

# Optimal Regional Partitioning for Wildfire Risk Characterization

Paolo Fiorucci, Francesco Gaetani, Riccardo Minciardi

*CIMA Centro di ricerca Interuniversitario in Monitoraggio Ambientale, Università degli Studi di Genova, via Cadorna, 7-17100 Savona, Italy. Name.Surname@unige.it*

**Abstract:** The efficiency of fire prevention and fire-fighting organization are strictly related with the identification of the elementary Administrative Units, whose physical characteristics and number/quality of resources affect significantly the occurrence and the dynamics of the considered phenomenon. In fact, wildland fire risk is determined by several factors such as vegetation, local climate condition, topography, socio-economical aspects, and, of course, wildfire fighting available resources. A power law distribution can be used to represent the frequency-area relationship empirical wildfire data. Thus, it is possible to think of characterizing a region by its wildland fire regime, which can be defined by the power law parameters estimated on the basis of the historical data relevant to fire occurrences. In this paper, a methodology for optimal regional partitioning, in connection with wildfire risk characterization, is proposed and discussed in detail. In particular, the paper defines a procedure to identify, on the basis of the available data set, over a wide regional area, a number of zones (regions) characterized by different wildfire regimes. These zones are composed by groups of contiguous elementary territorial units, which have the same characteristics, as regards the wildland fires phenomenon. Such a characterization can be used to identify the static risk distribution over the considered territory. Liguria, a small region (5400 km<sup>2</sup>) placed on the northwestern coastline of Italy and frequently affected by severe wildland fires occurrences, is the basis for case study relevant to the implementation of the proposed approach.

**Keywords:** Forest fire; Static risk assessment; Frequency-area statistics; Power-law distribution; Districting problems.

## 1. INTRODUCTION

The protection of the natural environment, and the supervision of the wildland urban interface, require the design and/or the management of an heterogeneous number of resources, both infrastructures (fire-breaks, hydraulic networks, rural roads,...) and mobile (trucks, engines, water-bombers,...).

Generally, the organization of the structure in charge of managing the forest fires emergency is based on a hierarchical scheme in which a central authority (the State or the Region) coordinates the local resources located in some administrative zones (i.e., a set of elementary units). The availability of such resources, and their effectiveness, is influenced by the seasonal nature of forest fires and, overall, by the wildfire risk dynamics in space and time, mostly related with the meteorological conditions. For these reasons,

the management planning has to be based on the sizing and location of a reliable number of mobile re-sources, along with a set of logistic/operational infra-structures needed in the emergency phase. Nevertheless, the high costs related to the management planning require a careful analysis on the trade-off between that costs and the benefits attainable by their use.

A deep knowledge on the actual forest fire regimes of each administrative unit is necessary in such analysis, in order to maximize the (global and local) effective-ness of the resources. In this connection, the present work has to be considered within the framework of the static risk assessment. Usually, in the literature, the evaluation and the distribution of the risk over the territory are carried out on the basis of topographic data and land use information (including vegetation cover), climate, and average fuel conditions (Nunez-Regueira et al., 1998). On the other hand, static risk

assessment can be based on the data series corresponding to historical forest fires occurred in the considered area of study (Forestry Canada Fire Danger Group, 1992).

Actually, static risk cannot be considered only in terms of “frequency of event”; in fact, as it will be discussed in the rest of the paper, forest fires, like other natural events (i.e., landslides, earthquakes, floods), are characterized by a large number of low-intensity events, whose management requires a limited resource effort, and few extreme events, whose occurrence implies the intervention and coordination of a great number of resources. Such a behavior is well represented by the so-called power-law statistical distribution, which allows taking into account explicitly frequency-magnitude relationships. The estimation of power-law distribution requires a large population of data, therefore at small spatial scale (i.e., elementary units) the available frequency of events doesn't allow to define consistent statistical distributions.

