL `http://www.envplan.com/abstract.cgi?id=a181143`.

[20] M. Batty and P. Longley. *Fractal Cities*. Academic Press, London, 1994. ISBN 9780124555709. URL `https://books.google.ch/books?id=gYR9QgAACAAJ`.

[21] M. Batty, P. Longley, and S. Fotheringham. Urban Growth and Form: Scaling, Fractal Geometry, and Diffusion-Limited Aggregation. *Environment and Planning A*, 21(11):1447–1472, 1989. doi: 10.1068/a211447. URL `http://www.envplan.com/abstract.cgi?id=a211447`.

[22] P.Z. Bermudez, J. Mendes, J.M.C. Pereira, K.F. Turkman, and M.J.P. Vasconcelos. Spatial and Temporal Extremes of Wildfire Sizes in Portugal (1984–2004). *International Journal of Wildland Fire*, 18:983—-991, 2009. doi: 10.1071/WF07044. URL `http://dx.doi.org/10.1071/WF07044`.

[23] J. Besag. Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2):192–236, 1974. URL `http://www.jstor.org/stable/2984812`.

[24] J. Besag. Contribution to the Discussion of Dr Ripley's Paper. *Journal of the Royal Statistical Society B*, 39(2):193–195, 1977. doi: citeulike:8236341. URL `http://www.citeulike.org/user/fredoss/article/8236341`.

[25] J. Besag and P.J. Diggle. Simple Monte Carlo Tests for Spatial Pattern. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 26(3):327–333, 1977. ISSN 00359254. URL `http://www.jstor.org/stable/2346974`.

[26] J. Besag and J. Newell. The Detection of Clusters in Rare Diseases. *Journal of the Royal Statistic Society. Series A (Statistics in Society)*, 154(1):143–155, 1991. URL `http://www.jstor.org/stable/2982708`.

[27] F. Bonneu. Exploring and Modeling Fire Department Emergencies with a Spatio-Temporal Marked Point Process, 2006/07. URL `http://www.bentley.edu/csbigs/documents/bonneu.pdf`.

[28] B. Boots, A. Okabe, and R. Thomas. *Modelling Geographical Systems: Statistical and Computational Applications*. Kluwer Academic Publishers, Dordrecht, 2002. ISBN 1-4020-0821-X. doi: 10.1007/978-94-017-2296-4. URL `http://dx.doi.org/10.1007/978-94-017-2296-4`.

[29] S. Borgani, G. Murante, A. Provenzale, and R. Valdarnini. Multifractal Analysis of the Galaxy Distribution: Reliability of Results from Finite Data Sets. *Physical Review E*, 47(6):3879–3888, 1993. doi: 10.1103/PhysRevE.47.3879. URL `http://link.aps.org/doi/10.1103/PhysRevE.47.3879`.

[30] D.M.J.S. Bowman, J.K. Balch, P. Artaxo, W.J. Bond, J.M. Carlson, M.A. Cochrane, C.M. D'Antonio, R.S. DeFries, J.C. Doyle, S.P. Harrison, F.H. Johnston, J.E. Keeley, M.A. Krawchuk, C.A. Kull, J.B. Marston, M.A. Moritz, I.C. Prentice, C.I. Roos, A.C. Scott, T.W. Swetnam, G.R. van der Werf, and S.J. Pyne. Fire in the Earth System. *Science*, 324(5926):481–484, 2009. doi: 10.1126/science.1163886. URL `http://www.sciencemag.org/content/324/5926/481.abstract`.

[31] D.M.J.S. Bowman, J. Balch, P. Artaxo, W.J. Bond, M.A. Cochrane, C.M. D'Antonio, R. DeFries, F.H. Johnston, J.E. Keeley, M.A. Krawchuk, C.A. Kull, M Mack, M.A. Moritz, S. Pyne, C.I. Roos, A.C. Scott, N.S. Sodhi, and T.W. Swetnam. The Human Dimension of Fire Regimes on Earth. *Journal of Biogeography*, 38(12):2223–2236, 2011. ISSN 1365-2699. doi: 10.1111/j.1365-2699.2011.02595.x. URL `http://dx.doi.org/10.1111/j.1365-2699.2011.02595.x`.

[32] G.A. Bradley. *Land Use and Forest Resources: The Urban/Forest Interface*. The Geo. S. Long Publication Series. University of Washington Press, Seattle, 1984. ISBN 978-0295961453. URL `https://books.google.ch/books?id=yS0o4xKz0pkC`.

[33] P. Brang, W. Schönenberger, M. Frehner, R. Schwitter, J.J. Thormann, and B. Wasser. Management of Protection Forests in the European Alps: an Overview. *Forest, Snow and Landscape Research Journal*, 80(1):23–44, 2006. URL `http://www.wsl.ch/dienstleistungen/publikationen/pdf/6947.pdf`.

[34] L. Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001. doi: 10.1023/A:1010933404324. URL `http://dx.doi.org/10.1023/A:1010933404324`.

[35] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone. *Classification and Regression Trees*. Wadsworth Statistics/Probability. Chapman and Hall/CRC; 1 edition, Boca Raton, 1984. ISBN 978-0412048418.

[36] D.R. Brillinger, H.K. Preisler, and J.W. Benoit. Probabilistic Risk Assessment for Wildfires. *Environmetrics*, 17(6):623–633, 2006. ISSN 1099-095X. doi: 10.1002/env.768. URL `http://dx.doi.org/10.1002/env.768`.

[37] M.T. Brunetti, F. Guzzetti, and M. Rossi. Probability Distributions of Landslide Volumes. *Nonlinear Processes in Geophysics*, 16(2):179–188, 2009. doi: 10.5194/npg-16-179-2009. URL `http://www.nonlin-processes-geophys.net/16/179/2009/`.

[38] R.E. Burgan and R.A. Hartford. Computer Mapping of Fire Danger and Fire Locations in the Continental United States. *Journal of Forestry*, 86(1): 25–30, 1988. URL `http://www.ingentaconnect.com/content/saf/jof/1988/00000086/00000001/art00008`.

[39] P.A. Burrough. Fractal Dimensions of Landscapes and Other Environmental Data. *Nature*, 294:240–142, 1981. doi: 10.1038/294240a0. URL `http://www.nature.com/nature/journal/v294/n5838/abs/294240a0.html`.

[40] K. Byrne. *Environmental Science*. Bath Advance Science. Nelson Thornes, Cheltenham, 2001. ISBN 9780174483052.

[41] C. Carcaillet, A.A. Ali, O. Blarquez, A. Genries, B. Mourier, and L. Bremond. Spatial Variability of Fire History in Subalpine Forests: From Natural to Cultural Regimes. *Ecoscience*, 16(1):1–12, 2009. doi: 10.2980/16-1-3189. URL `http://www.bioone.org/doi/abs/10.2980/16-1-3189`.

[42] J.A. Cardille, S.J. Ventura, and M.G. Turner. Environmental and Social Factors Influencing Wildfires in the Upper Midwest, United States. *Ecological Applications*, 11(1):111–127, 2001.

[43] G. Cello, B.D. Malamud, and Geological Society of London. *Fractal Analysis for Natural Hazards*. Geological Society Special Publication. Geological Society, 2006. ISBN 9781862392014. URL `https://books.google.ch/books?id=ntCjZybNOp4C`.

[44] I. Ceschi. *Il Bosco nel Canton Ticino*. Armando Dadò, Locarno, 2006.

[45] Y. Chen. Multifractals of Central Place Systems: Models, Dimension Spectrums, and Empirical Analysis. *Physica A: Statistical Mechanics and its Applications*, 402(0):266–282, 2014. ISSN 0378-4371. doi: 10.1016/j.physa.2014.01.061. URL `http://www.sciencedirect.com/science/article/pii/S037843711400082X`.

[46] Q. Cheng. Multifractal Modeling and Lacunarity Analysis. *Mathematical Geology*, 29(7):919–932, 1997. ISSN 0882-8121. doi: 10.1023/A:1022355723781. URL `http://dx.doi.org/10.1023/A%3A1022355723781`.

[47] Q. Cheng. Non-Linear Theory and Power-Law Models for Information Integration and Mineral Resources Quantitative Assessments. *Mathematical Geosciences*, 40(5):503–532, 2008. doi: 10.1007/s11004-008-9172-6. URL `http://dx.doi.org/10.1007/s11004-008-9172-6`.

[48] Q. Cheng and F.P. Agterberg. Multifractal Modeling and Spatial Point Processes. *Mathematical Geology*, 27(7):831–845, 1995. ISSN 0882-8121. doi: 10.1007/BF02087098. URL `http://dx.doi.org/10.1007/BF02087098`.

[49] T. Cheng and J. Wang. Integrated Spatio-Temporal Data Mining for Forest Fire Prediction. *Transactions in GIS*, 12(5):591–611, 2008. doi: 10.1111/j.1467-9671.2008.01117.x. URL `http://dx.doi.org/10.1111/j.1467-9671.2008.01117.x`.

[50] A. Chhabra and R.V. Jensen. Direct Determination of the $f(\alpha)$ Singularity Spectrum. *Physicl Review Letters*, 62(12):1327–1330, 1989. doi: 10.1103/PhysRevLett.62.1327. URL `http://link.aps.org/doi/10.1103/PhysRevLett.62.1327`.

[51] E. Chuvieco and R.G. Congalton. Application of Remote Sensing and Geographic Information Systems to Forest Fire Hazard Mapping. *Remote Sensing of Environment*, 29(2):147–159, 1989. doi: 10.1016/0034-4257(89)90023-0. URL `http://dx.doi.org/10.1016/0034-4257(89)90023-0`.

[52] S. Clar, B. Drossel, and F. Schwabl. Forest Fires and Other Examples of Self-Organized Criticality. *Journal of Physics: Condensed Matter*, 8(37):6803—-6824, 1996. doi: 10.1088/0953-8984/8/37/004. URL `http://dx.doi.org/10.1088/0953-8984/8/37/004`.

[53] J.S. Clark. Effect of Climate Change on Fire Regimes in Northwestern Minnesota. *Nature*, 334:233–235, 1988. doi: 10.1038/334233a0. URL `http://dx.doi.org/10.1038/334233a0`.

[54] C. Cleve, M. Kelly, F.R. Kearns, and M. Moritz. Classification of the Wildland–Urban Interface: A Comparison of Pixel- and Object-Based Classifications Using High-Resolution Aerial Photography. *Computers, Environment and Urban Systems*, 32(4):317–326, 2008. doi: 10.1016/j.compenvurbsys.2007.10.001. URL `http://www.sciencedirect.com/science/article/pii/S0198971507000737`.

[55] A.D. Cliff and J.K. Ord. *Spatial Autocorrelation*. Monographs in Spatial and Environmental Systems Analysis. Pion, 1973. ISBN 9780850860375. URL `http://books.google.ch/books?id=pBEJAQAAIAAJ`.

[56] A.D. Cliff and J.K. Ord. *Spatial Processes: Models and Applications*. Pion, London, 1981.

[57] J.D. Cohen. Preventing Disaster: Home Ignitability in the Wildland-Urban Interface. *Journal of Forestry*, 98(3):15–21, 2000. URL `http://www.ingentaconnect.com/content/saf/jof/2000/00000098/00000003/art00008`.

[58] J.D. Cohen. Relating Flame Radiation to Home Ignition Using Modeling and Experimental Crown Fires. *Canadian Journal of Forest Research*, 34(8):1616–1626, 2004. doi: 10.1139/x04-049. URL `http://dx.doi.org/10.1139/x04-049`.

[59] T.W. Collins. Households, Forest, and Fire Hazard Vulnerability in the American West: a Case Study of a California Community. *Environmental Hazards*, 6(1):23–37, 2005. doi: 10.1016/j.hazards.2004.12.003. URL `http://dx.doi.org/10.1016/j.hazards.2004.12.003`.

[60] M. Conedera. Incendi di Boschi in Canton Ticino: dallo Studio Pionieristico di Ceschi all Situazione Attuale. *Bolletino della Società Ticinese di Scienze Naturali*, 91(1-2):135–144, 2003.

[61] M. Conedera. *Implementing Fire History and Fire Ecology in Fire Risk Assessment: The Study Case of Canton Ticino (Southern Switzerland)*. Doctoral and habilitation theses, Universität Fridericiana zu Karlsruhe (TH), Karlsruhe, 2009. URL `http://nbn-resolving.org/urn:nbn:de:swb:90-118455`.

