INVESTIGATION OF WILDFIRES AT FORESTED LANDSCAPES: A NOVEL CONTRIBUTION TO NONPARAMETRIC DENSITY MAPPING AT REGIONAL SCALE

KUTER, N.[1] – KUTER, S.[2*]

[1]*Landscape Architecture Dept., Cankiri Karatekin University, Cankiri, Turkey*
(phone: +90-376-212-2757; fax: +90-376-213-6983)

[2]*Forest Engineering Dept., Cankiri Karatekin University, Cankiri, Turkey*
(phone: +90-376-212-2757; fax: +90-376-213-6983)

*Corresponding author
e-mail: semihkuter@karatekin.edu.tr

Abstract. Forest fires have been regarded as a common phenomenon that has endangered forests throughout the history of the world. It changes the vegetation and forest floor suddenly and drastically. Kernel density estimation (KDE) is a frequently employed approach in order to transform historical wildfire data into a smooth and continuous 2-D surface during the spatiotemporal analysis of forest fires. The most crucial step in KDE is to choose the appropriate bandwidth parameter. In this study, a novel approach to obtain an appropriate smoothing parameter is introduced by adapting a powerful nonparametric spline fitting methodology, namely, multivariate adaptive regression splines (MARS) into the KDE analysis of wildfires for the first time. Spatial and temporal analysis of wildfires between 2000 and 2017 in Mumcular Forest Subdistrict in Turkey is investigated by using the MARS-based KDE analysis. The proposed methodology in this paper produces a single smoothing parameter, instead of a range of values laying within a certain fixed interval, and it appears more robust to the positional uncertainties in the historical fire data, and more reliable than nonobjective visual analysis of maps generated by using alternating smoothing parameter values.
Keywords: *forest fires, kernel density, kernel bandwidth, MARS, Mumcular*

Introduction

Wildfires may often lead to serious damage to the ecosystems of forested landscapes, as well as to the diversity of flora and fauna. Preparing wildfire risk maps by using temporal fire occurrence data is very important for reducing these negative effects of forest fires. However, a wildfire event is generally recorded with its *x* and *y* coordinates by assuming each fire event as a point process, which negatively alters its 2-D exterior characteristics and propagation style. Additionally, temporal wildfire data mostly suffer from inappropriate or inadequate information and precision due to the fact that it is operationally and logistically hard to obtain the real coordinates of locations where a wildfire event starts.

This problematic issue significantly degrades the geolocational accuracy of wildfire data. These errors can eventually propagate wildfire liability estimates and result in unforeseeable negative impacts on the related risk maps. In order to tackle with this problem, two main approaches exist: *i*) to overlay a 10×10 km Universal Transverse Mercator (UTM) grid on administrative borderlines and to force the wildfire starting locations (i.e., points) to be confined in a polygonal area demarcated by the 10×10 km net, administrative borders and prescribed wild-land sites, or *ii*) to increase the resolution of lattice superimposed over the recorded wildfire starting locations in order

to decrease geolocational inaccuracies, which, in turn, may often cause complications regarding abstraction and reduction in geospatial variability (Koutsias et al., 2004).

Derivation of a density function by using the locations of recorded observations is often considered as an effective approach to create a continuous raster surface from point data. Kernel density estimation (KDE) is a commonly preferred approach to obtain such a 2-D raster surface. KDE is a nonparametric method and its output is a surface that models the hidden density function (Anderson, 2009). KDE is a widely-recognized methodology and used in wildlife ecology (Horne et al., 2007; Katajisto and Moilanen, 2006; Kie et al., 2010), in geology and geophysics (Kagan and Jackson, 2000; Lasocki and Orlecka-Sikora, 2008), in spatial epidemiology (Anderson and Titterington, 1997; Robertson et al., 2010), in transportation planning and management (Okabe et al., 2009; Xie and Yan, 2008), in crime analysis (Chainey et al., 2008; Nakaya and Yano, 2010), in civil and mechanical engineering (Chen et al., 2000; Worden et al., 2003; Yu and Su, 2012), and in medicine (Rossiter, 1991; Zou et al., 1997).

Preceding studies

Koutsias et al. (2014) conducted a study in order to obtain a map that reveals the fire occurrence zones of Greece at a national scale by using KDE. In the study, they employed 4 different kernel widths as 1000, 2000, 3000 and 4000 m, and statistically compared them by Monte Carlo randomization test. The results indicated that increasing the width of kernels above 1000 m did not have any influence on the density surfaces.

To investigate the relationship between the fuel phenology and the spatiotemporal patterns of wildfire events in Sardinia, Italy, Bajocco et al. (2017) used KDE to convert wildfire ignition points to a 2-D raster density surface. According to the results of the study, wildfire locations were significantly associated to both anthropogenic pressure and to the spatiotemporal variation of fuel conditions in the study area.

In the study of Camarero et al. (2018), the impact of anthropogenic factors and climate systems on historical fire incidence in Mediterranean black pine forests were investigated within a long time interval (i.e., 1800-2000). Three input parameters, i.e., charcoal accumulation rates, historical records of wild fire events and tree-ring data, were collected in Sierra de Gredos (central Spain), and employed in the analysis. The historic fire data in point format were transformed into a continuous 2-D density surface by using KDE approach. The results of the study pointed that extensive grazing and uncontrolled use of forests and grasslands, together with enormous warm spring temperatures could lead to increased wildfire occurrence in Mediterranean pine forests.

The studies of Koutsias et al. (2014), Bajocco et al. (2017) and Camarero et al. (2018) have revealed that KDE is an effective way to convert the temporal wildfire data with inherent point characteristic and a certain degree of uncertainty into a continuous raster surface for defining spatial arrangement of wildfire occurrences at regional dimensions and analyzing their spatial scope, and it has also been applied by many other researchers to investigate the spatiotemporal patterns of such events (Del Hoyo et al., 2011; Gonzalez-Olabarria et al., 2012; Gralewicz et al., 2012; Hély et al., 2010).