The aim of the present paper is to implement a procedure that, on the basis of available dataset of forest fires, is able to aggregate sets of elementary units de-noted by homogeneous values of static risk. A case study relevant to Liguria, a region placed on the northwestern coastline of Italy and frequently affected by severe wildland fires occurrences, is the basis for case study relevant to the implementation of the proposed approach.

## 2. CHARACTERIZING THE WILDFIRES REGIMES USING HISTORICAL DATA

Many works appeared in the last years in the literature (Telesca et al., 2005; Turcotte and Malamud, 2004; Song et al., 2001; Ricotta et al., 1999; Malamud and Turcotte, 1998) propose the so-called *power-law* (scale invariant) statistical distribution in order to characterize an heterogeneous set of natural hazard case studies. Such distribution imposes that the probability that a certain value occurs is proportional to some power of that value.

The definition of the structure of the power-law requires the discretization of the axis relevant to forest fire areas. This discretization must take into account the bi-logarithmic nature of the plots where power-laws are represented. The width of the discretization interval is named “bin” and its value is not constant on the abscissa-axis.

Then, following the problem formalization in Malamud et al. (2005), the so-called power-law state that the probability to have, in a given unit area, and in a given unit time, an expected number

of fires having an extension belonging to the interval  $[A_F - \varepsilon, A_F]$  (where  $A_F$  represent the burnt area and  $\varepsilon$  is the bin width) is given by

$$f(A_F) = \alpha A_F^{-\beta} \quad (1)$$

where  $\alpha$  and  $\beta$  are suitable real constants. In a bi-logarithmic plot, (1) corresponds to a straight line

$$\log[f(A_F)] = -\beta \log[A_F] + \log \alpha \quad (2)$$

Basing on a suitable large (significant) data set, the values of  $f(A_F)$ , for each bin, can be represented on a bi-logarithmic plane and linearly interpolated, using a least squares method, obtaining the values of  $\alpha$  and  $\beta$ , for a given region and a given time interval. Of course, the expected number of events with an area belonging to interval  $[A_F - \varepsilon, A_F]$  occurring in a time interval of length  $\tau$ , and within an area of extension  $A_R$ , is given by  $\tau A_R \alpha A_F^{-\beta}$ , provided that time and space homogeneity are ensured.

Thus, still referring to Malamud et al. (2005), the expected number  $N_F(A_F, A_R, \tau)$  of wildfires with burnt area greater than or equal to a given area  $A_F$  is given by

$$N_F(A_F, A_R, \tau) = \tau A_R \int_{A_F}^{\infty} f(x) dx = \tau A_R \frac{\alpha}{\beta - 1} A_F^{1-\beta} \quad (3)$$

Then, using the well-known Weibull's result, the recurrence interval for wildfires having burnt areas greater than or equal to a given area  $A_F$  may be obtained as

$$T(\geq A_F) = \frac{\tau + 1}{N_F(A_F, A_R, \tau)} = \frac{\tau + 1}{\tau} \frac{\beta - 1}{\alpha} \frac{A_F^{\beta-1}}{A_R} \quad (4)$$

Actually, the knowledge of such recurrence intervals can be considered as a fundamental information for the management planning, only if it is referred to a spatial area whose extension is reasonable for the planning of the resources utilization. In fact, a too large area could poorly discriminate among the different fire regimes, and thus would be of scarce practical utility in management planning. On the other hand, a very limited spatial area would provide information about events with a very low frequency, and thus not very significant for the management planning. Taking into account such exigencies, in the next paragraph a procedure will be introduced, capable to identify a set of zones, within a given region,

each one featured by a characteristic wildfire regime.

### 3. AN ITERATIVE DISTRICTING PROCEDURE FOR WILDFIRE RISK ASSESSMENT

The proposed procedure attempts to aggregate a number  $N$  of areal (elementary) units in  $K \leq N$  zones, through the optimization of a cost function, subject to constraints relevant to the topology of the zones, and the number of actually occurred wildfires in each zone.

The procedure is iterative, as the final partition of the considered set of areal units is achieved by a series of steps, starting from a given initial partition somehow determined.

