

Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

(OBIFS) isotropic image analysis for improving a predicting agent based systems

M. Dolores Muñoz*, Aitor Mata, Emilio Corchado, Juan Manuel Corchado

Departamento de Informática y Automática, Universidad de Salamanca, Plaza de la Merced, s/n, Salamanca, Spain

ARTICLE INFO

Keywords:

Isotropic image analysis
Hybrid intelligent systems
Organization-based system
Forest fires

ABSTRACT

In this interdisciplinary study a novel hybrid forecasting system is presented, in which an isotropic buffer operator is applied for case-based creation within the structure of the organization-based multi-agent system. Commonly used as an image analysis technique by commercial Geographic Information Systems (GIS), the buffer operator in this particular system calculates the area of a forest fire for prediction and visualization tasks. The use of the buffer operator improves the quality of the data used by the system and in consequence the quality of the results obtained. The system has been successfully tested using real historical data on forest fires evolution, by generating accurate predictions.

Published by Elsevier Ltd.

1. Introduction

Prediction and simulation of forest fires propagation is an important problem from the computational point of view (Denham et al., 2008) due to the complexity of the involved models, the necessity of numerical methods and the required resources for calculation. The phenomenon of forest fires has not only given as a result an important loss of forests and damage in the economy, but it has also seriously affected human health and environment (Pausas et al., 2009). The fire fighting should have at its disposal the most advanced resources and tools to help the use of available resources in the most efficient way to diminish fire effects as much as possible.

Once a fire is detected, it is crucial to visualize the terrain and monitor the evolution of the fire in real-time. Nowadays a wide variety of computer vision techniques have been developed for forest fire monitoring involving measurement of forest fire features (fire front, flame height, fame indication angle, fire base width) required for effective and safe fire fighting (Martínez-de Dios et al., 2008).

This study presents OBIFS, a novel Organization Based (Ferber et al., 2004) Isotropic Forecasting System. This system deploys a prediction model which makes use of intelligent agents (Nuxoll and Laird, 2011) and Case-Based Reasoning systems (CBR) to determine the possibility of detecting forest fires in a certain area once a fire has nearby started. OBIFS relies on meteorological and environmental observations, to predict the forest fire spread patterns and to develop adequate response strategies for fire fighting. The

system applies a distributed multiagent architecture based on Service Oriented Architectures (SOA) (Josuttis, 2007), modelling most of the system's functionalities as independent applications and services. These functionalities are invoked by deliberative agents (Carrascosa et al., 2007) acting as coordinators. The system presented in this research generates as a solution the future situation of an area, considering the present one. Predictions are created using a CBR system (Kolodner, 1993). The cases used by the CBR systems contain atmospheric data (wind, temperature and pressure) and may also contain information about a specific situation to be solved (e.g. forest fires). OBIFS combines artificial intelligence techniques in order to improve the efficiency of the CBR systems (Tung et al., 2010), thus generating better results.

OBIFS has been trained using historical data (Gestosa, 2000). The development of agents is an essential piece in the analysis of data from distributed sensors (Martínez-de Dios et al., 2008) and gives those sensors the ability to work together and analyze complex situations.

The term buffer operator, also known as the buffering or influence zone is used by all customized Geographical Information Systems (GIS) (Huang and Li, 2009; Liu, 1992). This operator is defined as the geometric space of the points that are at a shorter or similar distance to a given object (point, polyline or polygon) (Chou, 1997). This definition is isotropic or directionally uniform, since the distance of the object to the edge of the buffer is constant in any direction on the plane. Among other fields (Longley, 2005), this operator is used in the simulated display of environmental processes, such as surveys of pesticides and chemical fertilizers contamination of surface water in the watershed, the influence of nitrate levels sediments on the growth of local flora. It might also be a useful tool for the identification and evaluation of environmental impact on engineering projects, to identify high-risk seismic areas as well as in

* Corresponding author. Tel.: +34 609744149.

E-mail addresses: mariado@usal.es (M. Dolores Muñoz), aitor@usal.es (A. Mata), escorchado@usal.es (E. Corchado), corchado@usal.es (J.M. Corchado).

studying the spread of certain diseases (Dun et al., 2007). The joint work based on the use of GIS and decision support systems provides the information necessary to understand how certain factors may influence the environmental impact (Cetin and Yusuf, 2011; Herrero-Jiménez, 2012).