[62] M. Conedera and W. Tinner. The Interaction Between Forest Fires and Human Activity in Southern Switzerland. In J.L. Innes, M. Beniston, and M.M. Verstraete, editors, *Biomass Burning and Its Inter-Relationships with the Climate System*, volume 3 of *Advances in Global Change Research*, pages 247–261. Springer Netherlands, 2000. ISBN 978-90-481-5375-6. doi: 10.1007/0-306-47959-1_14. URL `http://dx.doi.org/10.1007/0-306-47959-1_14`.

[63] M. Conedera, M. Marozzi, B. Jud, D. Mandallaz, F. Chatelain, C. Frank, F. Kienast, P. Ambrosetti, and G. Corti. *Incendi Boschivi al Sud delle Alpi: Passato, Presente e Possibili Sviluppi Futuri.* Rapporto di Lavoro PNR 31 | Work Report of the Swiss National Research Program "Climate change and Natural Disasters" PNR 31. VDF Hochschulverlag AG an der ETH Zürich, Zürich, 1996. ISBN 9783728123480. URL `https://books.google.ch/books?id=20_eJ6ejsW8C`.

[64] M. Conedera, D. Mandallaz, R. Bolognesi, and P. Ambrosetti. Forest Fire Research in Switzerland. Part 2: Fire Danger Prediction in the Southern Part of Switzerland. *International Forest Fire News*, 16:2–6, 1997. URL `http://www.slf.ch/it/bellinzona/insubrisch/publications/pubblicazioni_incendi/Conedera_et_al_1997_IntForFire`.

[65] M. Conedera, P. Stanga, B. Oester, and P. Bachmann. The Natural Dynamics of Abandoned Chestnut Stands in Southern Switzerland. In *Recent Dynamics of the Mediterranean Vegetation and Landscape*, pages 237–247. John Wiley & Sons, Ltd, 2005. ISBN 9780470093719. doi: 10.1002/0470093714.ch20. URL `http://dx.doi.org/10.1002/0470093714.ch20`.

[66] M. Conedera, G. Cesti, G.B. Pezzatti, T. Zumbrunnen, and F. Spinedi. Lightning-Induced Fires in the Alpine Region: An Increasing Problem. In D.X. Viegas, editor, *Proceedings of the 5th International Conference on Forest Fire Research*, page 9S, Coimbra, 2006.

[67] M. Conedera, D. Torriani, C. Neff, C. Ricotta, S. Bajocco, and G.B. Pezzatti. Using Monte Carlo Simulations to Estimate Relative Fire Ignition Danger in a Low-to-Medium Fire-Prone Region. *Forest Ecology and Management*, 261(12):2179–2187, 2011. doi: 10.1016/j.foreco.2010.08.013. URL `http://www.sciendirect.com/science/article/pii/S0378112710004688`.

[68] M. Conedera, M. Tonini, L. Oleggini, C. Vega Orozco, M. Leuenberger, and G.B. Pezzatti. Geospatial Approach for Defining the Wildland-Urban Interface in the Alpine Environment. *Computers, Environment and Urban Systems*, 2015.

[69] A. Corral, A. Telesca, and R. Lasaponara. Scaling and Correlations in the Dynamics of Forest-Fire Occurrence. *Physical Review E*, 77(1):016101, 2008. doi: 10.1103/PhysRevE.77.016101. URL `http://dx.doi.org/10.1103/PhysRevE.77.016101`.

[70] M.A. Costa and R.M. Assunção. A Fair Comparison Between the Spatial Scan and the Besag-Newell Disease Clustering Tests. *Environmental and Ecological Statistics*, 12(3):301–319, 2005. doi: 10.1007/s10651-005-1515-6. URL `http://dx.doi.org/10.1007/s10651-005-1515-6`.

[71] M.R. Coughlan and A.M. Petty. Linking Humans and Fire: a Proporsal for a Transdisciplinary Fire Ecology. *International Journal of Wildland Fire*, 21(5): 477–487, 2012. doi: 10.1071/WF11048. URL `http://www.publish.csiro.au/nid/114/paper/WF11048.htm`.

[72] D.R. Cox and V. Isham. *Point Processes*. Chapman & Hall/CRC, London, 1980.

[73] N. Cressie. Geostatistics. *The American Statistician*, 43(4):197–202, 1989. doi: 10.1080/00031305.1989.10475658. URL `http://dx.doi.org/10.1080/00031305.1989.10475658`.

[74] N.A.C. Cressie. *Statistics for Spatial Data*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd, New York, 1993. ISBN 9780471002550.

[75] D.R. Cutler, T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler. Random Forests for Classification in Ecology. *Ecology*, 88:2783–2792, 2007. doi: 10.1890/07-0539.1. URL `http://www.esajournals.org/doi/abs/10.1890/07-0539.1`.

[76] L.R. da Silva, T. Stošić, and B.D. Stošić. Power Law Correlations in Time Series of Wild-Land and Forest Fires in Brazil. *International Journal of Remote Sensing*, 33 (7):2059–2067, 2012. doi: 10.1080/01431161.2011.605086. URL `http://dx.doi.org/10.1080/01431161.2011.605086`.

[77] G. Daccord, J. Nittman, and H.E. Stanley. Radial Viscous Fingers and Diffusion-Limited Aggregation: Fractal Dimension and Growth Sites. *Physical Review Letters*, 56(4):336–339, 1986. doi: 10.1103/PhysRevLett.56.336. URL `http://link.aps.org/doi/10.1103/PhysRevLett.56.336`.

[78] M.R.T. Dale, P. Dixon, M.J. Fortin, P. Legendre, D.E. Myers, and M.S. Rosenberg. Conceptual and Mathematical Relationships among Methods for Spatial Analysis.

Ecography, 25(5):558–577, 2002. ISSN 1600-0587. doi: 10.1034/j.1600-0587.2002.250506.x. URL `http://dx.doi.org/10.1034/j.1600-0587.2002.250506.x`.

[79] D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes.* Springer, New York, 1988.

[80] D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods.* Springer-Verlag, New York, second edition, 2003. ISBN 0-387-95541-0.

[81] D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure.* Springer-Verlag, New York, second edition, 2008.

[82] A. Dauphiné. *Géographie Fractale. Fractals Auto-Similaire et Auto-Affine.* Hermes Science Publications, 2011. ISBN 9782746237988. URL `https://books.google.ch/books?id=JB4EcxdlIR8C`.

[83] J.B. Davis. The Wildland-Urban Interface: Paradise or Battleground? *Journal of Forestry*, 88(1):26–31, 1990. URL `http://www.ingentaconnect.com/content/saf/jof/1990/00000088/00000001/art00008`.

[84] A. De Angelis, C. Ricotta, M. Conedera, and G.B. Pezzatti. Modelling the Meteorological Forest Fire Niche in Heterogeneous Pyrologic Conditions. *PLoS ONE*, 10(2):e0116875, 2 2015. doi: 10.1371/journal.pone.0116875. URL `http://dx.doi.org/10.1371%2Fjournal.pone.0116875`.

[85] M.L. De Keersmaecker, P. Frankhauser, and I. Thomas. Using Fractal Dimensions for Characterizing Intra-Urban Diversity: The Example of Brussels. *Geographical Analysis*, 35(4):310–328, 2003. doi: 10.1111/j.1538-4632.2003.tb01117.x. URL `http://dx.doi.org/10.1111/j.1538-4632.2003.tb01117.x`.

[86] M. de Torres Curth, C. Biscayart, L. Ghermandi, and G. Pfister. Wildland–Urban Interface Fires and Socioeconomic Conditions: A Case Study of a Northwestern Patagonia City. *Environmental Management*, 49(4):876–891, 2012. doi: 10.1007/s00267-012-9825-6. URL `http://dx.doi.org/10.1007/s00267-012-9825-6`.

[87] Ufficio di Statistica del Cantone Ticino and Dipartimento delle finanze e dell'economia. *Annuario Statistico Ticinese 2015.* Ufficio di Statistica, Repubblica e Cantone Ticino, Bellinzona, 2015. ISBN 978-88-8468-019-8.

[88] C. Díaz-Avalos, D.L. Peterson, E. Alvarado, S.A. Ferguson, and J.E. Besag. Space-Tim Modelling of Lightning-Caused Ignitions in the Blue Montains, Oregon. *Canadian Journal of Forest Research*, 31:1579–1593, 2001. doi: 10.1139/cjfr-31-9-1579. URL `http://dx.doi.org/10.1139/cjfr-31-9-1579`.

[89] P.J. Diggle. *Spatial Analysis of Spatial Point Patterns*. Arnold, London, second edition, 2003. ISBN 978-0340740705.

[90] P.J. Diggle. Spatio-Temporal Point Processes: Methods and Applications. In B. Finkenstädt, L. Held, and V. Isham, editors, *Statistical Methods for Spatio-Temporal Systems*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, chapter 1, pages 2–45. Chapman & Hall/CRC, Boca Raton, 2007. ISBN 978-1-58488-593-1. URL `https://books.google.ch/books?id=j-rtW1nVuCoC`.

[91] P.J. Diggle. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Chapman & Hall/CRC Press, Boca Raton, third edition, 2014. ISBN 978-1-4665-6023-9.

[92] P.J. Diggle and E. Gabriel. Spatio-Temporal Point Processes. In A.E. Gelfand, P.J. Diggle, M. Fuentes, and P. Guttorp, editors, *Handbook of Spatial Statistics*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, chapter 25, pages 449–461. Chapman & Hall/CRC Press, Boca Raton, 2010. ISBN 978-1-4200-7287-7.

[93] P.J. Diggle, A.G. Chetwynd, Häggkvist R., and S.E. Morris. Second-Order Analysis of Space-Time Clustering. *Statistical Methods in Medical Research*, 4:124–136, 1995. doi: 10.1177/096228029500400203. URL `http://dx.doi.org/10.1177/096228029500400203`.

[94] J.B. Drake and J.F. Weishampel. Simulating Vertical and Horizontal Multifractal Patterns of a Longleaf Pine Savanna. *Ecological Modelling*, 145(2-3):129–142, 2001. doi: 10.1016/S0304-3800(01)00398-2. URL `http://dx.doi.org/10.1016/S0304-3800(01)00398-2`.

[95] R. Durão and A. Soares. Integrating Meteorological Dynamic Data and Historical Data Into a Stochastic Model for Predicting Forest Fires Risk Maps. In P.M. Atkinson and C.D. Lloyd, editors, *geoENV VII – Geostatistics for Environmental Applications*, volume 16 of *Quantitative Geology and Geostatistics*, pages 77–88. Springer Netherlands, 2010. ISBN 978-90-481-2321-6. doi: 10.1007/978-90-481-2322-3_7. URL `http://dx.doi.org/10.1007/978-90-481-2322-3_7`.

[96] ESRI Environmental Systems Research Institute. Modelbuilder, 2014. URL `http://resources.arcgis.com/en/help/main/10.2/index.html#/What_is_ModelBuilder/002w00000001000000/`.

[97] M. Ester, H.P. Kriegel, J. Sander, and X. Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pages

226–231. AAAI Press, 1996. doi: 10.1.1.71.1980. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.71.1980.

[98] D.A. Falk, C. Miller, D. McKenzie, and A.E. Black. Cross-Scale Analysis of Fire Regimes. *Ecosystems*, 10(5):809–823, 2007. doi: 10.1007/s10021-007-9070-7. URL http://dx.doi.org/10.1007/s10021-007-9070-7.

[99] J. Feder. *Fractals (Physics of Solids and Liquids)*. Plenum Press, New York, 1988. ISBN 978-1-4899-2126-0. doi: 10.1007/978-1-4899-2124-6. URL http://link.springer.com/book/10.1007%2F978-1-4899-2124-6.

[100] B. Finkenstädt, L. Held, and V. Isham. *Statistical Methods for Spatio-Temporal Systems*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Chapman & Hall/CRC, 2007. ISBN 978-1-58488-593-1. URL https://books.google.ch/books?id=j-rtW1nVuCoC.