Basics of kernel density estimation

KDE is a fundamental data smoothing problem and a nonparametric statistical technique to estimate a real valued function. In order to calculate KDE for a finite data

sample, *i*) a grid with predefined spatial resolution is overlaid on the study area, *ii*) and the intensity at each crossing of the grid is predicted by putting a symmetrical surface on each of the point event locations (cf. *Fig. 1*). KDE in bivariate form can be defined as (*Eq. 1;* Hastie et al., 2009; Worden et al., 2003):

$$\hat{f}(\mathbf{z}) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{\mathbf{z} - \mathbf{Z}_i}{h}\right),$$ (Eq.1)

where $\hat{f}(\mathbf{z})$ is the prediction for the real intensity function, *n* denotes the number of total event records which are in point data format, *h* is the width of the kernel window (also referred as Parzen window (Hastie et al., 2009)), which is commonly known as a smoothing parameter, **z** gives the vector of coordinates at which the density is calculated (i.e., intersection of overlaid grid), $\mathbf{Z}_i$ represents the geolocation of each event point (i.e., *x* and *y* coordinates of wildfire events), and *K* indicates the kernel function that meets the following requirement (*Eq. 2*):

$$\int K(z)\,dz = 1.$$ (Eq.2)
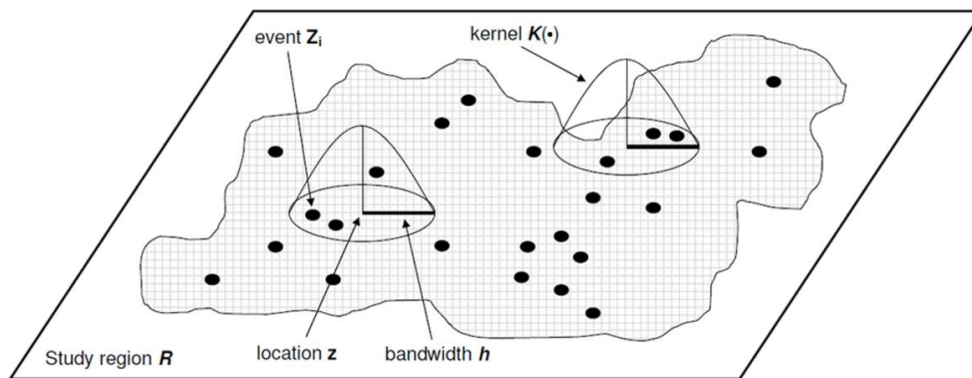
Figure 1. *KDE of a set of point observations in 2-D. (Adapted from Bailey and Gatrell, 1995)*

The choice for *K* can be made from a range of options such as triweight, uniform and normal. In this study, normal kernel (i.e., Gaussian), the widely recognized and employed kernel function, is preferred and its bivariate form is given as (*Eq. 3;* Wand and Jones, 1993)

$$K(\mathbf{x}) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{x}\right).$$ (Eq.3)

At this point, it is of value to mention that the type of the kernel function does not have a significant impact; however, the selection of smoothing parameter directly and strongly affects the estimated density values. While small choice of *h* causes insufficient smoothing and gives spikier density values rather than the underlying density function, large values of *h* result in over-smoothed estimates that obscure important features of the underlying structure (Bowman and Azzalini, 1997; Wand and Jones, 1993; Worden et al., 2003).

For KDE analysis, there are several data-driven methods introduced by researchers in order to calculate the bandwidth parameter (Bailey and Gatrell, 1995; Williamson et al., 2008; Worton, 1989), yet enumerating them all with details is out of our scope. Selecting the bandwidth value by visual inspection is another frequently used method in such analysis and this approach can be found satisfactory for many applications (Kie et al., 2010; Wand and Jones, 1995). However, a profound and extensive information and a high level of expertise regarding the area of study are required in this approach, which may not always be satisfied by the analyst.

Main motivation

In our study, as a novel contribution to KDE analysis of wildfire events, a state-of-the-art multivariate adaptive regression splines (MARS) (Friedman, 1991) algorithm is applied to predict a suitable value for smoothing parameter. For this purpose, Mumcular Forest Subdistrict located in City of Muğla, Turkey is selected as study area, and wildfire data of the region between years 2000 and 2017 are used. KDE maps with different bandwidths are produced, and then compared by correlation analysis. Correlation values are plotted against bandwidths for building the MARS model function. By using the generated MARS model function, the appropriate value of smoothing parameter is decided. Then, KDE maps of wildfire events in the study area are produced to expose the behavior of wildfires in both spatial and temporal domains in the region. Then, final density map is obtained and converted into a mean density map representing the average intensity values for each of the administrative polygons in the entire region.

The paper is organized as follows. The next section represents the materials and methods including a brief mathematical background of the MARS approach, study area, the wildfire data and the estimation of suitable smoothing parameter. The results then are discussed in the third section. And finally, the last section concludes our study.

Materials and methods

Multivariate adaptive regression splines (MARS)

In this subsection, a brief introduction to MARS algorithm is given based on Friedman (1991), Hastie et al. (2009), Nalcaci et al. (2018) and Özmen et al. (2018).

MARS is a nonparametric spline regression method and it uses one-dimensional piecewise linear basis functions (BFs) in order to define a relationship between a response variable and its predictors. MARS has many successful implementations in almost every branch of science and engineering (Alp et al., 2011; Çevik et al., 2017; Durmaz et al., 2010; Özmen et al., 2014; Quirós et al., 2009).

The BFs of MARS are in the following form (*Eq. 4*):

$$[x-t]_+ = \begin{cases} x-t, & \text{if } x > t, \\ 0, & \text{otherwise,} \end{cases}$$
$$[t-x]_+ = \begin{cases} t-x, & \text{if } x < t, \\ 0, & \text{otherwise,} \end{cases} \tag{Eq.4}$$

where t is a univariate knot determined using the dataset. The range of each predictor variable is cut into subsets of the full range by using knots which defines an inflection point along the range of a predictor. The MARS model in its general form can be expressed as (*Eq. 5*):

$$Y = \beta_0 + \sum_{m=1}^{M} \beta_m B_m \left(\boldsymbol{X}^m \right) + \varepsilon.$$ (Eq.5)

In *Equation 5, Y* is the response, β_0 is the intercept, β_m denotes an unknown coefficient of the mth BF, or the constant 1 ($m = 0$), M is given as the number of BFs in the present model, B_m is a BF or product of two or more BFs, $\boldsymbol{X}^m$ is the vector of predictor variables contributing to the function B_m, and finally, ε is an additive stochastic component with zero mean and finite variance.

The basic form of the mth BF is given as (*Eq. 6*):

$$B_m \left(\boldsymbol{X}^m \right) := \prod_{j=1}^{K_m} \left[s_{\kappa_j^m} \cdot \left(x_{\kappa_j^m} - \tau_{\kappa_j^m} \right) \right]_+ .$$ (Eq.6)

In *Equation 6, K_m* denotes the total number of truncated linear functions multiplied in the mth BF, $x_{\kappa_j^m}$ indicates the input variable corresponding to the kth truncated linear function in the mth BF, $\tau_{\kappa_j^m}$ is the knot location for $x_{\kappa_j^m}$ and, finally, $s_{\kappa_j^m} \in \{ \pm 1 \}$.