There are two methods for the generation of influence areas: Voronoi triangulation (Atsuyuki et al., 1992; Sun et al., 2011) and the Minkowski sum (González, 2000; Schou, 1978). In the latter method, a secondary polygon or generating polygon is defined as located on a point or moving on a polyline or polygon and generating a surface formed by the points over which the generating polygon moves. In the isotropic buffer, the generating polygon is a circle, which implies a constant distance between the border of the buffer and the object.

In the following section, a brief description of the forest fires problem is shown. Section 3 presents the main concepts of the organizations of agents. The fourth section develops the concept of isotropic image analysis, introducing the main terms of the directional statistics, after which the fourth section describes the use of the buffer operator in the existing GIS. The two following sections describe the OBIFS system and the final two sections present the results, the conclusions and future research lines.

2. Forest fires

Forest fires are a very serious hazard that, every year, causes significant damage around the world from the ecological, social, economical and human point of view (Long, 2001). These hazards are particularly dangerous when meteorological conditions are extreme with dry and hot seasons or strong wind. For example, fire is a recurrent factor in Mediterranean areas. Fires represent a complex environment, where multiple parameters are involved.

In this sub-section, a series of applications and possible solutions are explained. There are different approaches to the forest fire problems, including all the main phases existing in the evolution of this kind of problem (Alexandridis et al., 2008; Cheng and Wang, 2008; Denham et al., 2008). The OBIFS system presented here was applied to generate predictions in a forest fire scenario.

Forest fires represent a great environmental risk. The main approaches that have been used to solve this problem begin with the detection of the fires (Mazzeo et al., 2007) where different techniques have been applied. Once the fire is detected, it is important to generate predictions that should assist in making a decision in those contingency response situations (Iliadis, 2005). Finally, there are complex models that tackle the forest fire problem by trying to forecast its evolution and to minimize its associated risks (Serón et al., 2005). These models may also assist in planning the recuperation of the burner areas (Alonso-Betanzos et al., 2003). The data used to analyse the OBIFS system were a subset of the available data that had not been previously used in the training phase. The predicted situation (Cetin and Yusuf, 2011) was contrasted with the actual future situation as it was known (historical data was used to train the system and also to test its correction). The proposed solution, in most of the variables, had a near 90% accuracy rate.

To create the cases, the geographical area being analyzed was divided into small squares, each of which was considered as a case, with its associated parameters shown in Table 1. The squares determine the area to be considered in every case. The problem is represented by the current situation of the area (all its parameters and the presence, or lack thereof, of fire). The solution is represented by the situation in that area in a future moment (same location but parameters changed to those for the next day or the next step, if less than a day is considered in every step-). The data used are part of the SPREAD project, in particular the Gestosa field experiments that took place in 2002 and 2004 (Gestosa, 2005). The

Table 1
Variables that define a case, applying OBIFS to the forest fire problem.

Variable	Definition	Unit
Longitude	Geographical longitude	Degree
Latitude	Geographical latitude	Degree
Date	Day, month and year of the analysis	dd/mm/yyyy
Bottom pressure	Atmospheric pressure of the area	Newton/m ²
Temperature	Celsius temperature in the area	°C
Meridional wind	Meridional direction of the wind	m/s
Zonal wind	Zonal direction of the wind	m/s
Wind strength	Wind strength	m/s
Area of the fire	Surface covered by the fires	km ²

experiments of the Gestosa field began in 1998 and finished in December 2004. They were aimed at collecting experimental data to support the development of new concepts and models, and to validate existing methods or models in various fields of fire management.