[101] M.D. Flannigan, Y. Bergeron, O. Engelmark, and B.M. Wotton. Future Wildfire in Circumboreal Forests in Relation to Global Warming. *Journal of Vegetation Science*, 9(4):469–476, 1998. URL http://www.jstor.org/stable/3237261.

[102] A.S. Fotheringham, C. Brunsdon, and M.E. Charlton. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons Ltd, 1 edition, Hoboken, N.J., 2002. ISBN 978-0471496168.

[103] P. Frankhauser. The Fractal Approach: A New Tool for the Spatial Analysis of Urban Agglomerations. *Population*, 1:205–240, 1998. URL /web/revues/home/prescript/article/pop_0032-4663_1998_hos_10_1_6828.

[104] N. François, P. Frankhauser, and D. Pumain. Villes, Densité et Fractalité. Nouvelles Représentations de la Répartition de la Population. *Annales de la Recherche Urbaine*, 67:55–63, 1995. URL http://crdaln.documentation.developpement-durable.gouv.fr/document.xsp?id=Cdu-0067410.

[105] D.A. Freedman. *Statistical Models. Theory and Practice*. Cambridge University Press, 2 edition, New York, 2009. ISBN 978-0521743853.

[106] J.H. Friedman. Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1):1–67, 1991. doi: 10.1214/aos/1176347963. URL http://dx.doi.org/10.1214/aos/1176347963.

[107] S. Frontier. Applications of Fractal Theory to Ecology. In P. Legendre and L. Legendre, editors, *Develoments in Numerical Ecology*, volume 14 of *NATO ASI Series*, pages 335–378. Springer Berlin Heidelberg, 1987. ISBN 978-3-642-70882-4. doi: 10.1007/978-3-642-70880-0_9. URL http://dx.doi.org/10.1007/978-3-642-70880-0_9.

[108] E. Gabriel and P.J. Diggle. Second-Order Analysis in Inhomogeneous Spatio-Temporal Point Process Data. *Statistica Neerlandica*, 63(1):43–51, 2009. doi: 10.1111/j.1467-9574.2008.00407.x. URL `http://dx.doi.org/10.1111/j.1467-9574.2008.00407.x`.

[109] A. Gabrielli, F. Sylos Labini, M. Joyce, and L. Pietronero. *Satistical Physics for Cosmic Structures*. Springer-Verlag, Berlin, 2005.

[110] A. Gatrell and B. Rowlingson. Spatial Point Process Modelling in a GIS Environment. In S Fotheringham. and P. Rogerson, editors, *Spatial Analysis and GIS*. Taylor and Francis, 1994. ISBN 0748401032.

[111] A.C. Gatrell, T.C. Bailey, P.J. Diggle, and B.S. Rowlingson. Spatial Point Pattern Analysis and its Application in Geographical Epidemiology. *Transactions of the Institute of British Geographers*, 21(1):256–274, 1996. ISSN 00202754. URL `http://www.jstor.org/stable/622936`.

[112] A.E. Gelfand, P.J. Diggle, M. Fuentes, and P. Guttorp. *Handbook of Spatial Statistics*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Chapman & Hall/CRC Press, Boca Raton, 2010. ISBN 978-1-4200-7287-7.

[113] M. Gellrich, P. Baur, B. Koch, and N.E. Zimmermann. Agricultural Land Abandonment and Natural Forest Re-Growth in the Swiss Mountains: A Spatially Explicit Economic Analysis. *Agriculture, Ecosystems & Environment*, 118(1-4): 93–108, 2007. doi: 10.1016/j.agee.2006.05.001. URL `http://dx.doi.org/10.1016/j.agee.2006.05.001`.

[114] M.G. Genton, D.T. Butry, M.L. Gumpertz, and J.P. Prestemon. Spatio-Temporal Analysis of Wildfire Ignitions in the St Johns River Water Management District, Florida. *International Journal of Wildland Fire*, 15(1):87–97, 2006. doi: 10.1071/WF04034. URL `http://dx.doi.org/10.1071/WF04034`.

[115] L. Ghermandi, R. Lasaponara, and L. Telesca. Intra-Annual Time Dynamical Patterns of Fire Sequences Observed in Patagonia (Argentina). *Ecological Modelling*, 221(1):94–97, 2010. doi: 10.1016/j.ecolmodel.2008.11.011. URL `http://dx.doi.org/10.1016/j.ecolmodel.2008.11.011`.

[116] A.M. Gill and S.L. Stephens. Scientific and Social Challenges for the Management of Fire-Prone Wildland-Urban Interfaces. *Environmental Research Letters*, 4(3): 034014, 2009. doi: 10.1088/1748-9326/4/3/034014. URL `http://dx.doi.org/10.1088/1748-9326/4/3/034014`.

[117] N.P. Gillett, A.J. Weaver, F.W. Zwiers, and M.D. Flannigan. Detecting the Effect of Climate Change on Canadian Forest Fires. *Geophysical Research Letters*, 31 (18):L18211, 2004. ISSN 1944-8007. doi: 10.1029/2004GL020876. URL `http://dx.doi.org/10.1029/2004GL020876`.

[118] U. Gimmi, M. Bürgi, and M. Stuber. Reconstructing Anthropogenic Disturbance Regimes in Forest Ecosystems: A Case Study from the Swiss Rhone Valley. *Ecosystems*, 11(1):113–124, 2008. ISSN 1432-9840. doi: 10.1007/s10021-007-9111-2. URL `http://dx.doi.org/10.1007/s10021-007-9111-2`.

[119] J. Glaz, J. Naus, and S. Wallenstein. *Scan Statistics*. Graduate Texts in Mathematics. Springer, 2001. ISBN 9780387988191. URL `http://books.google.ch/books?id=CHUwtWl6zOYC`.

[120] J. Golay, M. Kanevski, C. Vega Orozco, and M. Leuenberger. The Multipoint Morisita Index for the Analysis of Spatial Patterns. *Physica A: Statistical Mechanics and its Applications*, 406:191–202, 2014. doi: 10.1016/j.physa.2014.03.063. URL `http://dx.doi.org/10.1016/j.physa.2014.03.063`.

[121] F. Gonzalez-Alonso, J.M. Cuevas, J.L. Casanova, A. Calle, and P. Illera. A Forest Fire Risk Assessment using NOAA AVHRR Images in the Valencia Area, Eastern Spain. *International Journal of Remote Sensing*, 18(10):2201–2207, 1997. doi: 10.1080/014311697217837. URL `http://dx.doi.org/10.1080/014311697217837`.

[122] J.R. Gonzalez-Olabarria, B. Mola-Yudego, T. Pukkala, and M. Palahi. Using Multiscale Spatial Analysis to Assess Fire Ignition Density in Catalonia, Spain. *Annals of Forest Science*, 68(4):861–871, 2011. doi: 10.1007/s13595-011-0082-2. URL `http://dx.doi.org/10.1007/s13595-011-0082-2`.

[123] M. Goodchild. A Spatial Analytical Perspective on Geographical Information Systems. *International Journal of Geographical Information Systems*, 1(4):327–334, 1987. doi: 10.1080/02693798708927820. URL `http://dx.doi.org/10.1080/02693798708927820`.

[124] M.F. Goodchild. *Spatial Autocorrelation*. CATMOG Series. Geo Books, 1986. URL `http://books.google.ch/books?id=2BYnAQAAIAAJ`.

[125] M.F. Goodchild and D.M. Mark. The Fractal Nature of Geographic Phenomena. *Annals of the Association of American Geographers*, 77(2):265–278, 1987. ISSN 1467-8306. doi: 10.1111/j.1467-8306.1987.tb00158.x. URL `http://dx.doi.org/10.1111/j.1467-8306.1987.tb00158.x`.

[126] P. Grassberger. Generalized Dimensions of Strange Attractors. *Physics Letters A*, 97(6):227–230, 1983. doi: 10.1016/0375-9601(83)90753-3. URL `http://dx.doi.org/10.1016/0375-9601(83)90753-3`.

[127] P. Grassberger and I. Procaccia. Measuring the Strangeness of Strange Attractors. *Physica D*, 9(1-2):189–208, 1983. doi: 10.1016/0167-2789(83)90298-1. URL `http://dx.doi.org/10.1016/0167-2789(83)90298-1`.

[128] J.D. Greenberg and G.A. Bradley. Analyzing the Urban-Wildland Interface with GIS: Two Case Studies. *Journal of Forestry*, 95(10):18–22, 1997. URL http://www.ingentaconnect.com/content/saf/jof/1997/00000095/00000010/art00008.

[129] J.P. Greenberg, P.R. Zimmerman, L.E. Heidt, and W.H. Pollock. Hydrocarbon and Carbon Monoxide Emissions from Biomass Burning in Brazil. *Journal of Geophysical Research*, 89(D1):1350–1354, 1984. doi: 10.1029/JD089iD01p01350. URL http://dx.doi.org/10.1029/JD089iD01p01350.

[130] D. Guglietta, M. Conedera, S. Mazzoleni, and C. Ricotta. Mapping Fire Ignition Risk in a Complex Anthropogenic Landscape. *Remote Sensing Letters*, 2(3):213–219, 2011. doi: 10.1080/01431161.2010.512927. URL http://dx.doi.org/10.1080/01431161.2010.512927.

[131] R.P. Guyette, R.M. Muzika, and D.C. Dey. Dynamics of an Anthropogenic Fire Regime. *Ecosystems*, 5(5):472–486, 2002. doi: 10.10071/s10021-002-0115-7. URL http://dx.doi.org/10.10071/s10021-002-0115-7.

[132] R.G. Haight, D.T. Cleland, R.B. Hammer, V.C. Radeloff, and T.S. Rupp. Assessing Fire Risk in the Wildland-Urban Interface. *Journal of Forestry*, 102(7):41–48, 2004. URL http://www.ingentaconnect.com/content/saf/jof/2004/00000102/00000007/art00007.

[133] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, and B.I. Shraiman. Fractal Measures and Their Singularities: The Characterization of Strange Sets. *Physical Review A*, 33(2):1141–1151, 1986. doi: 10.1103/PhysRevA.33.1141. URL http://link.aps.org/doi/10.1103/PhysRevA.33.1141.

[134] R.B. Hammer, S.I. Stewart, and V.C. Radeloff. Demographic Trends, the Wildland-Urban Interface, and Wildfire Management. *Society & Natural Resources*, 22(8):777–782, 2009. doi: 10.1080/08941920802714042. URL http://dx.doi.org/10.1080/08941920802714042.

[135] J. Han, M. Kamber, and J. Pei. *Data Mining: Concepts and Techniques*. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann Publishers Inc., Boston, 3rd edition, 2012. ISBN 978-0-12-381479-1. doi: 10.1016/B978-0-12-381479-1.00001-0. URL http://www.sciencedirect.com/science/article/pii/B9780123814791000010.

[136] W.M. Hao, M.-H. Liu, and P.J. Crutzen. Estimates of Annual and Regional Fire Releases of CO2 and other Trace Gases to the Atmosphere from Fires in the Tropics, Based on FAO Statistics for the Period 1975–1980. In J.G. Goldammer, editor, *Fires in Tropical Biota*, pages 440–462. Springer-Verlag, Berlin, 1990. ISBN 9783540521150.

[137] H.G.E. Hentschel and I. Procaccia. The Infinite Number of Generalized Dimensions of Fractals and Strange Attractors. *Physica D*, 8(3):435–444, 1983. doi: 10.1016/0167-2789(83)90235-X. URL `http://dx.doi.org/10.1016/0167-2789(83)90235-X`.

[138] S. Hergarten. *Self-Organized Criticality in Earth Systems*. Springer-Verlag Berlin Heidelberg, 2002. ISBN 978-3-540-43452-8.

[139] A.S. Hering, C.L. Bell, and M.G. Genton. Modeling Spatio-Temporal Wildfire Ignition Point Patterns. *Environmental and Ecological Statistics*, 16(2):225–250, 2009. doi: 10.1007/s10651-007-0080-6. URL `http://dx.doi.org/10.1007/s10651-007-0080-6`.

[140] P.F. Hessburg, J.K. Agee, and J.F. Franklin. Dry Forests and Wildland Fires of the Inland Northwest USA: Contrasting the Landscape Ecology of the Pre-Settlement and Modern Eras. *Forest Ecology and Management*, 211:117–139, 2005. doi: l0.1016/j.foreco.2005.02.0. URL `http://dx.doi.org/l0.1016/j.foreco.2005.02.0`.