Let the set of areal units be  $U = \{u_n\}$ ,  $n=1, \dots, N$ . Moreover, let  $Z_i$ ,  $i=1, \dots, K$ , the set of zones (each zone is defined as a subset of  $U$ ) corresponding to a generic partition of the considered region. Of course, the following constraints have to be satisfied

$$Z_i \neq \emptyset \quad i=1, \dots, K \quad (5)$$

$$Z_i \cap Z_s = \emptyset \quad i \neq s, \quad i=1, \dots, K, \quad s=1, \dots, K \quad (6)$$

$$\bigcup_{i=1}^K Z_i = U \quad (7)$$

A further constraint has to be introduced, as regards the zones of a partition, namely each zone has to be connected and may not be constituted by two or more separate geographical components. A  $N \times N$  adjacency matrix  $A$  is a priori known, whose generic element  $a_{ij} = a_{ji}$  is 1 if the areal unit  $u_i$  is directly neighbouring with  $u_j$ , and 0 otherwise. Simple algorithms for graph theory (Deo, 1974) can be used in order to test the connectedness of a given set  $Z_i$  of areal units.

The problem of determining a suitable partition of the considered area can be set within the framework of the well known set-partitioning problem (Garfinkel and Nemhauser, 1972), that is a classical problem in combinatorial optimization. In this framework, *districting* problems (Kalcsics et al., 2005) is one of the classes of problems that have received a major attention in the literature. The statement of a districting problem as a set-partitioning problem is the following

$$\min \underline{C}^T \underline{x} \quad (8)$$

s.t.

$$B\underline{x} = \underline{1} \quad (9)$$

$$x_j \in \{0, 1\} \quad j=1, 2, \dots, N \quad (10)$$

where  $x_j$  is a binary variable associated with a generic zone  $Z_j$ , indicating whether that zone is

selected ( $x_j=1$ ) or not ( $x_j=0$ ) in the solution of the problem,  $j=1, \dots, M$ , being  $M$  a certain number of zones purposely a priori determined among which an optimal partition has to be found. Matrix  $B$  is a  $N \times M$  matrix, whose generic element  $b_{ij}$  is 1 whether the areal unit  $i$  belongs to the  $j$ -th zone, and 0 otherwise.  $\underline{C} = \text{col}[c_1, \dots, c_M]$  is a vector whose generic component  $c_j$  represents the ‘‘cost’’ of the  $j$ -th zone.

Then, in order to state the considered problem, regarding the partition of the region in relation to wildfire risk characterization, it is necessary: a) to determine a suitable large number of admissible zones (satisfying also the connectedness constraint), b) to define a quality measure for a certain zone, in order to assess the value of coefficient  $c_j$  for any of the determined areas.

Considering the second of the above issues, it is worth recalling that, for a given zone  $Z_j$ , the optimal estimates of coefficient  $\alpha_j$  and  $\beta_j$  characterizing the power law for that zone can be determined via the ordinary Least Squares method. Namely, let  $A_{F,h}$ ,  $h=1, \dots, H$  the right extreme of the discretization intervals (bins) a-priori defined for the considered power laws representation. Let  $n_{F,j,h}$  the number of wildfires occurred in the set of areal units belonging to zone  $Z_j$  (normalized with respect to the area of the zone and the length of observation time), which falls in bin  $[A_{F,h-\varepsilon}, A_{F,h}]$ . Then, the optimal estimates  $\alpha_j$  and  $\beta_j$  are determined by solving the following unconstrained minimization problem

$$\min_{\alpha_j, \beta_j} \frac{1}{H} \sum_{h=1}^H \left[ n_{F,j,h} + \beta_j \log[A_{F,h}] - \log \alpha_j \right]^2 \quad (11)$$

Let  $\sigma_j^2$  the optimal value of the cost function obtained via the minimization in (11). Then, the

‘‘cost’’  $c_j$  can be set equal to  $\sigma_j^2 \sum_{h=1}^H n_{F,j,h}$ .