3. Organizations of agents

In the multiagent field, the term organization (García et al., 2008; Ye et al., 2012) has been mainly used to describe a set of agents that, using some kind of roles, interact with each other coordinating themselves to achieve the global objectives of the system. (Gasser, 2001) assumes that organizations are structured systems with activity, knowledge, culture, history, and ability pattern, different of any particular agent (Gestosa, 2000). Organizations exist in a completely different level than individual agents that make up the organizations themselves. Individual agents are replaceable. Organizations are established in a space; it either be geographical, temporal, symbolic, etc. So, an organization of agents provides a kind of workspace for the activity and interaction of the agents by defining roles, behavioural expectative and relations. Ferber indicates that organizations proportionate a way to divide the system, creating groups or units that form the interaction context of the agents (Ferber et al., 2004). The organization is then based in two main aspects: structural and dynamic. The structure of the organization represents the remaining components when the individual elements enter or leave the organization. The organization is composed by the set of relationships that allow seeing a number of different elements as unique. The structure defines the way the agents are grouped in organizational units and how those units are related with each other. The roles needed to develop the activities of the organization are also defined in the structure, as long as the relationships and restrictions. The organizational dynamics is focused in the interaction patterns defined for the roles, describing the way to get into or to leave the organization, the parameters of the roles and the way the roles are assigned to the agents. Finally, V. Dignum, affirms that the organizations of agents assume that there are global objectives, different from the individual agents' objectives (Dignum and Dignum, 2007). Roles represent organizational positions that help to achieve those global objectives. Agents may have their own objectives and decide if they take any specific system role or not, determining which among the available protocols are more suitable to achieve their chosen objectives.

4. Directional statistics

Directional statistics (Nishino, 2009) is concerned with data that represent vectors on a plane or in 3D space. In the former case, the sample space is a circle and in the second, a sphere. In order to take the structure of these sampling spaces into account, specific statistical methods are needed (Batschelet, 1981; Mardia and Jupp, 2000). Some examples of circular data are the directions of the predominant winds, the flight of migratory birds; indeed, any data

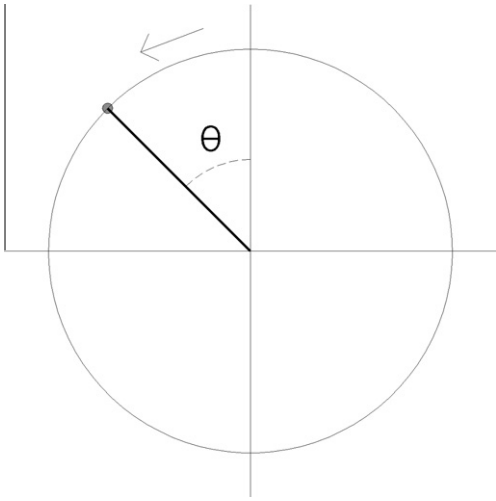


Fig. 1. Representation of a unitary vector on the plane on a circular graphic.

that can be measured and converted into degrees or radius may be considered circular data.

Circular data may be represented as a point on a unit circle or as a unitary vector on the plane. Take a direction as the origin and choosing a direction, each circular datum may be specified as having an angle θ between the initial direction and the direction that corresponds to the datum (Fig. 1). The direction of the vector is characterized by de angle θ . The simplest way of representing circular data is to draw them as points on the unit circle and, when a direction repeats, place the news points outside the circle on the corresponding radius (Fig. 2).

Circular variables can be defined (Batschelet, 1981; Downs and Mardia, 2002; Jammalamadaka and Sengupta, 2001; Singh et al., 2002; Zar, 1999) as those that represent directions on the plane, which is quantified by angles that vary from 0 to 2π . One of the most important differences in comparison with linear variables is that while these can take values of the whole straight line, circular variables take cyclic values and, consequently, the sum or difference of observations can surpass 360° or can even take a negative value. It begins possible in these cases to find an equivalent value in the interval $0\text{--}360^\circ$. This characteristic means that circular variables may be treated in a different way to the linear ones, through statistical methods correlation analysis and specific distributions for this type of variables.

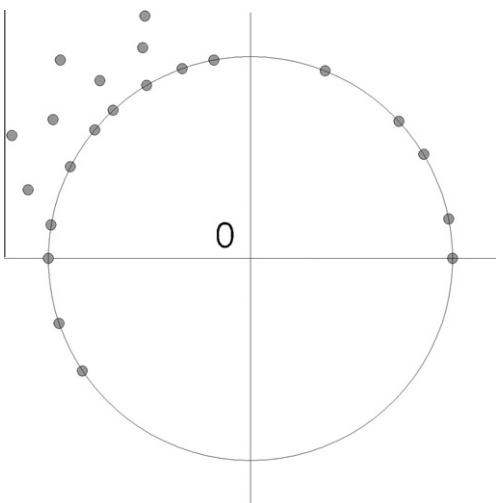


Fig. 2. Representation of a circular data as points on the unit circle.