[141] F. Höchtl, S. Lehringer, and W. Konold. “Wilderness”: What It Means When It Becomes a Reality — A Case Study From The Southwestern Alps. *Landscape and Urban Planning*, 70(1–2):85–95, 2005. ISSN 0169-2046. doi: 10.1016/j.landurbplan.2003.10.006. URL `http://www.sciencedirect.com/science/article/pii/S0169204603002123`.

[142] S.H. Hurlbert. Spatial Distribution of the Montane Unicorn. *Oikos*, 58(3):257–271, 1990. ISSN 00301299. URL `http://www.jstor.org/stable/3545216`.

[143] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley & Sons Ltd, Chichester, 2008. ISBN 978-0-470-01491-2.

[144] H.E. Isaaks and R.M. Srivastava. *An Introduction to Applied Geostatistics*. Oxford University Press, New York, 1989.

[145] V. Isham. Spatial Point Process Models. In A.E. Gelfand, P.J. Diggle, M. Fuentes, and P. Guttorp, editors, *Handbook of Spatial Statistics*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, pages 283–298. Chapman & Hall/CRC Press, Boca Raton, 2010. ISBN 978-1-4200-7287-7.

[146] E.T. Jaynes. Information Theory and Statistical Mechanics. *The Physical Review*, 106(4):620–630, 1957. doi: 10.1103/PhysRev.106.620. URL `http://link.aps.org/doi/10.1103/PhysRev.106.620`.

[147] D. Joshi, A. Samal, and L.-K. Soh. Spatio-Temporal Polygonal Clustering with Space and Time as First-Class Citizens. *GeoInformatica*, 17(2):387–412, 2013.

ISSN 1384-6175. doi: 10.1007/s10707-012-0157-8. URL `http://dx.doi.org/10.1007/s10707-012-0157-8`.

[148] C. Kaiser, M. Kanevski, A. Da Cunha, and V. Timonin. Emergence of Swiss Metropole and Scaling Properties of Urban Clusters. In *S4 International Conference on Emergence in Geographical Space, 23-25 November*, pages 23–25, Paris, 2009. S4. URL `http://www.google.ch/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CCcQFjAB&url=http%3A%2F%2Fs4.csregistry.org%2Ftiki-download_file.php%3FfileId%3D88&ei=VgyMVcS-NIHpUvypofgF&usg=AFQjCNHqwV5R-Gh-rSlCPfFOuuTdsjmT0w&bvm=bv.96782255,d.bGg`.

[149] P. Kaltenrieder, G. Procacci, B. Vannière, and W. Tinner. Vegetation and Fire History of the Euganean Hills (Colli Euganei) as Recorded by Lateglacial and Holocene Sedimentary Series from Lago della Dosta (Northeastern Italy). *The Holocene*, 20(5):679–695, 2010. doi: 10.1177/0959683609358911. URL `http://hol.sagepub.com/content/20/5/679.abstract`.

[150] M. Kanevski. *Advanced Mapping of Environmental Data: Geostatistics, Machine Learning and Bayesian Maximum Entropy.* Iste/Wiley, London, 2008.

[151] M. Kanevski and M. Maignan. *Analysis and Modelling of Spatial Environmental Data.* Environmental sciences. EPFL Press, Lausanne, 2004. ISBN 9780824759810.

[152] I. Karafyllidis. Design of a Dedicated Pparallel Processor for the Prediction of Forest Fire Spreading Using Cellular Automata and Genetic Algorithms. *Engineering Applications of Artificial Intelligence*, 17(1):19–36, 2004. doi: 10.1016/j.engappai.2003.12.001. URL `http://dx.doi.org/10.1016/j.engappai.2003.12.001`.

[153] A. Karr. *Point Processes and their Statistical Inference.* Marcel Dekker, New York, second edition, 1991.

[154] M.A. Krawchuk, M.A. Moritz, M-A Parisien, J. Van Dorn, and K. Hayhoe. Global Pyrogeography: the Current and Future Distribution of Wildfire. *PLoS ONE*, 4 (4):e5102, 2009. doi: 10.1371/journal.pone.0005102. URL `http://dx.doi.org/10.1371%2Fjournal.pone.0005102`.

[155] P. Krebs, G.B. Pezzatti, S. Mazzoleni, L.M. Talbot, and M. Conedera. Fire Regime: History and Definition of a Key Concept in Disturbance Ecology. *Theory in Biosciences*, 129(1):53–69, 2010. doi: 10.1007/s12064-010-0082-z. URL `http://dx.doi.org/10.1007/s12064-010-0082-z`.

[156] M. Kulldorff. A Spatial Scan Statistic. *Communications in Statistics - Theory and Methods*, 26(6):1481–1496, 1997. doi: 10.1080/03610929708831995. URL `http://dx.doi.org/10.1080/03610929708831995`.

[157] M. Kulldorff. Spatial Scan Statistics: Models, Calculations, and Applications. In J. Glaz and N. Balakrishnan, editors, *Scan Statistics and Applications*, Statistics for Industry and Technology, pages 303–322. Birkhäuser Boston, 1999. ISBN 978-1-4612-7201-4. doi: 10.1007/978-1-4612-1578-3_14. URL http://dx.doi.org/10.1007/978-1-4612-1578-3_14.

[158] M. Kulldorff. SaTScan^{TM} User Guide for version 9.4. Technical Documentation. Technical report, Kulldorff, M. and Management Services Inc, 2015. URL http://www.satscan.org/cgi-bin/satscan/register.pl/SaTScan_Users_Guide.pdf?todo=process_userguide_download.

[159] M. Kulldorff, W.F. Athas, E.J. Feurer, B.A. Miller, and C.R. Key. Evaluating Cluster Alarms: A Space-Time Scan Statistic and Brain Cancer in Los Alamos, New Mexico. *American Journal of Public Health*, 88(9):1377–1380, 1998. URL http://www.ncbi.nlm.nih.gov/pubmed/9736881.

[160] M. Kulldorff, R. Heffernan, J. Hartman, R. Assunção, and F. Mostashari. A Space–Time Permutation Scan Statistic for Disease Outbreak Detection. *PLoS Medicine*, 2(3):e59, 2005. doi: 10.1371/journal.pmed.0020059. URL http://dx.doi.org/10.1371/journal.pmed.0020059.

[161] C. Lampin-Maillet, M. Jappiot, M. Long, D. Morge, and J.P. Ferrier. Characterization and Mapping of Dwelling Types for Forest Fire Prevention. *Computers, Environment and Urban Systems*, 33(3):224–232, 2009. doi: 10.1016/j.compenvurbsys.2008.07.003. URL http://dx.doi.org/10.1016/j.compenvurbsys.2008.07.003.

[162] C. Lampin-Maillet, M. Jappiot, M. Long, Ch. Bouillon, D. Morge, and J.P. Ferrier. Mapping Wildland-Urban Interfaces at Large Scales Integrating Housing Density and Vegetation Aggregation for Fire Prevention in the South of France. *Journal of Environmental Management*, 91(3):732–741, 2010. doi: 10.1016/j.jenvman.2009.10.001. URL http://dx.doi.org/10.1016/j.jenvman.2009.10.001.

[163] C. Lampin-Maillet, M. Long-Fournel, A. Ganteaume, M. Jappiot, and J.P. Ferrier. Land Cover Analysis in Wildland–Urban Interfaces According to Wildfire Risk: A Case Study in the South of France. *Forest Ecology and Management*, 261(12):2200–2213, 2011. doi: 10.1016/j.foreco.2010.11.022. URL http://www.sciencedirect.com/science/article/pii/S0378112710006924.

[164] R. Lasaponara, A. Santulli, and L. Telesca. Time-Clustering Analysis of Forest-Fire Sequences in Southern Italy. *Chaos, Solitons & Fractals*, 24(1):139–149, 2005. doi: 10.1016/j.chaos.2004.07.025. URL http://dx.doi.org/10.1016/j.chaos.2004.07.025.

[165] D. Latham and E. Williams. Lightning and Forest Fires. In E.A. Johnson and K. Miyanishi, editors, *Forest Fires: Behaviour and Ecological Effects*, chapter 11, pages 375–418. Academic Press, San Diego, 2001. ISBN 978-0-12-386660-8.

[166] R. Law, J. Illian, D.F.R.P. Burslem, G. Gratzer, C.V.S. Gunatilleke, and I.A.U.N. Gunatilleke. Ecological Information from Spatial Patterns of Plants: Insights from Point Process Theory. *Journal of Ecology*, 97(4):616–628, 2009. ISSN 1365-2745. doi: 10.1111/j.1365-2745.2009.01510.x. URL `http://dx.doi.org/10.1111/j.1365-2745.2009.01510.x`.

[167] C.-K. Lee. Multifractal Characteristics in Air Pollutant Concentration Time Series. *Water, Air, and Soil Pollution*, 135(1-4):389–409, 2002. ISSN 0049-6979. doi: 10.1023/A:1014768632318. URL `http://dx.doi.org/10.1023/A%3A1014768632318`.

[168] A.S. Lefohn, H.P. Knudsen, J.A. Logan, J. Simpson, and C. Bhumralkar. An Evaluation of the Kriging Method to Predict 7-h Seasonal Mean Ozone Concentrations for Estimating Crop Losses. *JAPCA*, 37(5):595–602, 1987. doi: 10.1080/08940630.1987.10466247. URL `http://dx.doi.org/10.1080/08940630.1987.10466247`.

[169] A. Liaw and M. Wiener. Classification and Regression by randomForest. *R News*, 2(3):18–22, 2002. URL `http://CRAN.R-project.org/doc/Rnews/`.

[170] R. Lopes and N. Betrouni. Fractal and Multifractal Analysis: A Review. *Medical Image Analysis*, 13(4):643–649, 2009. doi: 10.1016/j.media.2009.05.003. URL `http://www.medicalimageanalysisjournal.com/article/S1361-8415%2809%2900039-5/abstract`.

[171] S. Lovejoy, D. Schertzer, and P. Ladoy. Fractal Charaterization of Inhomogeneous Geophysical Measuring Networks. *Nature*, 319(6048):43–44, 1986. doi: 10.1038/319043a0. URL `http://www.nature.com/nature/journal/v319/n6048/abs/319043a0.html`.

[172] S.B. Lowen and M.C. Teich. Estimation and Simulation of Fractal Stochastic Point Processes. *Fractals*, 3(1):183–210, 1995. doi: 10.1142/S0218348X95000151. URL `http://dx.doi.org/10.1142/S0218348X95000151`.

[173] B.D. Malamud, G. Morein, and D.L. Turcotte. Forest Fires: An Example of Self-Organized Critical Behavior. *Science*, 281(5384):1840–1842, 1998. doi: 10.1126/science.281.5384.1840. URL `http://dx.doi.org/10.1126/science.281.5384.1840`.

[174] B.D. Malamud, D.L. Turcotte, F. Guzzetti, and P. Reichenbach. Landslide Inventories and Their Statistical Properties. *Earth Surface Processes and Landforms*, 29(6):687–711, 2004. ISSN 1096-9837. doi: 10.1002/esp.1064. URL `http://dx.doi.org/10.1002/esp.1064`.

[175] B.D. Malamud, J.D.A. Millington, and G.L.W. Perry. Characterizing Wildfire Regimes in the United States. *Proceedings of the National Academy of Sciences of the United States of America*, 102(13):4694–4699, 2005. doi: 10.1073/pnas.0500880102. URL `http://dx.doi.org/10.1073/pnas.0500880102`.

[176] D. Mandallaz and R. Ye. Prediction of Forest Fires with Poisson Models. *Canadian Journal of Forest Research*, 27(10):1685–1694, 1997. doi: 10.1139/x97-103. URL `http://www.nrcresearchpress.com/doi/abs/10.1139/x97-103`.

[177] B. Mandelbrot. How Long is the Coast of Bretain? Statistical Self-Similarity and Fractional Dimension. *Science*, 156(3775):636–638, 1967. doi: 10.1126/science.156.3775.636. URL `http://dx.doi.org/10.1126/science.156.3775.636`.

[178] B.B Mandelbrot. *The Fractal Geometry of Nature*. Freeman and Company, New York, 1983.