Instead, as regards the first of the above mentioned issues, i.e., the a-priori generation of a suitably large set of admissible zones, things are much more difficult. In fact, the generation of a quite limited set of zones would heavily affect the further application of an optimal decision approach, like that consisting in the solution of a set-partitioning problem (through well-established algorithms developed for such a problem (Garfinkel and Nemhauser, 1972)). On the other hand, the generation of an extensive set of admissible zones would imply the use of time-consuming combinatorial generation procedure, made heavier by the necessity of checking, for each zone, whether the admissibility constraints, and, in particular, that concerning the

connectedness of the zone are fulfilled. For this reason, an heuristic procedure has been developed and applied, whose aim is not that of finding a set of zones for the statement of the set-partitioning problem, but directly that of determining a solution (i.e., a partition of the considered region) whose quality can be considered as “reasonably good”.

To this end, the proposed iterative procedure requires the initial generation of a partition of the region into a set of admissible zones (each zone must have an overall number of occurred fires not lower than a certain number  $Q$ , unless the statistical significance of the power law for that region is questionable). Besides, each of such zones is a connected set of areal units. Then, a set of *unary* perturbations are generated to the considered solutions, in which only an areal unit is involved; each perturbation consists in the generation of a new partition in which, with respect to the original partition, a single areal unit has been moved from one zone to another (of course, having care of fulfilling the connectedness constraints) to each of those  $P$  partitions, generating an additional set of  $P$  (admissible) partitions. Then, the whole set of  $2P$  partitions is investigated, as regards the quality of each partition, that is evaluated through a cost function like that in (8), where each coefficient is computed

as  $c_j = \sigma_j^2 \sum_{h=1}^H n_{F,j,h}$ . The “best”  $P$  partitions are then

selected to “survive”, and the remaining  $P$  are discarded. To each of the survived partitions, a many (admissible) perturbation is then applied, and so on. The procedure ends whenever the difference between the quality of the best partition of a certain “generation” (the procedure has close similarities with genetic optimization) and the quality of the best partition of the previous generation, is below a certain pre-specified threshold.

For the sake of brevity, the details of the proposed procedure are not reported here. Instead, in the next section, the results of its application to a real case study will be provided and discussed.

#### 4. CASE STUDY

Liguria (Italy) is a small region placed on the northwestern coastline of Tyrrhenian Sea. Although its relative northern latitude, in this Mediterranean region the wildfires are quite common both in winter and in summer seasons and, on average, 500 events that burn 50 km<sup>2</sup> are reported every year.

These high values are due to the extensive vegetated area (65% of the administrative area), the climate, and a high population density, whose presence determines an extended and spread wildland urban interface.

The available data (CFS, 1987-2004) refer to 14730 wildfires that burnt 1077 km<sup>2</sup> of forests and shrublands in Liguria between 1987 and 2004. The information considered by the procedure is the date, the location, and the burnt area of each record; furthermore, a deeper analysis of the data, suggested to consider only the records denoted by burnt area  $A_F \geq 0.01$  km<sup>2</sup> (1 hectare). In fact, it is assumed that under this threshold, the data are not enough representative of the phenomenon, because of the lack or incomplete information on small or microscopic fires.

The original inventory was partitioned into two sets, one relevant to winter fires, from November to April, and one relevant to the summer fires, from May to October. In this way, it is possible to distinguish the seasonal behavior of each zone and, therefore, to define different seasonal management planning. The resulting dataset is composed by  $N_{FT} = 4391$  winter wildfires and  $N_{FT} = 3996$  summer wildfires.

The districting procedure introduced in Section 3 has been implemented in order to define a finite number of zones denoted by homogeneous values of  $\alpha$  and  $\beta$ , starting from a set of  $N=235$  elementary areal units.

In Figure 1, a representation of Liguria region subdivided in zones is given. In this case, the application of the procedure in summer season allows obtaining 13 different zones, which show peculiar and well-differentiated wildfire regimes.