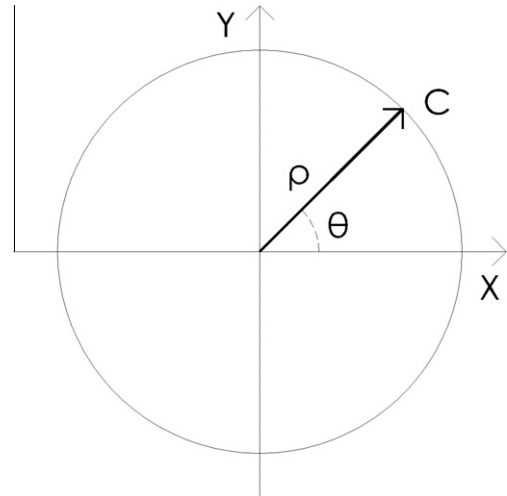


Fig. 3. Circular distribution with mean vector of module r and mean angle θ . C is the mass centre of the distribution.

Conceptually, a circular distribution can be considered a bivariated linear distribution where the total probability (or total mass) is dispersed around the circle unit. Therefore, as in the bivariated linear statistic, a mean vector \bar{m} of module r and mean angle $\bar{\theta}$ exists in circular statistics, at the tip of which is the mass centre C of the distribution (Fig. 3).

4.1. Mean vector

The mean statistical vector \bar{m} is calculated by assigning to each of the points n of the unit radius circle in Fig. 4, a mass M , from which the centre of mass or gravity C is calculated. The projections \bar{x} and \bar{y} of \bar{m} are giving by the following equations:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^j n_i \cos \Phi_i \quad (1)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^j n_i \sin \Phi_i \quad (2)$$

$$n = \sum_{i=1}^j n_i \quad (3)$$

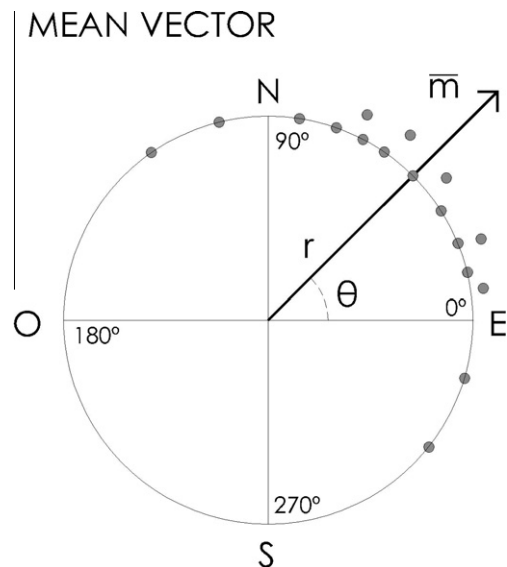


Fig. 4. Mean vector.

Let $Z(\Phi)$ be the random variable. If it is taken a mono-modal sample of frequencies n_1, n_2, \dots, n_j in the directions $\Phi_1, \Phi_2, \dots, \Phi_j$, the mean vector $\bar{m}(r, \bar{\Phi})$ is defined by the following expressions:

$$r = \sqrt{\bar{x}^2 + \bar{y}^2} \tag{4}$$

$$\bar{\Phi} = \begin{cases} \text{Arc tan}[\bar{y}/\bar{x}] & \text{si } \bar{x} > 0 \\ 180 + \text{Arc tan}[\bar{y}/\bar{x}] & \text{si } \bar{x} < 0 \end{cases} \tag{5}$$

where \bar{x} and \bar{y} are the projections of \bar{m} on the axes X and Y respectively. The direction of the vector is found on the straight line that joins C with the coordinates of the origin O . When the data are grouped in j arcs with a length of $\lambda = \frac{2\pi}{j}$, the values of \bar{x} and \bar{y} are as follows:

$$\bar{x} = 1/n[n_1 \cos \Phi_1 + n_2 \cos \Phi_2 + \dots + n_j \cos \Phi_n] \tag{6}$$

$$\bar{y} = 1/n[n_1 \text{sen} \Phi_1 + n_2 \text{sen} \Phi_2 + \dots + n_j \text{sen} \Phi_n] \tag{7}$$

where, n_1, n_2, \dots, n_j are the frequencies of the mean points $\Phi_1, \Phi_2, \dots, \Phi_j$ of the arcs. The value θ is estimated in the same way as for the grouped samples. However, it needs a correction factor. Without correction, r tends to be a little smaller. Therefore, r has to be multiplied by a factor $c > 1$. The corrected module $r_c = c \cdot r$, where the correction factor c is:

$$c = \lambda/2 / \sin(\lambda/2) \quad (\lambda \text{ expressed in radians}) \tag{8}$$