[179] B.B. Mandelbrot. An Introduction to Multifractal Distribution Functions. In H.Eugene Stanley and Nicole Ostrowsky, editors, *Random Fluctuations and Pattern Growth: Experiments and Models*, volume 157 of *NATO ASI Series*, pages 279–291. Springer Netherlands, Dordrecht,The Netherlands, 1988. ISBN 978-0-7923-0073-1. doi: 10.1007/978-94-009-2653-0_40. URL `http://dx.doi.org/10.1007/978-94-009-2653-0_40`.

[180] M. Marcozzi, M. Conedera, and B. Jud. Forest Fire Research in Switzerland. *International Forest Fire News*, 12:32–33, 1995. URL `http://www.slf.ch/it/bellinzona/insubrisch/publications/pubblicazioni_incendi/Marcozzi_et_al_1995_IntForFire`.

[181] M. Martínez, C. Vega-Garcia, and E. Chuvieco. Human-Caused Wildfire Risk Rating for Prevention Planning in Spain. *Journal of Environmental Management*, 90(2):1241–1252, 2009. doi: 10.1016/j.jenvman.2008.07.005. URL `http://www.sciencedirect.com/science/article/pii/S0301479708001758`.

[182] A. Massada, A.D. Syphard, S.I. Stewart, and V.C. Radeloff. Wildfire Ignition-Distribution Modelling: a Comparative Study in the Huron–Manistee National Forest, Michigan, USA. *International Journal of Wildland Fire*, pages 174—183, 2013. doi: 10.1071/WF11178. URL `http://www.publish.csiro.au/?paper=WF11178`.

[183] B. Matérn. *Spatial Variation: Stochastic Models and their Applications to some Problems in Forest Surveys and other Sampling Investigations*. Meddelanden från Statens Skogsforskningsinstitut: Statens Skogsforskningsinstitut. Skogsforskningsinstiut, 1960.

[184] B. Matérn. *Spatial Variation*. Lecture Notes in Statistics. Springer-Verlag, 1986. ISBN 9783540963653. URL `http://books.google.ch/books?id=s-xczaXRptoC`.

[185] G. Matheron. Principles of Geostatistics. *Economic Geology*, 58(8):1246–1266, 1963. doi: 10.2113/gsecongeo.58.8.1246. URL `http://dx.doi.org/10.2113/gsecongeo.58.8.1246`.

[186] T.K. McGee. Urban Residents' Approval of Management Measure to Mitigate Wildland-Urban Interface Fire Riks in Edmonton, Canada. *Landscape and Urban Planning*, 82(4):247–256, 2007. doi: 10.1016/j.landurbplan.2007.03.001. URL `http://dx.doi.org/10.1016/j.landurbplan.2007.03.001`.

[187] P. Meakin, A. Coniglio, H.E. Stanley, and T.A. Witten. Scaling Properties for the Surfaces of Fractal and Non-Fractal Objects: an Infinite Hierarchy of Critical Exponents. *Physical Review A*, 34:3325, 1986. doi: 10.1103/PhysRevA.34.3325. URL `http://inspirehep.net/record/243897?ln=en`.

[188] W.E. Mell, S.L. Manzello, A. Maranghides, D. Butry, and R.G. Rehm. The Wildland-Urban Interface Fire Problem - Current Approaches and Research Needs. *International Journal of Wildland Fire*, 19(2):238–251, 2010. doi: 10.1071/WF07131. URL `http://dx.doi.org/10.1071/WF07131`.

[189] J.M. Mendes, P.C. de Zea Bermudez, J. Pereira, K.F. Turkman, and M.J.P. Vasconcelos. Spatial Extremes of Wildfire Sizes: Bayesian Hierarchical Models for Extremes. *Environmental and Ecological Statistics*, 17(1):1–28, 2010. ISSN 1352-8505. doi: 10.1007/s10651-008-0099-3. URL `http://dx.doi.org/10.1007/s10651-008-0099-3`.

[190] M. Mermoz, T. Kitzberger, and T. Veblen. Landscape Influences on Occurrence and Spread of Wildfires in Patagonian Forests and Shrublands. *Ecology*, 86(10): 2705–2715, 2005. doi: 10.1890/04-1850. URL `http://dx.doi.org/10.1890/04-1850`.

[191] N. Micheletti, L. Foresti, S. Robert, M. Leuenberger, A. Pedrazzini, M. Jaboyedoff, and M. Kanevski. Machine Learning Feature Selection Methods for Landslide Susceptibility Mapping. *Mathematical Geosciences*, 46(1):33–57, 2014. ISSN 1874-8961. doi: 10.1007/s11004-013-9511-0. URL `http://dx.doi.org/10.1007/s11004-013-9511-0`.

[192] J. Møller and R.P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Chapman & Hall/CRC Press, Boca Raton, 2004.

[193] J. Møller and R.P. Waagepetersen. Modern Statistics for Spatial Point Processes. *Scandinavian Journal of Statistics*, 34(4):643–684, 2007. doi: 10.1111/j.1467-9469.2007.00569.x. URL `http://dx.doi.org/10.1111/j.1467-9469.2007.00569.x`.

[194] P.A.P. Moran. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1-2):17–23, 1950. doi: 10.1093/biomet/37.1-2.17. URL `http://biomet.oxfordjournals.org/content/37/1-2/17.short`.

[195] M.V. Moreno, B.D. Malamud, and E. Chuvieco. Wildfire Frequency-Area Statistics in Spain. *Procedia Environmental Sciences*, 7(0):182–187, 2011. doi: 10.1016/j.proenv.2011.07.032. URL `http://www.sciencedirect.com/science/article/pii/S1878029611001599`.

[196] M.A. Moretti. *Effect of Fire on the Invertebrate Communities in Chestnut Forests in Southern Switzerland*. Doctoral and habilitation theses, ETH Zürich, Zürich, 2004. URL `http://dx.doi.org/10.3929/ethz-a-004751917`.

[197] P. Morgan, G.H. Aplet, J.B. Haufler, H.C. Humphries, M.M. Moore, and W.D. Wilson. Historical Range of Variability: A Useful Tool for Evaluating Ecosystem Change. *Journal of Sustainable Forestry*, 2(1-2):87–111, 1994. doi: 10.1300/J091v02n01_04. URL `http://dx.doi.org/10.1300/J091v02n01_04`.

[198] M. Morisita. Measuring the Dispersion of Individuals and Analysis of the Distributional Patterns. *Memoires of the Faculty of Science, Kyushu University, Series E (Biology)*, 2(4):215–235, 1959.

[199] M.A. Moritz. Analyzing Extreme Disturbance Events: Fire in Los Padres National Forest. *Ecological Applications*, 7(4):1252–1262, 1997. doi: 10.1890/1051-0761(1997)007[1252:AEDEFI]2.0.CO;2. URL `http://dx.doi.org/10.1890/1051-0761(1997)007[1252:AEDEFI]2.0.CO;2`.

[200] M.A. Moritz, M.A. Parisien, E. Batllori, M.A. Krawchuk, J. Van Dorn, D.J. Ganz, and K. Hayhoe. Climate Change and Disruptions to Global Fire Activity. *Ecosphere*, 3(6):art.49, 2012. doi: 10.1890/ES11-00345.1. URL `http://dx.doi.org/10.1890/ES11-00345.1`.

[201] A. Nandi and A. Shakoor. A GIS-Based Landslide Susceptibility Evaluation Using Bivariate and Multivariate Statistical Analyses. *Engineering Geology*, 110(1–2):11–20, 2010. ISSN 0013-7952. doi: http://dx.doi.org/10.1016/j.enggeo.2009.10.001. URL `http://www.sciencedirect.com/science/article/pii/S0013795209002646`.

[202] J.I. Naus. Clustering of Random Points in Two Dimensions. *Biometrika*, 52(1-2):263–266, 1965. doi: 10.1093/biomet/52.1-2.263. URL `http://dx.doi.org/10.1093/biomet/52.1-2.263`.

[203] J.I. Naus. The Distribution of the Size of the Maximum Cluster of Points on a Line. *Journal of the American Statistical Association*, 60(310):532–538, 1965. doi: 10.1080/01621459.1965.10480810. URL `http://dx.doi.org/10.1080/01621459.1965.10480810`.

[204] US Department of Agriculture and US Department of the Interior. Urban Wildland Interface Communities within Vicinity of Federal Lands That Are at

High Risk from Wildfire. *Federal Register*, 66(3):751–777, 2001. URL https://federalregister.gov/a/01-52.

[205] Swiss Federal Statistical Office and Swiss Federal Department of Home Affairs. Forestry in Switzerland. Pocket Statistics 2009. Technical report, Swiss Confederation, Neuchâtel, 2009.

[206] S. Oliveira, F. Oehler, J. San-Miguel-Ayanz, A. Camia, and J.M.C. Pereira. Modeling Spatial Patterns of Fire Occurrence in Mediterranean Europe Using Multiple Regression and Random Forest. *Forest Ecology and Management*, 275:117–129, 2012. ISSN 0378-1127. doi: 10.1016/j.foreco.2012.03.003. URL http://www.sciencedirect.com/science/article/pii/S0378112712001272.

[207] S. Openshaw, M. Charlton, C. Wymer, and A. Craft. A Mark 1 Geographical Analysis Machine for the Automated Analysis of Point Data Sets. *International Journal of Geographical Information Systems*, 1(4):335–358, 1987. doi: 10.1080/02693798708927821. URL http://dx.doi.org/10.1080/02693798708927821.

[208] J. Ozik, B. Hunt, and E. Ott. Formation of Multifractal Population Patterns from Reproductive Growth and Local Resettlement. *Physical Review E*, 72(4):046213, 2005. doi: 10.1103/PhysRevE.72.046213. URL http://link.aps.org/doi/10.1103/PhysRevE.72.046213.

[209] G. Paladin and A. Vulpiani. Anomalous Scaling Laws in Multifractal Objects. *Physics Reports (Review Section of Physics Letters)*, 156(4):147–225, 1987. doi: doi:10.1016/0370-1573(87)90110-4. URL http://dx.doi.org/doi:10.1016/0370-1573(87)90110-4.

[210] J.G. Pausas and J.E. Keeley. A Burning Story: The Role of Fire in the History of Life. *BioScience*, 59(7):593–601, 2009. doi: 10.1525/bio.2009.59.7.10. URL http://bioscience.oxfordjournals.org/content/59/7/593.abstract.

[211] R.D. Peng, F.P. Schoenberg, and J.A. Woods. A Space-Time Conditional Intensity Model for Evaluating a Wildfire Hazard Index. *Journal of the American Statistical Association*, 100:26–35, 2005. URL http://EconPapers.repec.org/RePEc:bes:jnlasa:v:100:y:2005:p:26-35.

[212] M.G. Pereira, R.M. Trigo, C.C. da Camara, J.M.C. Pereira, and S.M. Leite. Synoptic Patterns Associated with Large Summer Forest Fires in Portugal. *Agricultural and Forest Meteorology*, 129(1–2):11–25, 2005. doi: 10.1016/j.agrformet.2004.12.007. URL http://dx.doi.org/10.1016/j.agrformet.2004.12.007.

[213] M.G. Pereira, B.D. Malamud, R.M. Trigo, and P.I. Alves. The History and Characteristics of the 1980–2005 Portuguese Rural Fire Database. *Natural Hazards and Earth System Science*, 11(12):3343–3358, 2011. doi: 10.5194/nhess-11-3343-2011. URL http://www.nat-hazards-earth-syst-sci.net/11/3343/2011/.

[214] M.G. Pereira, L. Caramelo, C.D. Vega Orozco, R. Costa, and M. Tonini. Space-Time Clustering Analysis Performance of an Aggregated Dataset: The Case of Wildfires in Portugal. *Environmental Modelling and Software*, 72:239–249, 2015. doi: 10.1016/j.envsoft.2015.05.016. URL http://www.sciencedirect.com/science/article/pii/S1364815215001644.

[215] P. Pereira, K.F. Turkman, M.A. Amaral Turkman, A. Sá, and J.M.C. Pereira. Quantification of Annual Wildfire Risk; A Spatio-Temporal Point Process Approach. *Statistica*, 73:55–68, 2013. doi: 10.6092/issn.1973-2201/3985. URL http://rivista-statistica.unibo.it/article/view/3985.