The final MARS model is determined with a two-step process including *forward pass* and *backward pass* stages. Forward pass generally creates an over-fit model. At the backward stage, the over-fitted model is simplified in terms of its complexity without degrading the overall fit to the data by imposing a *lack-of-fit* criterion defined by the following *generalized cross-validation* (GCV) formula (*Eq. 7*):

$$\text{LOF}(\hat{f}_\alpha) = \text{GCV}(\alpha) = \frac{\sum_{i=1}^{N} (y_i - \hat{f}_\alpha(\boldsymbol{X}_i))^2}{(1 - Q(\alpha)/N)^2},$$ (Eq.7)

where $Q(\alpha)$ denotes the effective number of parameters in the model, and N is the number of observations.
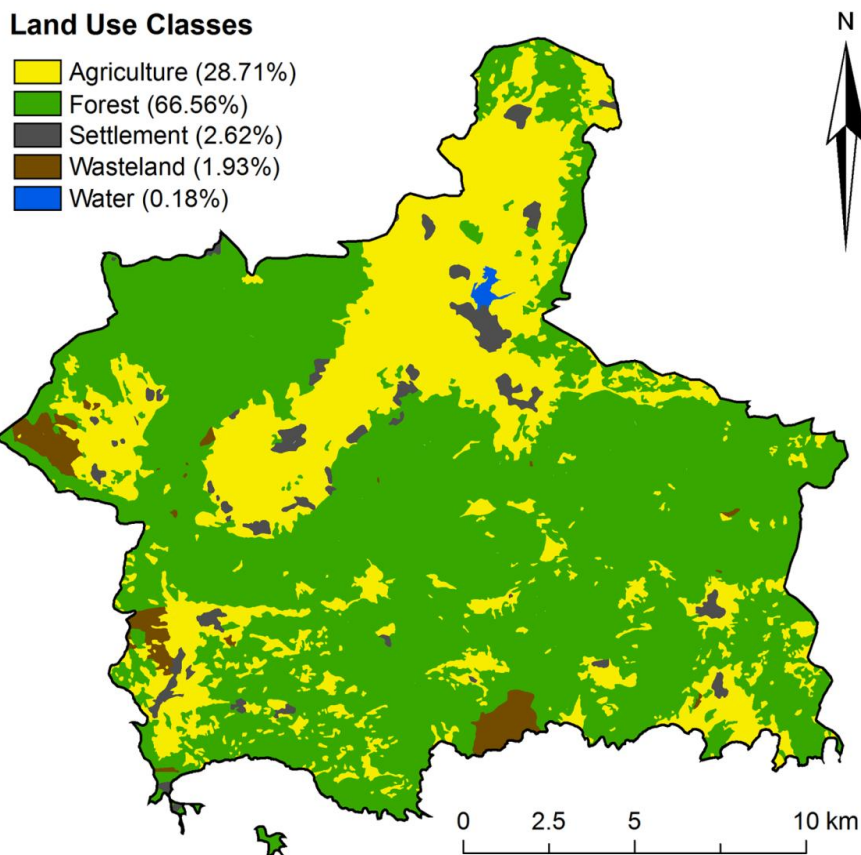
Study area

In the Mediterranean and Aegean regions of Turkey, forest fire is still one of the greatest natural hazard problems. Forest Subdistrict of Mumcular in Muğla province (cf. *Fig. 2*) located in the south-west Aegean region of Turkey is selected, since this region has experienced large number of wildfire events for the last two decades. The study site covers an area of nearly 322 km^2, and its altitude varies between 0 to 870 m above mean sea level. Forested lands constitute nearly 67% (214 km^2) of the study area, the canopy density of 43% of the forested lands ranges from 41 to 100%. Cultivated lands cover 29% of the area (i.e., 93 km^2). The land classification is derived by using the latest digital stand map of the subdistrict in ArcMap shape file format and can be seen in *Figure 3*. "*Pinus brutia*" is the prevailing tree species in the subdistrict region and it is

extremely flammable and fire-prone due to its resinous structure. The detailed info on study area is given in *Table 1*.
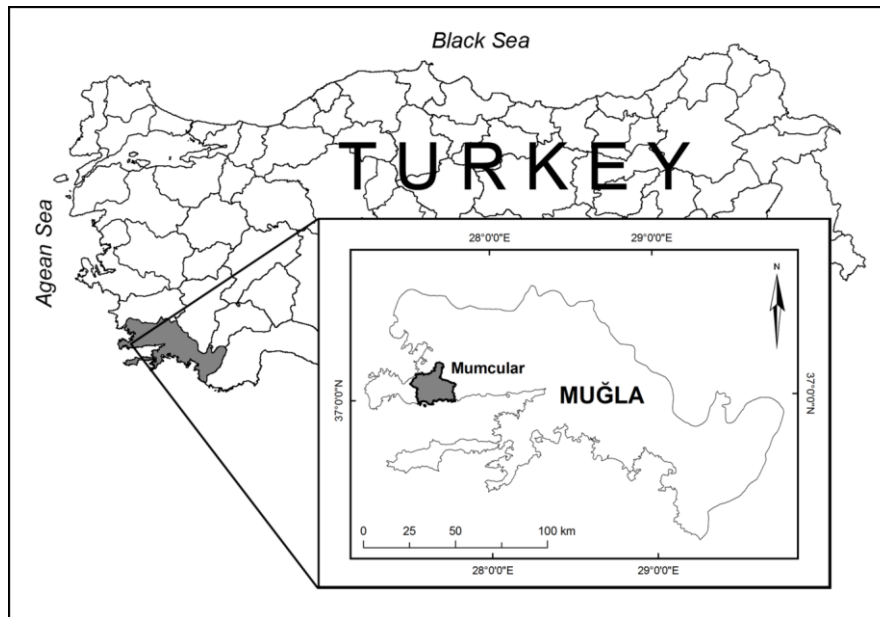

Figure 2. *Study area*

Figure 3. *Land-use map of the study area (Produced from the latest version of the digital stand map of the area supplied by Mumcular Forest Subdistrict)*

Table 1. *Basic information on the forest subdistrict area*

Surface area	321.7 km^2
Number of compartments (polygons)	373
Average value of wildfire occurrence per compartment	0.582
Number of wildfire events (years 2000-2017)	217
Min. compartment area	0.0029 km^2
Max. compartment area	9.43 km^2
Average compartment area	0.83 km^2