The calculation described for \bar{m} is valid for mono-modal samples. There exist many algorithms that allow to obtain the better fitting circles to the data graphically shown in Fig. 4. In Ladrón de Guevara et al. (2011), an analysis of currently used algorithms used to make that adjustment is done and a new one is proposed, based on left and right side partial derivatives. However, the experience shows that the phenomena linked to orographic discontinuities may be plurimodal. The process of ν -modal samples (where ν is the number of modes) differs from what is described, as these should be treated as if they were samples generated by ν unimodal distributions, which is why it can be considered from a mixture of distributions point of view.

The ν -modal samples should be considered as extracted from a distribution that is generated by the overlap of ν monomodal distributions. When the distances between modes are arbitrary, no standard methods exist to breakdown a ν -modal sample into ν -mono-modal samples; in nature, plurimodal samples usually appear as bimodal and diametrically opposed. In this case, it is possible to reduce the bimodal sample to a monomodal sample by duplicating the angles. With the new angles, the mean vector $\bar{m}_2(r_2, \bar{\Phi}_2)$ is calculated by using expressions (1)–(3). In order to obtain the symmetric modal angle $\bar{\Phi}_1$ from the original sample, the effect of the duplication of the angles has to be cancelled, as follows:

$$\bar{\Phi}_1 = \bar{\Phi}_2/2 \quad \text{or} \tag{9}$$

$$\bar{\Phi}_1 = \bar{\Phi}_2/2 + 180^\circ \tag{10}$$

4.2. Von Mises distribution

Among the existing circular distributions (Schou, 1978; Taylor, 2008), one of the most widely used for the modeling of circular variables is the Von Mises distribution, in which the density function for ν -modal and symmetric samples is showed by the following expression

$$f(\Phi) = 1/2\pi I_0(k)\text{Exp}[k \cos \nu(\Phi - \theta)] \tag{11}$$

where I_0 represented the Bessel function of imaginary pure argument of order 0, ν is the number of modes, and k is the concentration parameter (Ros et al., 2007; Taylor, 2008), that indicates the

extent to which the distribution around the dominant direction θ is concentrated.

For $k = 0$, $f(\Phi)$ degenerates in an uniform distribution. Mardia demonstrated (Mardia and Jupp, 2000) that the maximum likelihood estimation $\hat{\theta}$ and $\hat{\rho}$ for parameters θ and ρ of a Von Mises distribution are respectively $\bar{\Phi}$ and r . Likewise,

$$I_1(\hat{k})/I_0(\hat{k}) = r \tag{12}$$

is fulfilled. Hence, the solution to this expression is the maximum likelihood \hat{k} .

4.3. Minkowski sum

Given two images A and B in R^2 , the sum of Minkowski is defined as (Schou, 1978).

$$A \oplus B := \bigcup_{b \in B} A + b \tag{13}$$

where A is the generating polygon, and B the skeleton or primary element (point, polyline, or polygon. $A \oplus B$ is generated moving A though each element $b \in B$, and then by adding the result of all the translations later on. The translation of the generating polygon A trough the element $b \in B$ is defined as

$$A + b := \{a + b, a \in A\} \tag{14}$$

If a circle is taken as the generating polygon A , and the group of points $B = \{(2,3), (3,4), (2,5), (1,5)\}$ as the primary element, then the Minkowski sum is represented by the following expression:

$$A \oplus B = [(A + (2, 3)) \cup (A + (3, 4)) \cup (A + (2, 5)) \cup (A + (1, 5))] \tag{15}$$

Fig. 5 shows the result, $A \oplus P$, additions which have taken polygon P as primary elements.