[216] E. Perfect, R.W. Gentry, M.C. Sukop, and J.E. Lawson. Multifractal Sierpinski Carpets: Theory and Application to Upscaling Effective Saturated Hydraulic Conductivity. *Geoderma*, 134(3-4):240–252, 2006. doi: 10.1016/j.geoderma.2006.03.001. URL http://dx.doi.org/10.1016/j.geoderma.2006.03.001.

[217] G.B. Pezzatti, M. Conedera, and A. Kaltenbrunner. Die Neue Waldbranddatenbank. *Bündnerwald*, 58(6):37–39, 2005.

[218] G.B. Pezzatti, S. Bajocco, D. Torriani, and M. Conedera. Selective Burning of Forest Vegetation in Canton Ticino (Southern Switzerland). *Plant Biosystems*, 143(3):609–620, 2009. doi: 10.1080/11263500903233292. URL http://dx.doi.org/10.1080/11263500903233292.

[219] G.B. Pezzatti, M. Reinhard, and M. Conedera. Swissfire: die Neue Schweizerische Waldbranddatenbank | Swissfire: the New Swiss Forest Fire Database. *Schweiz Z Forstwes*, 161(12):465–469, 2010. doi: 10.3188/szf.2010.0465. URL http://dx.doi.org/10.3188/szf.2010.0465.

[220] G.B. Pezzatti, T. Zumbrunnen, M. Bürgi, P. Ambrosetti, and M. Conedera. Fire Regime Shifts as a Consequence of Fire Policy and Socio-Economic Development: An Analysis Based On The Change Point Approach. *Forest Policy and Economics*, 29(0):7–18, 2013. ISSN 1389-9341. doi: 10.1016/j.forpol.2011.07.002. URL http://www.sciencedirect.com/science/article/pii/S1389934111001043. The FIRE PARADOX project: Setting the basis for a shift in the forest fire policies in Europe.

[221] R.V. Platt. The Wildland-Urban Interface: Evaluating the Definition Effect. *Journal of Forestry*, 108(1):9–15, 2010. URL http://www.ingentaconnect.com/content/saf/jof/2010/00000108/00000001/art00005.

[222] S.R. Plevel. Fire Policy at the Wildland-Urban Interface: A Local Responsibility. *Journal of Forestry*, 95(10):12–17, 1997. URL http://www.ingentaconnect.com/content/saf/jof/1997/00000095/00000010/art00007.

[223] R.E. Plotnick, R.H. Gardner, W.W. Hargrove, K. Prestegaard, and M. Perlmutter. Lacunarity Analysis: A General Technique for the Analysis of Spatial Patterns. *Physical Review E*, 53(5):5461–5468, 1996. doi: 10.1103/PhysRevE.53.5461. URL `http://link.aps.org/doi/10.1103/PhysRevE.53.5461`.

[224] J. Podur, D.L. Martell, and F. Csillag. Spatial Patterns of Lightning-Caused Forest Fires in Ontario, 1976-1998. *Ecological Modelling*, 164(1):1–20, 2003. doi: 10.1016/S0304-3800(02)00386-1. URL `http://dx.doi.org/10.1016/S0304-3800(02)00386-1`.

[225] B. Porterie, J.L. Consalvi, J.C. Loraud, F. Giroud, and C. Picard. Dynamics of Wildland Fires and Their Impact on Structures. *Combustion and Flame*, 149(3):314–328, 2007. doi: 10.1016/j.combustflame.2006.12.017. URL `http://dx.doi.org/10.1016/j.combustflame.2006.12.017`.

[226] J.P. Prestemon and D.T. Butry. Time to Burn: Modeling Wildland Arson as an Autoregressive Crime Function. *American Journal of Agricultural Economics*, 87(3):756–770, 2005. URL `http://www.jstor.org/stable/3697911`.

[227] J.P. Prestemon, J.M. Pye, D.T. Butry, T.P. Holmes, and D.E. Mercer. Understanding Broadscale Wildfire Risks in a Human-Dominated Landscape. *Forest Science*, 48(4):685–693, 2002. URL `http://www.ingentaconnect.com/content/saf/fs/2002/00000048/00000004/art00006`.

[228] S.J. Pyne. *World Fire: The Culture of Fire on Earth.* Cycle of fire. University of Washington Press, 1997. ISBN 9780295975931. URL `https://books.google.ch/books?id=55EVdnVWFjIC`.

[229] S.J. Pyne, P.L. Andrews, and R.D. Laven. *Introduction to Wildland Fire.* Wiley, New York, second edition, 1996. ISBN 978-0471549130.

[230] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2013. URL `http://www.R-project.org/`.

[231] V.C. Radeloff, R.B. Hammer, and S.I. Stewart. Rural and Suburban Sprawl in the U.S. Midwest from 1940 to 2000 and its Relation to Forest Fragmentation. *Conservation Biology*, 19(3):793–805, 2005. doi: 10.1111/j.1523-1739.2005.00387.x. URL `http://dx.doi.org/10.1111/j.1523-1739.2005.00387.x`.

[232] V.C. Radeloff, R.B. Hammer, S.I. Stewart, J.S. Fried, S.S. Holcomb, and J.F. McKeefry. The Wildland Urban Interface in the United States. *Ecological Appications*, 15(3):799–805, 2005. doi: 10.1890/04-1413. URL `http://dx.doi.org/10.1890/04-1413`.

[233] M.A. Reams, T.K. Haines, C.R. Renner, M.W. Wascom, and H. Kingre. Goals, Obstacles and Effective Strategies of Wildfire Mitigation Programs in the Wildland-Urban Interface. *Forest Policy and Economics*, 7(5):818–826, 2005. doi: 10.1016/j.forpol.2005.03.006. URL `http://dx.doi.org/10.1016/j.forpol.2005.03.006`.

[234] M. Rebetez. Twentieth Century Trends in Droughts in Southern Switzerland. *Geophysical Research Letters*, 26(6):755–758, 1999. doi: 10.1029/1999GL900075. URL `http://dx.doi.org/10.1029/1999GL900075`.

[235] W.J. Reed and K.S. McKelvey. Power-Law Behaviour and Parametric Models for the Size-Distribution of Forest Fires. *Ecological Modelling*, 150(3):239–254, 2002. doi: 10.1016/S0304-3800(01)00483-5. URL `http://dx.doi.org/10.1016/S0304-3800(01)00483-5`.

[236] B. Reineking, P. Weibel, M. Conedera, and H. Bugmann. Environmental Determinants of Lightning- v. Human-Induced Forest Fire Ignitions Differ in a Temperate Mountain Region of Switzerland. *International Journal of Wildland Fire*, 19(5): 541–557, 2010. doi: 10.1071/WF08206. URL `http://www.publish.csiro.au/paper/WF08206.htm`.

[237] M. Reinhard, M. Rebetez, and R. Schlaepfer. Recent Climate Change: Rethinking Drough in the Context of Forest Fire Research in Ticino, South of Switzerland. *Theoretical and Applied Climatology*, 82(1-2):17–25, 2005. doi: 10.1007/s00704-005-0123-6. URL `http://dx.doi.org/10.1007/s00704-005-0123-6`.

[238] C. Ricotta, G. Avena, and M. Marchetti. The Flaming Sandpile: Self-Organized Criticality and Wildfires. *Ecological Modelling*, 119(1):73–77, 1999. doi: 10.1016/S0304-3800(99)00057-5. URL `http://dx.doi.org/10.1016/S0304-3800(99)00057-5`.

[239] B.D. Ripley. The Second-Order Analysis of Stationary Point Processes. *Journal of Applied Probability*, 13(2):255–266, 1976. doi: 10.2307/3212829. URL `http://www.jstor.org/stable/3212829`.

[240] B.D. Ripley. Modelling Spatial Patterns. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(2):172–212, 1977. doi: 10.2307/2984796. URL `http://www.jstor.org/stable/2984796`.

[241] B.D. Ripley. *Spatial Statistics*. Wiley, New York, 1981.

[242] B.D. Ripley. *Statistical Inference for Spatial Processes*. Cambridge University Press, 1988. ISBN 9780511624131. URL `http://dx.doi.org/10.1017/CBO9780511624131`.

[243] B.D. Ripley. *Spatial Statistics*. Wiley Series in Probability and Statistics. Wiley, 2005. ISBN 9780471725206. URL `https://books.google.ch/books?id=eB_XYz3B7qQC`.

[244] A. Rényi. *Probability Theory.* Akadémiai Kiadò, Budapest, 1970.

[245] G.P. Robertson. Geostatistics in Ecology: Interpolating with Known Variance. *Ecology*, 68(3):744–748, 1987. URL http://www.jstor.org/stable/1938482.

[246] I. Rodríguez-Iturbe and A. Rinaldo. *Fractal River Basins: Chance and Self-Organization.* Cambridge University Press, UK, 1997. ISBN 9780521004053. URL http://www.cambridge.org/ch/academic/subjects/earth-and-environmental-science/hydrology-hydrogeology-and-water-resources/fractal-river-basins-chance-and-self-organization.

[247] M.L. Rorig and S.A. Ferguson. Characteristics of Lightning and Wildland Fire Ignition in the Pacific Northwest. *Journal of Applied Meteorology*, 38:1565–1575, 1999. doi: 10.1175/1520-0450(1999)038⟨1565:COLAWF⟩2.0.CO;2. URL http://dx.doi.org/10.1175/1520-0450(1999)038<1565:COLAWF>2.0.CO;2.

[248] D.A. Russell, J.D. Hanson, and E. Ott. Dimension of Strange Attractors. *Physical Review Letters*, 45(14):1175–1178, 1980. doi: 10.1103/PhysRevLett.45.1175. URL http://link.aps.org/doi/10.1103/PhysRevLett.45.1175.

[249] G. Salvadori, S.P. Ratti, and G. Belli. Fractal and Multifractal Approach to Environmental Pollution. *Environmental Science and Pollution Research*, 4(2): 91–98, 1997. ISSN 0944-1344. doi: 10.1007/BF02986286. URL http://dx.doi.org/10.1007/BF02986286.

[250] R.C. Sambrook and R.F. Voss. Fractal Analysis of US Settlement Patterns. *Fractals*, 9(3):241–250, 2001. doi: 10.1142/S0218348X01000749. URL http://www.worldscientific.com/doi/abs/10.1142/S0218348X01000749.

[251] J. Sander, M. Ester, H.P. Kriegel, and X. Xu. Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and its Applications. *Data Mining and Knowledge Discovery*, 2(2):169–194, 1998. ISSN 1384-5810. doi: 10.1023/A:1009745219419. URL http://dx.doi.org/10.1023/A:1009745219419.

[252] C. Schär, P.L. Vidale, D. Lüthi, C. Frei, C. Häberli, M.A. Liniger, and C. Appenzeller. The Role of Increasing Temperature Variability in European Summer Heatwaves. *Nature*, 427:332–336, 2004. doi: 10.1038/nature02300. URL http://dx.doi.org/10.1038/nature02300.

[253] S. Schumacher. *The Role of Large-Scales Disturbances and Climate for the Dynamics of Forested Landscapes in the European Alps.* Doctoral and habilitation theses, ETH Zürich, Zürich, 2004. URL http://dx.doi.org/10.3929/ethz-a-004818825.

[254] S. Schumacher and H. Bugmann. The Relative Importance of Climatic Effects, Wildfires and Management for Future Forest Landscape Dynamics in the Swiss Alps. *Global Change Biology*, 12(8):1435–1450, 2006. doi: 10.1111/j.1365-2486.2006.01188.x. URL `http://dx.doi.org/10.1111/j.1365-2486.2006.01188.x`.

[255] A. Sebastián-López, R. Salvador-Civil, J. Gonzalo-Jiménez, and J. SanMiguel-Ayanz. Integration of Socio-Economic and Environmental Variables for Modelling Long-Term Fire Danger in Southern Europe. *European Journal of Forest Research*, 127(2):149–163, 2008. ISSN 1612-4669. doi: 10.1007/s10342-007-0191-5. URL `http://dx.doi.org/10.1007/s10342-007-0191-5`.