Dataset

The temporal wildfire events data supplied by Mumcular Forest Subdistrict are in spread sheet format and belong to years between 2000 and 2017. There are 217 fire ignition locations in total. The coordinate system of the digital stand map of the subdistrict conforms with WGS84/UTM projection. The map comprises 373 compartments, which can be defined as administrative polygons (cf. *Fig. 4*). As a preliminary preparation prior to the analysis, the complete wildfire data are merged with the digital stand map in *ArcMap 10.3*.
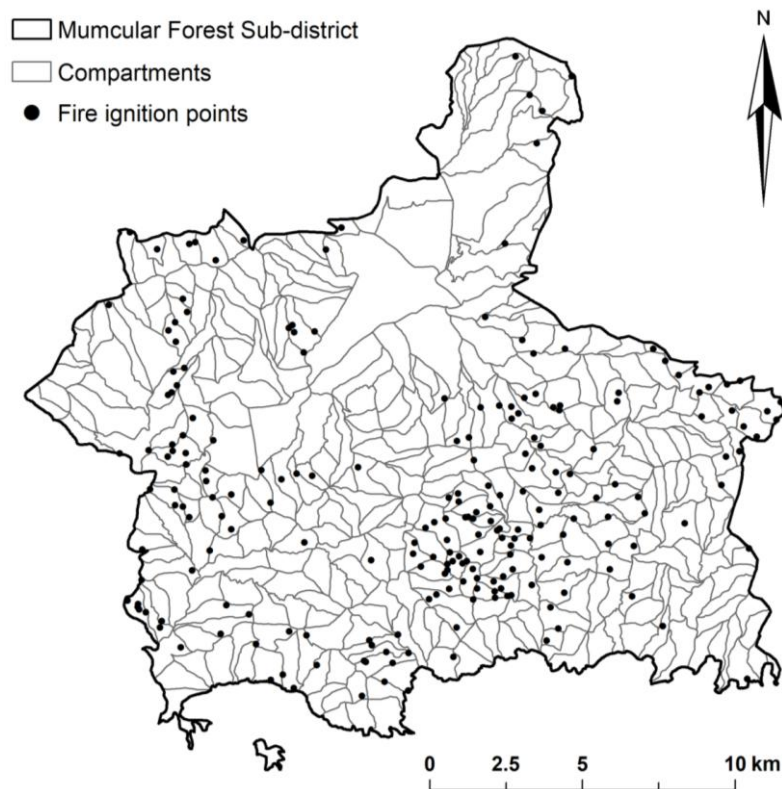
Figure 4. *Compartments of the digital stand map and ignition points*

Choice of KDE smoothing parameter

Although true geolocations of wildfires are not known, the true number of wildfire events is available for each compartment. Therefore, instead of utilizing above

mentioned traditional methods, three completely unique and distinct random point distributions are created (i.e., R1, R2 and R3), and by this way, exact number of wildfire occurrences in each of the administrative polygons (i.e., compartments) is realized.

This approach lets us not only provide the ability of introducing randomness into the analysis, but also confine the wildfire starting points in relatively smaller administrative areas (i.e., compartments in the stand map) as compared to the traditional method (i.e., superimposing 10×10 km net over the area). As a result of this process, an improvement over positional inaccuracies is achieved.

To select an appropriate smoothing parameter (i.e., h), a different and a novel strategy is followed. First, KDE maps for various h taking values within the set {1,2,…5000} are obtained for each random point distribution, i.e., R1, R2 and R3, which means that 15000 KDE maps are generated on aggregate. Then, mutual comparison between each pair of KDE maps for each value of h is carried out via a simple analysis of correlation. Correlation values for each random point distribution for several h values (i.e., 100, 500, 2500 and 5000 m) are given in *Table 2*.

Table 2. Correlation coefficients vs. smoothing parameter h

| Random distributions→ $h↓$	R1-R2	R1-R3	R2-R3
100 m	0.4015275	0.4427352	0.4259254
500 m	0.9225721	0.9245119	0.9205818
2500 m	0.9993016	0.9990487	0.9994646
5000 m	0.9990197	0.9990199	0.9990199

Next, in order to include more stochasticity in the process, random noise is added to the correlation values of each pair of random point distributions. This noise mimics the uncertainties inherited in the coordinates of the wildfire locations which may eventually propagate to the correlation analysis between each pair of point distributions. By this way, extra randomness in the positional inaccuracies of wildfire events is introduced. The standard deviation of the noise added is equal to the standard deviation of the correlation values of that specific pair. The plots of the correlation values with noise with respect to h for three random point distributions are depicted in *Figure 5*.
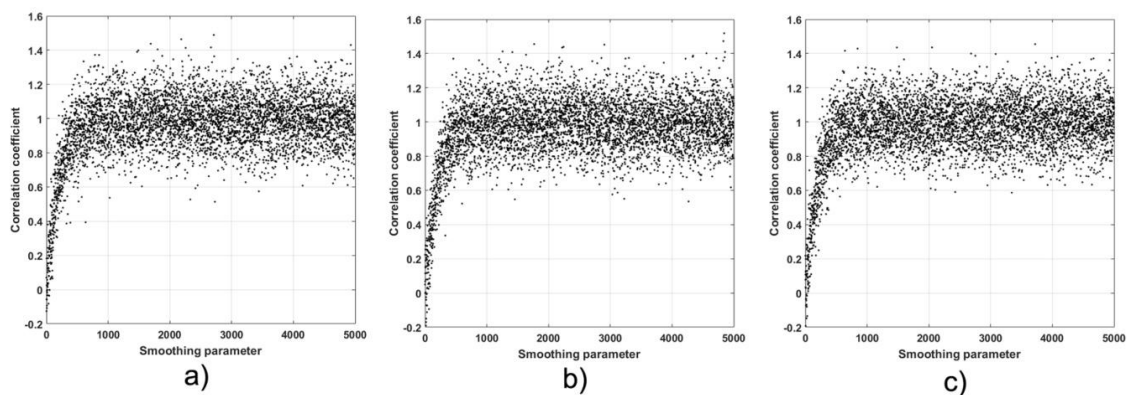
Figure 5. Correlation coefficients with stochastic noise vs. smoothing parameter for a) R1-R2, b) R1-R3, and c) R2-R3. (Please note that the range of correlation values exceeds the usual range of [0, 1] due to the added stochastic noise)

Then, the corresponding MARS model that gives the best fit to the correlation coefficient data with noise for each pair of random point distributions according to GCV criterion given in *Equation 7* is obtained. The graphs of individual MARS models are available in *Figure 6*.
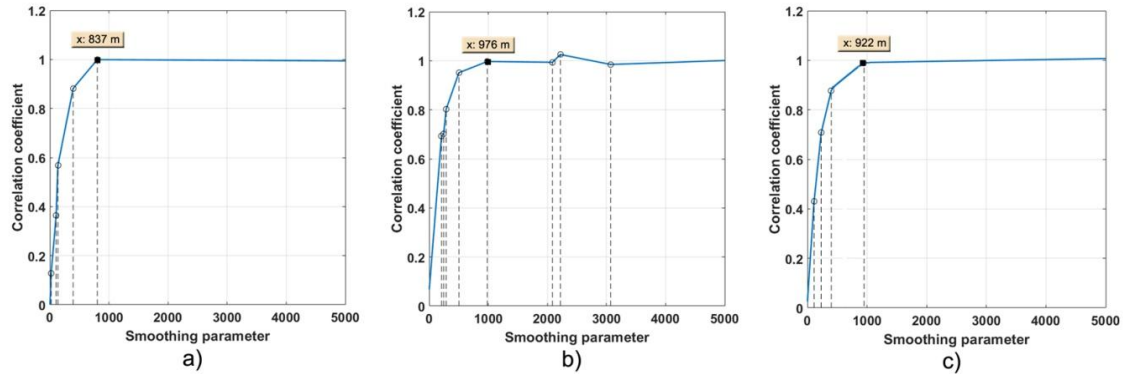
Figure 6. *Plots of MARS model functions for a) R1-R2, b) R1-R3, and c) R2-R3. In the figure, o indicates the location of knot points (i.e., t) between two consecutive BFs, and ■ denotes the position of the knot location where further increase in h results in over-smoothing in the corresponding KDEs*