In Table 1, the values of parameters  $\log \alpha$ , and  $\beta$  along with the coefficient of determination  $R^2$ , for each zone are reported for the summer season case study. The last column of Table 1 report the recurrence intervals  $T(\geq A_F)$  representing the average time between wildfires  $\geq 0.01$  km<sup>2</sup>.

#### 5. CONCLUSIONS

The application of the proposed procedure allows assessing wildfire static risk through the identification of specific regimes relevant to a finite set of zones.

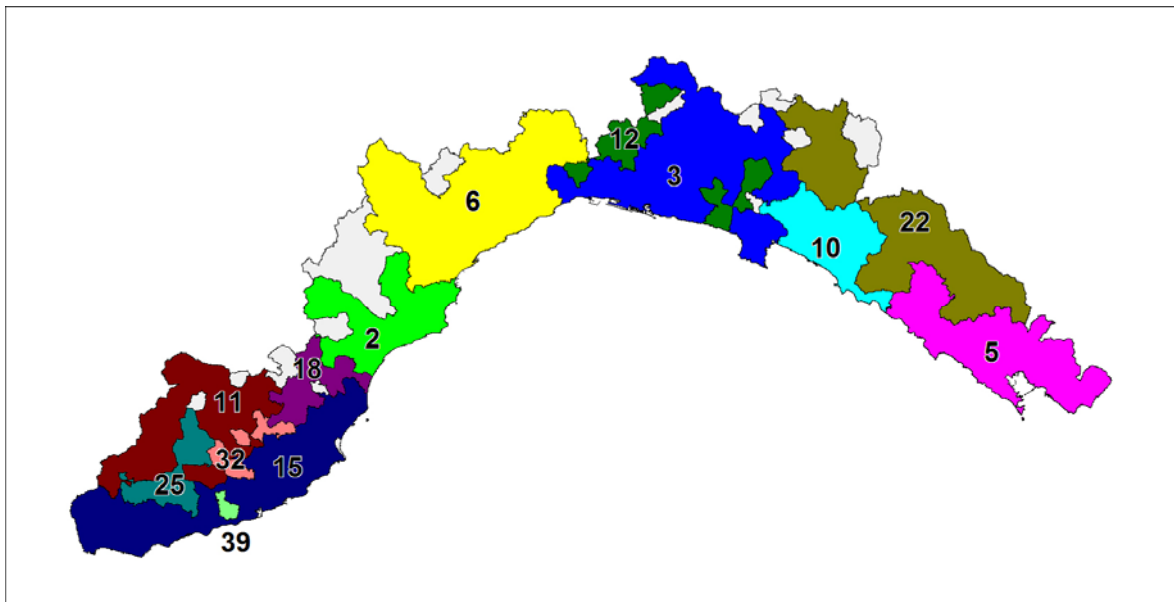
Such zones represent the optimal territorial units that can be used in order to locate new resources and infrastructures for the prevention and suppression of wildfires. In fact, since the wildfire regime of a given zone is influenced by the availability and the typology of the resources usually operational in that zone, it seems not sensible to relocate existing resources from a zone to another. Actually, it is very difficult to identify the effects that already available resources have on fire regimes, which could be obtained only by means of a very large database including detailed information about the resources intervened on the fire front. Nevertheless, the wildfire regime can be considered as the result of the (independent)

effects of several and heterogeneous territorial features. In this connection, a correlation analysis among the parameters  $\alpha$  and  $\beta$  of each zone and the set of features that can be considered as significant for wildfire phenomena (topography,

vegetation, population, ect.) can be carried out in order to isolate the contribution of each single feature to the regime and, therefore, the influence that firefighting resources have on fire behavior.

**Table 1.** Results by the obtained zones in Liguria region (Italy) for frequency area and recurrence interval analyses of wildfires.