[256] L. Serra, P. Juan, D. Varga, J. Mateu, and M. Saez. Spatial Pattern Modelling of Wildfires in Catalonia, Spain 2004–2008. *Environmental Modelling & Software*, 40(0):235–244, 2013. ISSN 1364-8152. doi: 10.1016/j.envsoft.2012.09.014. URL `http://www.sciencedirect.com/science/article/pii/S1364815212002460`.

[257] L. Seuront. *Fractals and Multifractals in Ecology and Aquatic Science*. CRC Press, Boca Raton, 2010.

[258] T.G. Smith Jr., W.B. Marks, G.D. Lange, W.H. Sheriff Jr., and E.A. Neale. A Fractal Analysis of Cell Images. *Journal of Neuroscience Methods*, 27(2):173–180, 1989. ISSN 0165-0270. doi: 10.1016/0165-0270(89)90100-3. URL `http://www.sciencedirect.com/science/article/pii/0165027089901003`.

[259] W. Song, F. Weicheng, W. Binghong, and Z. Jianjun. Self-Organized Criticality of Forest Fire in China. *Ecological Modelling*, 145(1):61–68, 2001. doi: 10.1016/S0304-3800(01)00383-0. URL `http://dx.doi.org/10.1016/S0304-3800(01)00383-0`.

[260] W.P. Sousa. The Role of Disturbance in Natural Communities. *Annual Review of Ecology and Systematics*, 15:353–391, 1984. doi: 10.1146/annurev.es.15.110184.002033. URL `http://www.jstor.org/stable/2096953`.

[261] F. Spinedi and F. Isotta. Il Clima del Ticino. In Elio Venturelli et al., editor, *Dati - Statistiche e Società, Il Clima del Ticino negli Ultimi 50 Anni*, volume A.IV n.2, pages 4–39. Ustat, Bellinzona, June 2004. doi: 10.1007/s10707-012-0161-z. URL `http://dx.doi.org/10.1007/s10707-012-0161-z`.

[262] H. Stanley and P. Meakin. Multifractal Phenomena in Physics and Chemistry. *Nature*, 335(6189):405–409, 1988. doi: 10.1038/335405a0. URL `http://www.nature.com/nature/journal/v335/n6189/abs/335405a0.html`.

[263] S.I. Stewart, V.C. Radeloff, R.B. Hammer, and T.J. Hawbaker. Defining the Wildland-Urban Interface. *Journal of Forestry*, 105(4):201–207, 2007. URL `http://www.ingentaconnect.com/content/saf/jof/2007/00000105/00000004/art00012`.

[264] S.I. Stewart, B. Wilmer, R.B. Hammer, G.H. Aplet, T.J. Hawbaker, C. Miller, and V.C. Radeloff. Wildland-Urban Interface Maps Vary with Purpose and Context. *Journal of Forestry*, 107(2):78–83, 2009. URL http://www.ingentaconnect.com/content/saf/jof/2009/00000107/00000002/art00009.

[265] B.J. Stocks, B.D. Lawson, M.E. Alexander, C.E. Van Wagner, R.S. McAlpine, T.J. Lynham, and D.E. Dubé. The Canadian Forest Fire Danger Rating System: An Overview. *Forestry Chronicle*, 65(6):450–457, 1989. URL https://cfs.nrcan.gc.ca/publications?id=11347.

[266] D. Stoyan and H. Stoyan. Non-Homogeneous Gibbs Process Models for Forestry — A Case Study. *Biometrical Journal*, 40(5):521–531, 1998. ISSN 1521-4036. doi: 10.1002/(SICI)1521-4036(199809)40:5⟨521::AID-BIMJ521⟩3.0.CO;2-R. URL http://dx.doi.org/10.1002/(SICI)1521-4036(199809)40:5<521::AID-BIMJ521>3.0.CO;2-R.

[267] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. John Wiley & Sons, Ltd, Chichester, second edition, 1995.

[268] M. Stuber and M. Bürgi. Agrarische Waldnutzungen in der Schweiz 1800–1950. Waldweide, Waldheu, Nadel- und Laubfutter — Agricultural Use of Forest in Switzerland 1800–1950. Wood Pasture, Wood Hay Collection, and The Use of Leaves and Needles for Fodder. *Schweizerische Zeitschrift fur Forstwesen*, 152 (12):490–508, 2001. doi: 10.3188/szf.2001.0490. URL http://dx.doi.org/10.3188/szf.2001.0490.

[269] A.D. Syphard and J. Franklin. Species Traits Affect the Performance of Species Distribution Models for Plants in Southern California. *Journal of Vegetation Science*, 21(1):177–189, 2010. ISSN 1654-1103. doi: 10.1111/j.1654-1103.2009.01133.x. URL http://dx.doi.org/10.1111/j.1654-1103.2009.01133.x.

[270] A.D. Syphard, V.C. Radeloff, N.S. Keuler, R.S. Taylor, T.J. Hawbaker, S.I. Stewart, and M.K. Clayton. Predicting Spatial Patterns of Fire on a Southern California Landscape. *International Journal of Wildland Fire*, 17(5):602–613, 2008. URL http://silvis.forest.wisc.edu/publications/Predicting-spatial-patterns-fire-southern-California-landscape.

[271] C. Tannier and D. Pumain. Fractals in Urban Geography: A Theoretical Outline and An Empirical Example. *Cybergeo: European Journal of Geography [En ligne], Systèmes, Modélisation, Géostatistiques*, document 307, 2005. doi: 10.4000/cybergeo.3275. URL http://cybergeo.revues.org/3275.

[272] A.M. Tarquis, K.J. McInnes, J.R. Key, A. Saa, M.R. García, and M.C. Díaz. Multiscaling Analysis in a Structured Clay Soil Using 2D Images. *Journal of Hydrology*, 322(1–4):236–246, 2006. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2005.03.005. URL http://www.sciencedirect.com/science/article/pii/S0022169405001186.

[273] T. Tél, A. Fülöp, and T. Vicsek. Determination of Fractal Dimensions for Geometrical Multifractals. *Physica A: Statistical Mechanics and its Applications*, 159(2):155–166, 1989. doi: 10.1016/0378-4371(89)90563-3. URL http://dx.doi.org/10.1016/0378-4371(89)90563-3.

[274] L. Telesca and R. Lasaponara. Emergence of Temporal Regimes in Fire Sequences. *Physica A: Statistical Mechanics and its Applications*, 360(2):543–547, 2006. doi: 10.1016/j.physa.2005.04.045. URL http://dx.doi.org/10.1016/j.physa.2005.04.045.

[275] L. Telesca and M.G. Pereira. Time-Clustering Investigation of Fire Temporal Fluctuations in Portugal. *Natural Hazards and Earth System Sciences*, 10(4):661–666, 2010. doi: 10.5194/nhess-10-661-2010. URL http://dx.doi.org/10.5194/nhess-10-661-2010.

[276] L. Telesca, G. Amatucci, R. Lasaponara, M. Lovallo, and A. Santulli. Time-Scaling Properties in Forest-Fire Sequences Observed in Gargano Area (Southern Italy). *Ecological Modelling*, 185(2-4):531–544, 2005. doi: 10.1016/j.ecolmodel.2005.01.009. URL http://dx.doi.org/10.1016/j.ecolmodel.2005.01.009.

[277] L. Telesca, G. Amatucci, R. Lasaponara, M. Lovallo, and M.J. Rodrigues. Space-Time Fractal Properties of the Forest-Fire Series in Central Italy. *Communications in Nonlinear Science and Numerical Simulation*, 12(7):1326–1333, 2007. doi: 10.1016/j.cnsns.2005.12.003. URL http://dx.doi.org/10.1016/j.cnsns.2005.12.003.

[278] L. Telesca, G. Amatucci, R. Lasaponara, M. Lovallo, and A. Santulli. Identifying Spatial Clustering Properties of the 1997-2003 Liguria (Northern Italy) Forest-Fire Sequence. *Chaos, Solitons & Fractals*, 32(4):1364–1370, 2007. doi: 10.1016/j.chaos.2005.11.075. URL http://dx.doi.org/10.1016/j.chaos.2005.11.075.

[279] L. Telesca, M. Kanevski, M. Tonini, G.B. Pezzatti, and M. Conedera. Temporal Patterns of Fire Sequences Observed in Canton of Ticino (Southern Switzerland). *Natural Hazards and Earth System Sciences*, 10(4):723–728, 2010. doi: 10.5194/nhess-10-723-2010. URL http://dx.doi.org/10.5194/nhess-10-723-2010.

[280] D. Theobald and B. Romme. Comment on "Wildland-Urban Interface Maps Vary with Purpose and Context". *Journal of Forestry*, 107(5):232–234, 2009. doi: 10.1016/j.landurbplan.2007.06.002. URL http://dx.doi.org/10.1016/j.landurbplan.2007.06.002.

[281] D.M. Theobald and W.H. Romme. Expansion of the US Wildland-Urban Interface. *Landscape and Urban Planning*, 83(4):340–354, 2007. doi: http://dx.doi.org/10.1016/j.landurbplan.2007.06.002. URL http://www.sciencedirect.com/science/article/pii/S0169204607001491.

[282] W. Tinner, P. Hubschmid, M. Wehrli, B. Ammann, and M. Conedera. Long-Term Forest Fire Ecology and Dynamics in Southern Switzerland. *Journal of Ecology*, 87(2):273–289, 1999. doi: 10.1046/j.1365-2745.1999.00346.x. URL http://dx.doi.org/10.1046/j.1365-2745.1999.00346.x.

[283] W. Tinner, B. Allgöwer, B. Ammann, M. Conedera, E. Gobet, A.F. Lotter, and M. Stähli. Ausmass und Auswirkungen der Waldbrände auf die Vegetation der Schweiz im Laufe der Jahrtausende | Relevance and Effects of Fire Disturbance on Vegetation in Switzerland During the Past Millennia. *Schweizerische Zeitschrift fur Forstwesen*, 156(9):325–330, 2005. doi: 10.3188/szf.2005.0325. URL http://dx.doi.org/10.3188/szf.2005.0325.

[284] W. Tinner, M. Conedera, B. Ammann, and A.F. Lotter. Fire Ecology North and South of the Alps Since the Last Ice Age. *The Holocene*, 15(8):1214–1226, 2005. doi: 10.1191/0959683605hl892rp. URL http://hol.sagepub.com/content/15/8/1214.abstract.

[285] W.R. Tobler. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46(0):234–240, 2010. doi: 10.2307/143141. URL http://www.jstor.org/stable/143141.

[286] M. Tonini, D. Tuia, and F. Ratle. Detection of Clusters Using Space-Time Scan Statistics. *International Journal of Wildland Fires*, 18(7):830–836, 2009. doi: 10.1071/WF07167. URL http://dx.doi.org/10.1071/WF07167.

[287] M. Tonini, A. Pedrazzini, I. Penna, and M. Jaboyedoff. Spatial Pattern of Landslides in Swiss Rhone Valley. *Natural Hazards*, 73(1):97–110, 2014. ISSN 0921-030X. doi: 10.1007/s11069-012-0522-9. URL http://dx.doi.org/10.1007/s11069-012-0522-9.

[288] L. Trabaud and J. Lepart. Changes in the Floristic Composition of a Quercus Coccifera L. Garrigue in Relation to Different Fire Regims. In P. Poissonet, F. Romane, M.A. Austin, E. van der Maarel, and W. Schmidt, editors, *Vegetation Dynamics in Grasslands, Healthlands and Mediterranean Ligneous Formations*, volume 4 of *Advances in Vegetation Science*, pages 105–116. Springer Netherlands, 1981. ISBN 978-94-009-7993-2. doi: 10.1007/978-94-009-7991-8_10. URL http://dx.doi.org/10.1007/978-94-009-7991-8_10.

[289] T.M. Trigo, J.M.C. Pereira, M.G. Pereira, B. Mota, T.J. Calado, C.C. Dacamara, and F.E. Santo. Atmospheric Conditions Associated with the Exceptional Fire Season of 2003 in Portugal. *International Journal of Climatology*, 26(13):1741–1757, 2006. ISSN 1097-0088. doi: 10.1002/joc.1333. URL http://dx.doi.org/10.1002/joc.1333.

[290] D. Tuia and M. Kanevski. Environmental Monitoring Network Characterization and Clustering. In M. Kanevski, editor, *Advanced Mapping of Environmental Data: Geostatistics, Machine Learning and Bayesian Maximum Entropy*, chapter 2, pages 19–46. ISTE/Wiley, London/Hoboken(USA), 2008. ISBN 9780470611463. doi: 10.1002/9780470611463.ch2. URL http://dx.doi.org/10.1002/9780470611463.ch2.