The obtained MARS models for R1-R2, R1-R3 and R2-R3 are given as:

for R1-R2:
$BF_1 = \max[0, 395 - x_1]$,
$BF_2 = \max[0, 137 - x_1]$,
$BF_3 = \max[0, x_1 - 837]$,
$BF_4 = \max[0, x_1 - 23]$,
$BF_5 = \max[0, x_1 - 107]$ and,
$Y = 1.1146 - 0.00092793 \cdot BF_1 - 0.0056183 \cdot BF_2 - 0.00028838 \cdot BF_3 - 0.0037535 \cdot BF_4 + 0.0040407 \cdot BF_5$,

for R1-R3:
$BF_1 = \max[0, x_2 - 976]$,
$BF_2 = \max[0, 209 - x_2]$,
$BF_3 = \max[0, 503 - x_2]$,
$BF_4 = \max[0, 287 - x_2]$,
$BF_5 = \max[0, 2219 - x_2]$,
$BF_6 = \max[0, x_2 - 2087]$,
$BF_7 = \max[0, x_2 - 2531]$,
$BF_8 = \max[0, 2531 - x_2]$ and,
$Y = 0.97569 - 0.00009659 \cdot BF_1 - 0.00148852 \cdot BF_2 - 0.000631463 \cdot BF_3 + 0.00072293 \cdot BF_4 - 0.00052987 \cdot BF_5 + 0.0003719 \cdot BF6 - 0.00027321 \cdot BF7 + 0.00043687 \cdot BF8$,

and finally for R2-R3:
$BF_1 = \max[0, x_3 - 401]$,
$BF_2 = \max[0, x_3 - 113]$,

$BF_3 = \max[0, 113 - x_3]$,
$BF_4 = \max[0, x_3 - 233]$,
$BF_5 = \max[0, 922 - x_3]$ and,
$Y = 0.68602 - 0.00050397 \cdot BF_1 + 0.0018125 \cdot BF_2 - 0.0031154 \cdot BF_3 - 0.0013047 \cdot BF_4 - 0.00050019 \cdot BF_5$.

In the above equations, x_1, x_2 and x_3 denote the correlation values of R1-R2, R1-R3 and R2-R3, respectively.

As indicated by the correlation coefficient values in *Table 2*, resemblance between R1, R2 and R3 becomes more obvious when the value of smoothing parameter gets larger. Actually, this is an expected result since over-smoothed KDE estimates are generated with increasing values of *h*, which results in less variability, or i.e., higher similarity between the maps. On the other hand, for narrower choices of smoothing parameter, spikier estimates containing lots of spurious local structure are obtained. This behavior is illustrated in *Figure 7*.
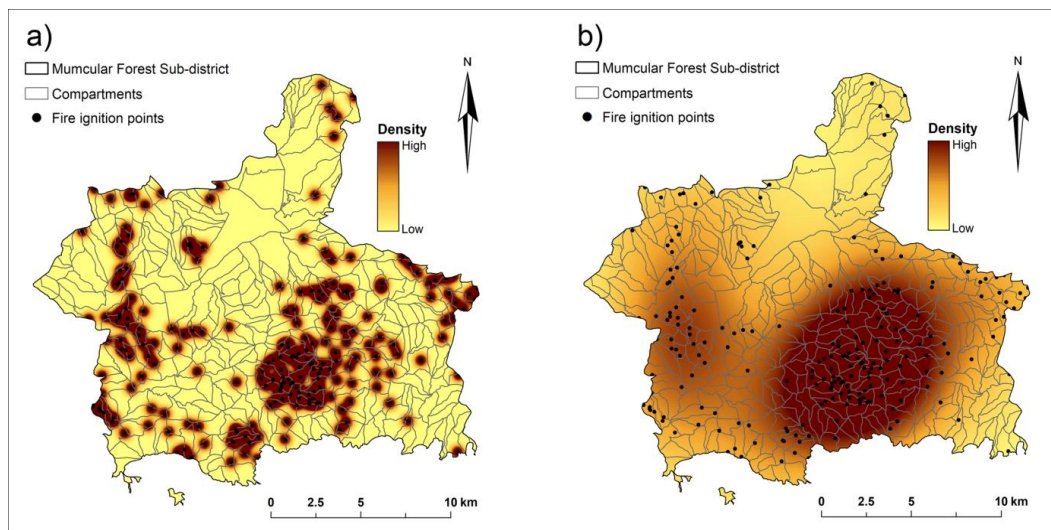
Figure 7. KDE maps of the study area: a) h = 250 m, and b) h = 2500 m

Results and discussion

Although the increments between successive correlation values with increasing *h* values are not so distinctive as depicted in *Table 2*, the amount of variation against increasing values of *h* can be observed very clearly in *Figure 6* for R1-R2, R1-R3 and R2-R3. A very sharp increase in the rate of change in correlation values are observed, and then the each curve exhibits a stabilized behavior above a certain *h* value (i.e., over-smoothing prevails).

The advantage of using MARS approach which employs piecewise linear splines in modelling is that the selection of knot locations is completely data-driven, i.e., specific to the dataset. Since the BFs that give smallest increase in the residual sum of squares are omitted from the model at each iterative step, an optimal model is generated at the end. The mean square error (MSE) versus GCV values for each MARS model shown in *Figure 6* is also illustrated in *Figure 8*. The change in MSE and GCV values during the iterations in the backward step can easily be observed in *Figure 8*. For a typical MARS

model building process, lines of MSE and GCV seem close to each other at the early stages. However, as the number of BFs included in the model increases, GCV diverges from MSE and rises up. The corresponding MSE and GCV values of the final MARS models are given *Table 3*.

Table 3. MSE and GCV values of the final MARS models

MARS model	MSE	GCV
R1-R2	0.01804	0.01812
R1-R3	0.01829	0.01842
R2-R3	0.01771	0.01779

As it is obvious in *Figure 6,* the specific knot location between the two consecutive BFs where the MARS model curve forms a kink shape indicates the specific value of *h* for that random point distribution. Further increase in *h* brings no significant change in the correlation values which simply indicates that the over-smoothing effect dominates in that specific KDE.
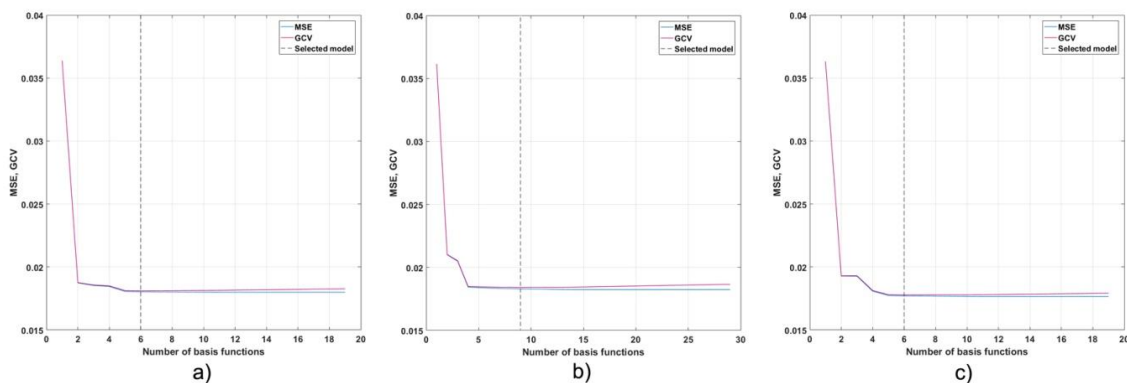
Figure 8. MSE vs. GCV values for the MARS models of a) R1-R2, b) R1-R3, and c) R2-R3