<i>Zone</i>	<i>A<sub>R</sub> km<sup>2</sup></i>	<i>N<sub>FT</sub></i>	<i>β</i>	<i>log(α)</i>	<i>R<sup>2</sup></i>	<i>T(A<sub>F</sub>≥0.01 km<sup>2</sup>) days</i>
2	343	81	-1.461	-3.075	0.963	73.8
3	766.82	349	-1.741	-3.097	0.978	15.36
5	494.63	227	-1.532	-2.876	0.967	26.9
6	806.15	189	-1.494	-3.193	0.964	37.92
10	279.44	95	-1.471	-2.917	0.969	61.5
11	431.22	63	-1.384	-3.146	0.969	82.14
12	200.33	94	-1.594	-2.755	0.994	42.22
15	595.18	741	-1.703	-2.639	0.964	7.78
18	142.02	51	-1.403	-2.740	0.979	94.28
22	669.63	86	-1.901	-3.785	0.956	50
25	142.02	43	-1.356	-2.625	0.993	79.33
32	65.47	51	-1.389	-2.377	0.969	91.23
39	15.51	24	-1.209	-1.822	0.967	132.22



**Figure1** The results obtained after the application of the procedure in case of summer fire season.

Turcotte, D.L., Malamud, B.D., 2004. Landslides, forest fires, and earthquakes: examples of self-organized critical behavior. *Physica A* 340 (4), 580-589.

## 6. REFERENCES

- CFS, Corpo Forestale dello Stato, Ministero per le Politiche Agricole e Forestali. *Foglio Notizie Antincendio Boschivo*, years 1987-2004, Rome (Italy).
- Deo, N., 1974. Graph theory with applications to engineering and computer science. Prentice-Hall Series in Automatic Computation. Prentice-Hall, Inc., Englewood Cliffs, N.J. 0 , pp. 1-478
- Forestry Canada Fire Danger Group 1992 Development and structure of the Canadian Forest Fire Behavior Prediction System. Information Report ST-X-3 Forestry CA, Ottawa, ON.
- Fons, W.T. 1946. Analysis of fire spread in light forest fuels. *J. Agric. Res.* 72(3):93-121.
- Garfinkel, R.S., Nemhauser, G.L., 1972. Integer Programming. John Wiley & Sons, New York.
- Kalcsics, J., Nickel, S., Schroder, M., 2005. Towards a Unified Territorial Design Approach. Applications, Algorithms and GIS Integration. *Top* (2005) Vol. 13, No. 1, pp. 1-74.
- MAF/ISAF, 1988. Inventario Forestale Nazionale. Sintesi metodologica e risultati. Ministero dell'Agricoltura e delle Foreste. Istituto Sperimentale per l'Assestamento forestale e per l'Alpicoltura, Trento (Italy).
- Malamud B.D., Turcotte D., 1998. Self organized criticality applied to natural hazards. *Natural Hazard v. 20*, pp. 93-116.
- Malamud B.D., Millington J.D.A., Perry G.L.W., 2005, Characterizing wildfire regimes in the United States. *PNAS*, 102, n. 13, pp. 4694-4699.
- Nunez-Regueira L, Rodriguez J, Proupin J, Mourino B. 1999. Design of forest biomass energetic maps as a tool to fight forest wildfires. *Termochimica Acta*, 328, pp. 111-120.
- Ricotta, C., Avena, G., Marchetti, M., 1999. The flaming sandpile: self-organized criticality and wildfires. *Ecological Modelling*. 119, 73-77.
- Rothermel, R.C. 1972. A mathematical model for predicting fire spread in wildland fuels. *USDA For. Serv. Res. Pap. INT-115*.
- Song, W.G., Fan, W.C., Wang, B.H., Zhou, J.J., 2001. Self-organized criticality of forest fire in China. *Ecological Modelling*. 145(1), 61-68.
- Telesca, L., Amatulli, G., Lasaponara, R., Lo vallo, M., Santulli, A., 2005. Time-scaling properties in forest-fire sequences observed in Gargano area (southern Italy). *Ecological Modelling*. 185, 531-544.