[291] D. Tuia, R. Lasaponara, L. Telesca, and M. Kanevski. Identifying Spatial Clustering Phenomena in Forest-Fire Sequences. *Physica A: Statistical Mechanics and its Applications*, 376:596–600, 2007. doi: 10.1016/j.physa.2006.10.102. URL http://dx.doi.org/10.1016/j.physa.2006.10.102.

[292] D. Tuia, R. Lasaponara, L. Telesca, and M. Kanevski. Emergence of Spatio-Temporal Patterns in Forest-Fire Sequences. *Physica A: Statistical Mechanics and its Applications*, 387(13):3271–3280, 2008. doi: 10.1016/j.physa.2008.01.057. URL http://dx.doi.org/10.1016/j.physa.2008.01.057.

[293] D. Tuia, F. Ratle, R. Lasaponara, L. Telesca, and M. Kanevski. Scan Statistics Analysis of Forest Fire Clusters. *Communications in Nonlinear Science and Numerical Simulation*, 13(8):1689–1694, 2008. doi: 10.1016/j.cnsns.2007.03.004. URL http://dx.doi.org/10.1016/j.cnsns.2007.03.004.

[294] D.L. Turcotte. *Fractals and Chaos in Geology and Geophysics*. Cambridge University Press, Cambridge, United Kingdom, second edition, 1997. ISBN 9781139174695. doi: 10.1017/CBO9781139174695. URL http://dx.doi.org/10.1017/CBO9781139174695.

[295] D.L. Turcotte and B.D. Malamud. Landslides, Forest Fires, and Earthquakes: Examples of Self-Organized Ccritical Behavior. *Physica A: Statistical Mechanics and its Applications*, 340(4):580–589, 2004. doi: 10.1016/j.physa.2004.05.009. URL http://dx.doi.org/10.1016/j.physa.2004.05.009.

[296] K.F. Turkman, M.A. Amaral Turkman, and J.M. Pereira. Asymptotic Models and Inference for Extremes of Spatio-Temporal Data. *Extremes*, 13(4):375–397, 2010. ISSN 1386-1999. doi: 10.1007/s10687-009-0092-8. URL http://dx.doi.org/10.1007/s10687-009-0092-8.

[297] B.W. Turnbull, E.J. Iwano, W.S. Burnett, H.L. Howe, and L.C. Clark. Monitoring for Clusters of Disease: Application to Leukemia Incidence in Upstate New York. *American Journal of Epidemiology*, 132(supp1):136–143, 1990. URL http://aje.oxfordjournals.org/content/132/supp1/136.abstract.

[298] R. Turner. Point Patterns of Forest Fire Locations. *Environmental and Ecological Statistics*, 16(2):197–223, 2009. ISSN 1352-8505. doi: 10.1007/s10651-007-0085-1. URL http://dx.doi.org/10.1007/s10651-007-0085-1.

[299] E. Valese, M. Conedera, A.C. Held, and D. Ascoli. Fire, Humans and Landscape in the European Alpine Region During the Holocene. *Anthropocene*, 6(0):63–74, 2014. ISSN 2213-3054. doi: 10.1016/j.ancene.2014.06.006. URL http://www.sciencedirect.com/science/article/pii/S2213305414000332.

[300] H.J. Vaux. Forestry's Hotseat: The Urban/Forest Interface. *American Forests*, 88 (5):36–46, 1982.

[301] A. Vázquez and J.M. Moreno. Patterns of Lightning-, and People-Caused Fires in Peninsular Spain. *International Journal of Wildland Fire*, 8(2):103–115, 1998. doi: 10.1071/WF9980103. URL http://dx.doi.org/10.1071/WF9980103.

[302] A. Vázquez and J.M. Moreno. Spatial Distribution of Forest Fires in Sierra de Gredos (Central Spain). *Forest Ecology and Management*, 147(1):55–65, 2001. doi: 10.1016/S0378-1127(00)00436-9. URL http://dx.doi.org/10.1016/S0378-1127(00)00436-9.

[303] T.T. Veblen, T. Kitzberger, R. Villalba, and J. Donnegan. Fire History in Northern Patagonia: The Roles of Humans and Climatic Variation. *Ecological Monographs*, 69(1):47–67, 1999. URL http://www.jstor.org/stable/2657194.

[304] C. Vega-García, P.M. Woodard, S.J. Titus, W.L. Adamowicz, and B.S. Lee. A Logit Model for Predicting the Daily Occurrence of Human Caused Forest-Fires. *International Journal of Wildland Fire*, 5:101–111, 1995. doi: 10.1071/WF9950101. URL http://dx.doi.org/10.1071/WF9950101.

[305] C. Vega Orozco, M. Tonini, M. Conedera, and Kanevskin M. Cluster Recognition in Spatial-Temporal Sequences: the Case of Forest Fires. *Geoinformatica*, 16(4): 653–673, 2012. doi: 10.1007/s10707-012-0161-z. URL http://dx.doi.org/10.1007/s10707-012-0161-z.

[306] C.D. Vega Orozco, J. Golay, and M. Kanevski. Multifractal Portrayal of the Swiss Population. *Cybergeo: European Journal of Geography [Online], Section: Systems, Modelling, Geostatistics*, 714, 2015. doi: 10.4000/cybergeo.26829. URL http://cybergeo.revues.org/26829.

[307] T. Vicsek. Mass Multifractals. *Physica A: Statistical Mechanics and its Applications*, 168(1):490–497, 1990. doi: 10.1016/0378-4371(90)90401-D. URL http://www.sciencedirect.com/science/article/pii/037843719090401D.

[308] D.X. Viegas, G. Bovio, A. Ferreira, A. Nosenzo, and B. Sol. Comparative Study of Various Methods of Fire Danger Evaluation in Southern Europe. *International Journal of Wildland Fire*, 9(4):235–246, 1999. doi: 10.1071/WF00015. URL http://www.publish.csiro.au/paper/WF00015.htm.

[309] M Vilà, F. Lloret, E. Ogheri, and J. Terradas. Positive Fire-Grass Feedback in Mediterranean Basin Woodlands. *Forest Ecology and Management*, 147(1):3–14, 2001. doi: 10.1016/S0378-1127(00)00435-7. URL http://dx.doi.org/10.1016/S0378-1127(00)00435-7.

[310] G. Voronoi. Nouvelles Applications des Paramètres Continus à la Théorie des Formes Quadratiques. Premier Mémoire. Sur Quelques Propriétés des Formes Quadratiques Positives Parfaites. *Journal für die Reine und Angewandte Mathematik*, 1908(133):97–102, 1908. doi: 10.1515/crll.1908.133.97. URL http://dx.doi.org/10.1515/crll.1908.133.97.

[311] L.A. Waller and C.A. Gotway. *Applied Spatial Statistics for Public Health Data*. Wiley Series in Probability and Statistics. Wiley, 2004. ISBN 9780471662679. URL http://books.google.ch/books?id=OuQwgShUdGAC.

[312] S. Wang and C.F. Eick. A Polygon-Based Clustering and Analysis Framework for Mining Spatial Datasets. *GeoInformatica*, 18(3):569–594, 2014. ISSN 1384-6175. doi: 10.1007/s10707-013-0190-2. URL http://dx.doi.org/10.1007/s10707-013-0190-2.

[313] P. Weibel. *Modelling and Assessing Fire Regimes in Mountain Forests of Switzerland*. Doctoral and habilitation theses, Swiss Federal Institute of Technology ETH Zürich, Zürich, 2009. URL http://dx.doi.org/10.3929/ethz-a-006088494.

[314] A.L. Westerling, H.G. Hidalgo, D.R. Cayan, and T.W. Swetnam. Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity. *Science*, 313(5789):940–943, 2006. doi: 10.1126/science.1128834. URL http://www.sciencemag.org/content/313/5789/940.abstract.

[315] C. Whitlock, P.E. Higuera, D.B. McWethy, and C.E. Briles. Paleoecological Perspectives on Fire Ecology: Revisiting the Fire-Regime Concept. *The Open Ecology Journal*, 3(0):6–23, 2010. doi: 10.2174/1874213001003020006. URL http://dx.doi.org/10.2174/1874213001003020006.

[316] T. Wiegand and K.A. Moloney. *Handbook of Spatial Point-Pattern Analysis in Ecology*. Chapman & Hall/CRC Press, Boca Raton, 2014. ISBN 978-1-4200-8254-8.

[317] A. Witt, B.D. Malamud, M. Rossi, F. Guzzetti, and S. Peruccacci. Temporal Correlations and Clustering of Landslides. *Earth Surface Processes and Landforms*, 35(10):1138–1156, 2010. ISSN 1096-9837. doi: 10.1002/esp.1998. URL http://dx.doi.org/10.1002/esp.1998.

[318] B.M. Wotton and D.L. Martell. A Lightning Fire Occurrence Model for Ontario. *Canadian Journal of Forest Research*, 35(6):1389–1401, 2005. doi: 10.1139/x05-071. URL http://dx.doi.org/10.1139/x05-071.

[319] J. Yang, H.S. Healy, S.R. Shifley, and E.J. Gustafson. Spatial Patterns of Modern Period Human-Caused Fire Occurrence in the Missouri Ozark Highlands. *Forest Science*, 53(1):1–15, 2007. URL http://www.ingentaconnect.com/content/saf/fs/2007/00000053/00000001/art00001.

[320] S. Yassemi, S. Dragićević, and M. Schmidt. Design and Implementation of an Integrated GIS-Based Cellular Automata Model to Characterize Forest Fire Behaviour. *Ecological Modelling*, 210(1-2):71–84, 2008. doi: 10.1016/j.ecolmodel.2007.07.020. URL http://dx.doi.org/10.1016/j.ecolmodel.2007.07.020.

[321] O. Zammit, X. Descombes, and J. Zerubia. Burnt Area Mapping Using Support Vector Machine. *Forest Ecology and Management*, 234, Supplement(0):S240, 2007. doi: 10.1016/j.foreco.2006.08.269. URL http://dx.doi.org/10.1016/j.foreco.2006.08.269.

[322] T.B. Zeleke and B.C. Si. Scaling Properties of Topographic Indices and Crop Yield. *Agronomy Journal*, 96:1082–1090, 2004. doi: 10.2134/agronj2004.1082. URL http://dx.doi.org/10.2134/agronj2004.1082.

[323] B. Zierl. Simulating the Tmpact of Climate Change on Drought in Swiss Forest Stands. In G. Visconti, M. Beniston, E.D. Iannorelli, and D. Barba, editors, *Global Change and Protected Areas*, volume 9 of *Advances in Global Change Research*, pages 229–243. Springer Netherlands, 2001. ISBN 978-90-481-5686-3. doi: 10.1007/0-306-48051-4_22. URL http://dx.doi.org/10.1007/0-306-48051-4_22.

[324] T. Zumbrunnen, H. Bugmann, M. Conedera, and M. Bürgi. Linking Forest Fire Regimes and Climate - A Historical Analysis in a Dry Inner Alpine Valley. *Ecosystems*, 12(1):73–86, 2009. doi: 10.1007/s10021-008-9207-3. URL http://dx.doi.org/10.1007/s10021-008-9207-3.

[325] T. Zumbrunnen, G.B. Pezzatti, P. Menéndez, H. Bugmann, M. Bürgi, and M. Conedera. Weather and Human Impacts on Forest Fires: 100 Years of Fire History in Two Climatic Regions of Switzerland. *Forest Ecology and Management*, 261(12):2188–2199, 2011. doi: 10.1016/j.foreco.2010.10.009. URL http://dx.doi.org/10.1016/j.foreco.2010.10.009.

[326] T. Zumbrunnen, P. Menéndez, H. Bugmann, M. Conedera, U. Gimmi, and M. Bürgi. Human Impacts on Fire Occurrence: A Case Study of Hundred Years of Forest Fires in a Dry Alpine Valley in Switzerland. *Regional Environmental Change*, 12(4):935–949, 2012. doi: 10.1007/s10113-012-0307-4. URL http://dx.doi.org/10.1007/s10113-012-0307-4.