According to Silverman (1998), an appropriate choice of smoothing parameter is a crucial factor for the success of a KDE, and plotting several density estimates with various *h* values and choosing the one in accordance with the user's idea about the density often gives quite satisfactory results. This approach is also known as subjective choice of smoothing parameter (i.e., smoothing by eye). However, as again indicated in Silverman (1998), an inexperienced user may need somewhat an automated way to find a value of *h* which can be used as a reliable starting point for further subjective adjustment. In the studies of Koutsias et al. (2014) and Bajocco et al. (2017), smoothing parameter was chosen based on nearest neighbor distance method; however, this approach only generates rough estimates for the value of smoothing parameter when the inevitable positional uncertainties in the wildfire locations are considered.

On the other hand, the use of MARS has unique advantages in the modelling of complex environmental dynamics with uncertainties at some degree. Since the method is nonparametric, it does not make any specific assumption about the underlying functional relationship between the dependent and independent variables. Additionally, the selection of BFs is specific to the dataset in MARS which makes it and adaptive

procedure to handle large and complex datasets. During MARS-model building, the BFs are directly obtained from the observations and their space partitioning property results in an adaptive model. By this way, the flexible piecewise spline BFs of MARS can smoothly approximate the discrete data with noise and model its inherent characteristics. As it is obvious in *Figure 6*, our KDE approach integrated with MARS gives unique smoothing parameter values which can definitely be employed as reference point for further adjustment by the user.

The specific *h* values for R1-R2, R1-R3 and R2-R3 are 837, 976 and 922 m, respectively. Very similar *h* values from the three different random point distributions reveal that even though these three random point distributions are spatially distinctive from each other and their pair wise correlation values include stochastic noise, MARS is able to capture the inherent characteristic of the dataset by using smoothing and data-driven space partitioning properties of the piecewise linear BFs.

Thus, the final smoothing parameter value of the Gaussian KDE for the study area is taken as the average of these three, which is equal to 912 m. The seasonal KDE maps and the final KDE map of the study area are prepared accordingly by using *CrimeStat 3.3* software (Levine, 2004).

When the seasonal KDE maps are analyzed (cf. *Fig. 9*), fire events in the study area temporally and spatially exhibit different cluster patterns.
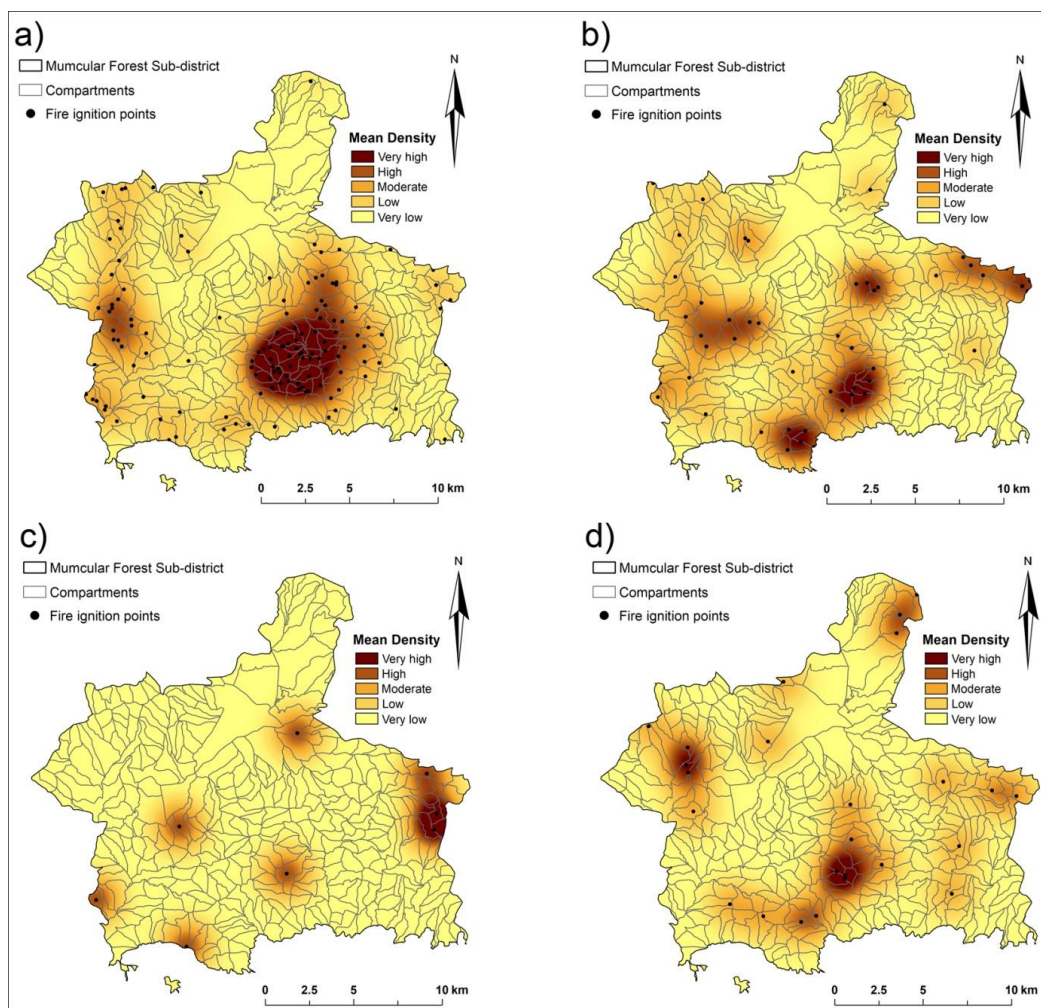
***Figure 9.*** *Seasonal KDE maps: a) summer, b) fall, c) winter, and d) spring*

The majority of forest fires (i.e., 67.6%) occur in the summer seasons (cf. *Fig. 9a*), which is obviously expected, and they are manly concentrated in the south-east central part, and small clusters are observed along the western part. Fire incidents that take place in summer are mostly located either in dense forested areas with closed canopy structure or around agricultural lands that are also close to settlement areas. Thus, these fire events would be associated with two factors: *i*) higher rate of human activities for recreational purposes during summer season, and *ii*) fires ignited on purpose by farmers for the removal of husk, which is often encountered in agricultural regions of Turkey.

Fire events that happen in fall have a 19.5% share in total incidents and exhibit a bit more homogenous spatial characteristic than summer season events (cf. *Fig. 9b*). They are mainly observed along the west side, in both south central and south parts, and small clusters exist along eastern side. A few events are observed in the dense forested areas; however, most of the incidents are located around the agricultural areas in the west and the north. This kind of spatial structure may arise from the raise in straw fire events.

Fire events in winter have the smallest share in all the seasons (i.e., 2.8%) and only limited number of events is observed in the central region, as well as the middle and upper parts of the western border (cf. *Fig. 9c*). Following the winter, spring fire events have an increasing trend in number (10.1%) and they are clustered in the central-south and eastern areas of the region, a few events are observed at northern cape as well (cf. *Fig. 9d*). Since the harvest season is not started yet in spring, these events can hardly be related with straw fires, but can be attributed to increasing recreational activities.

As the final step, the mean density map that shows the average intensity for each administrative polygon (i.e., compartment) is generated by employing the final intensity map of the area (cf. *Fig. 10*).
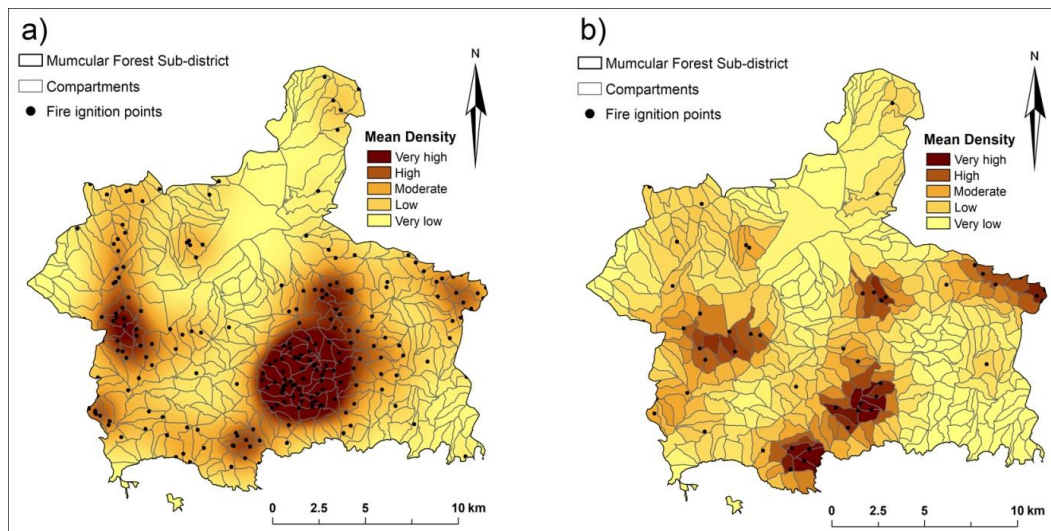
Figure 10. *a) Final density map, and b) mean density map of the study area*

Conclusion and outlook

Fire management planners always demand for elaborately prepared maps, which reveal the relation between wildfire events and several contributing characteristic attributes of the area of study, like topography, demographic structure, meteorological factors etc., for proper analysis of forest fire risk. However, historical fire data used in

this kind of analysis are mostly in point format and seriously suffer from positional inaccuracies. KDE method is often preferred for converting such data into continuous raster surface, and finding an appropriate smoothing parameter value that suits to the area under investigation would be of vital importance.

In this study, above mentioned points are addressed within the frame of Mumcular Forest Subdistrict case, and quite effective solutions are introduced in order to deal with them. These methods can easily be applied on different geographical areas.

The method employed to reduce the positional uncertainties is simply based on utilizing the temporal wildfire records at a finer spatial resolution, i.e., compartments in the stand map. Even though the geolocation of wildfire starting points at this spatial resolution may still include some level of uncertainty, an improvement is achieved by confining the wildfire starting locations into smaller administrative polygons.

In order to determine an appropriate smoothing parameter value h, a novel methodology is proposed by utilizing MARS for the first time in the KDE-based analysis of historical wildfire data. The results of our study indicate that MARS is an effective method for modelling dynamic and complex environmental processes, as in the case of wildfires, particularly, when expert judgment is not readily available. The results show that MARS is a proper choice to select suitable value of smoothing parameter for the KDE of wildfire events, and we definitely consider that the effectiveness of the MARS approach for dynamic modelling of wildfires should be further investigated by other case studies.

MARS provides important results on modelling the KDE of wildfires in Mumcular Forest Subdistrict in a 18-year period. It can enable us to obtain valuable and reliable information on spatiotemporal patterns of wildfires. This information is quite important for fire management planners' analytic studies on short- and long-term planning. As illustrated in this study, incorporating effective application of modern applied mathematics within geospatial analysis tools would have a significant contribution to fire departments and help them take more effective and timely measures for more risky areas.

In future researches, this study can be extended in order to produce a comprehensive wildfire assessment map of the study area by the inclusion of significant contributors to wildfires such as topographical, meteorological and anthropogenic factors together with KDE analysis. Within this phase of the study, topographic factors such as elevation, slope and aspect would be derived from the digital elevation model of the area. Meteorological data (i.e., temperature, rainfall and wind) would be acquired from the local ground stations in the area and interpolated by using kriging technique in order to produce the corresponding 2-D raster surfaces. For anthropogenic factors, distance from road networks and residential areas with certain population densities would be considered and realized by creating associated buffer zones.

REFERENCES

[1]   Alp, Ö. S., Büyükbebeci, E., Çekiç, A. İ., Özkurt, F. Y., Taylan, P., Weber, G.-W. (2011): CMARS and GAM & CQP-Modern optimization methods applied to international credit default prediction. – Journal of Computational and Applied Mathematics 235: 4639-4651.

[2]   Anderson, T. K. (2009): Kernel density estimation and K-means clustering to profile road accident hotspots. – Accident Analysis & Prevention 41(3): 359-364.

[3]   Anderson, N., Titterington, D. (1997): Some methods for investigating spatial clustering, with epidemiological applications. – Journal of the Royal Statistical Society: Series A (Statistics in Society) 160(1): 87-105.

[4]   Bailey, T. C., Gatrell, A. C. (1995): Interactive Spatial Data Analysis. – Longman Scientific & Technical, Essex.

[5]   Bajocco, S., Koutsias, N., Ricotta, C. (2017): Linking fire ignitions hotspots and fuel phenology: The importance of being seasonal. – Ecological Indicators 82: 433-440.

[6]   Bowman, A. W., Azzalini, A. (1997): Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations. – Clarendon Press, Oxford.

[7]   Camarero, J. J., Sangüesa-Barreda, G., Montiel-Molina, C., Seijo, F., López-Sáez, J. A. (2018): Past growth suppressions as proxies of fire incidence in relict Mediterranean black pine forests. – Forest Ecology and Management 413: 9-20.

[8]   Çevik, A., Weber, G.-W., Eyüboğlu, B. M., Oğuz, K. K. (2017): Voxel-MARS: a method for early detection of Alzheimer's disease by classification of structural brain MRI. – Annals of Operations Research 258(1): 31-57.

[9]   Chainey, S., Tompson, L., Uhlig, S. (2008): The utility of hotspot mapping for predicting spatial patterns of crime. – Security Journal 21(1-2): 4-28.

[10]   Chen, Q., Wynne, R., Goulding, P., Sandoz, D. (2000): The application of principal component analysis and kernel density estimation to enhance process monitoring. – Control Engineering Practice 8(5): 531-543.

[11]   Del Hoyo, L. V., Isabel, M. P. M., Vega, F. J. M. (2011): Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. – European Journal of Forest Research 130(6): 983-996.

[12]   Durmaz, M., Karslıoğlu, M. O., Nohutcu, M. (2010): Regional VTEC modeling with multivariate adaptive regression splines. – Advances in Space Research 46: 180-189.

[13]   Friedman, J. H. (1991): Multivariate adaptive regression splines. – The Annals of Statistics 19(1): 1-67.

[14]   Gonzalez-Olabarria, J. R., Brotons, L., Gritten, D., Tudela, A., Teres, J. A. (2012): Identifying location and causality of fire ignition hotspots in a Mediterranean region. – International Journal of Wildland Fire 21(7): 905-914.

[15]   Gralewicz, N. J., Nelson, T. A., Wulder, M. A. (2012): Spatial and temporal patterns of wildfire ignitions in Canada from 1980 to 2006. – International Journal of Wildland Fire 21(3): 230-242.

[16]   Hastie, T., Tibshirani, R., Friedman, J. (2009): The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). – Springer, New York.

[17]   Hély, C., Girardin, M. P., Ali, A. A., Carcaillet, C., Brewer, S., Bergeron, Y. (2010): Eastern boreal North American wildfire risk of the past 7000 years: A model-data comparison. – Geophysical Research Letters 37(14).

[18]   Horne, J. S., Garton, E. O., Krone, S. M., Lewis, J. S. (2007): Analyzing animal movements using Brownian bridges. – Ecology 88(9): 2354-2363.

[19]   Kagan, Y. Y., Jackson, D. D. (2000): Probabilistic forecasting of earthquakes. – Geophysical Journal International 143(2): 438-453.

[20]   Katajisto, J., Moilanen, A. (2006): Kernel-based home range method for data with irregular sampling intervals. – Ecological Modelling 194(4): 405-413.

[21]   Kie, J. G., Matthiopoulos, J., Fieberg, J., Powell, R. A., Cagnacci, F., Mitchell, M. S., Gaillard, J.-M., Moorcroft, P. R. (2010): The home-range concept: are traditional estimators still relevant with modern telemetry technology? – Philosophical Transactions of the Royal Society B: Biological Sciences 365(1550): 2221-2231.

[22]   Koutsias, N., Kalabokidis, K. D., Allgöwer, B. (2004): Fire occurrence patterns at landscape level: beyond positional accuracy of ignition points with kernel density estimation methods. – Natural Resource Modeling 17(4): 359-375.

[23] Koutsias, N., Balatsos, P., Kalabokidis, K. (2014): Fire occurrence zones: kernel density estimation of historical wildfire ignitions at the national level, Greece. – Journal of Maps 10(4): 630-639.

[24] Lasocki, S., Orlecka-Sikora, B. (2008): Seismic hazard assessment under complex source size distribution of mining-induced seismicity. – Tectonophysics 456(1-2): 28-37.

[25] Levine, N. (2004): CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations (Version 3.3). – Ned Levine & Associates, Houston (TX)/National Institute of Justice, Washington, DC.

[26] Nakaya, T., Yano, K. (2010): Visualising crime clusters in a space-time cube: an exploratory data-analysis approach using space-time kernel density estimation and scan statistics. – Transactions in GIS 14(3): 223-239.

[27] Nalcaci, G., Özmen, A., Weber, G. W. (2018): Long-term load forecasting: models based on MARS, ANN and LR methods. – Central European Journal of Operations Research 1-17. https://doi.org/10.1007/s10100-018-0531-1.

[28] Okabe, A., Satoh, T., Sugihara, K. (2009): A kernel density estimation method for networks, its computational method and a GIS-based tool. – International Journal of Geographical Information Science 23(1): 7-32.

[29] Özmen, A., Batmaz, İ., Weber, G.-W. (2014): Precipitation modeling by polyhedral RCMARS and comparison with MARS and CMARS. – Environmental Modeling & Assessment 19(5): 425-435.

[30] Özmen, A., Yılmaz, Y., Weber, G.-W. (2018): Natural gas consumption forecast with MARS and CMARS models for residential users. – Energy Economics 70: 357-381.

[31] Quirós, E., Felicísimo, Á. M., Cuartero, A. (2009): Testing Multivariate Adaptive Regression Splines (MARS) as a method of land cover classification of TERRA-ASTER satellite images. – Sensors 9: 9011-9028.

[32] Robertson, C., Nelson, T. A., MacNab, Y. C., Lawson, A. B. (2010): Review of methods for space–time disease surveillance. – Spatial and spatio-temporal epidemiology 1(2-3): 105-116.

[33] Rossiter, J. (1991): Calculating centile curves using kernel density estimation methods with application to infant kidney lengths. – Statistics in Medicine 10(11): 1693-1701.

[34] Silverman, B. W. (1998): Density Estimation for Statistics and Data Analysis. – Chapman & Hall/CRC, USA.

[35] Wand, M. P., Jones, M. C. (1993): Comparison of smoothing parameterizations in bivariate kernel density estimation. – Journal of the American Statistical Association 88(422): 520-528.

[36] Wand, M. P., Jones, M. C. (1995): Kernel Smoothing. – Chapman & Hall, London.

[37] Williamson, D., McLafferty, S., Goldsmith, J., Mallenkopf, J., McGuire, P. (2008): A better method to smooth crime incident data. – ESRI ArcUser Magazine January-March: 1-5.

[38] Worden, K., Manson, G., Allman, D. (2003): Experimental validation of a structural health monitoring methodology: Part I. Novelty detection on a laboratory structure. – Journal of Sound and Vibration 259(2): 323-343.

[39] Worton, B. J. (1989): Kernel methods for estimating the utilization distribution in home-range studies. – Ecology 70(1): 164-168.

[40] Xie, Z., Yan, J. (2008): Kernel density estimation of traffic accidents in a network space. – Computers, Environment and Urban Systems 32(5): 396-406.

[41] Yu, L., Su, Z. (2012): Application of kernel density estimation in lamb wave-based damage detection. – Mathematical Problems in Engineering 2012: #406521. http://dx.doi.org/10.1155/2012/406521.

[42] Zou, K. H., Hall, W., Shapiro, D. E. (1997): Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests. – Statistics in Medicine 16(19): 2143-2156.