Conceptually, the Minkowski sum is a dilation or expansion of the primary image B , whose form is determined by the generating polygon A . In the previous example a circle has been chosen as the generating image. The expansion of the primary image is directionally uniform or isotropic, since the generating image is a symmetrical figure with regard to both axes. Most of the times, the complexity is in the choice of the centre coordinates within the buffers. Normally, an aleatory point is chosen to become the centre of the first buffer. Each point is then assigned to the nearest centre, and each centre is recalculated as the mean of all points assigned to it. In most cases, some not necessary calculations are done in (Bretin et al., 2011) proposing an algorithm based on left and right side partial derivatives that allow to accelerate the method to choose the centres of the new buffers.

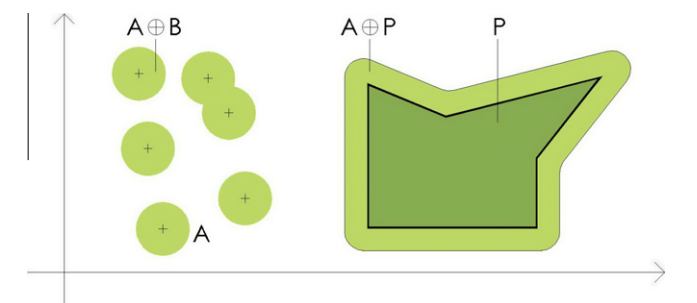


Fig. 5. Examples of the Minkowski sum taking a circle as the generating geometric object.

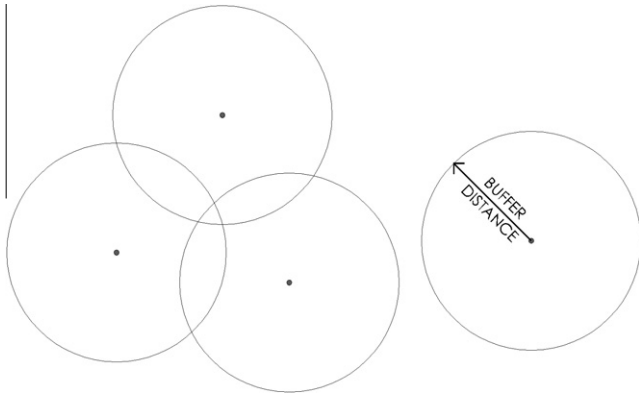


Fig. 6. Area of influence of specific data with a fixed distance.

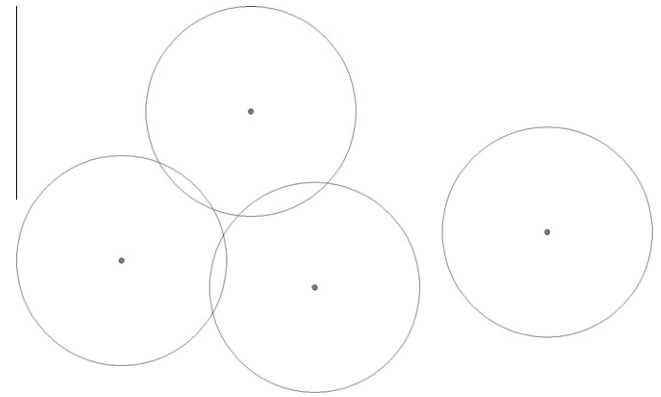


Fig. 8. Calculation of the intersection of areas of influence.

5. Buffers in the commercial GIS

The majority commercial GIS have one or various modules that generate areas of isotropic influence (Esri, 2012). There are three methods to generate the buffer or the zone of influence: the creation of areas of influence of specific data, the creation of areas of influence of linear data and the creation of areas of influence of polygonal data. This study is focused on the creation of areas of influence of specific data.

5.1. The creation of areas of influence of specific data

Very frequently, GIS requires the generation of areas of influence in certain operations that analyze spatial data. The simplest area of influence is generated around specific data, as the process only implies the creation of “circular” polygon around each point, with a radius that is equivalent to the distance from the buffer. There are two ways of assigning the width of the buffer. The first one applies a fixed buffer distance (specified by the user) to all points of a layer (Fig. 6).

The second one (Fig. 7) consists of assigning an individual distance to each point based on the attributes of another layer of the system (its weighting or weight). The width attributes of the area of influence in this case will be stored in a feature attribute table as shown on Fig. 7. The number of the first row of the table is used to identify the buffer, and the second row, indicates the width of the buffer measured in meters or kilometres.

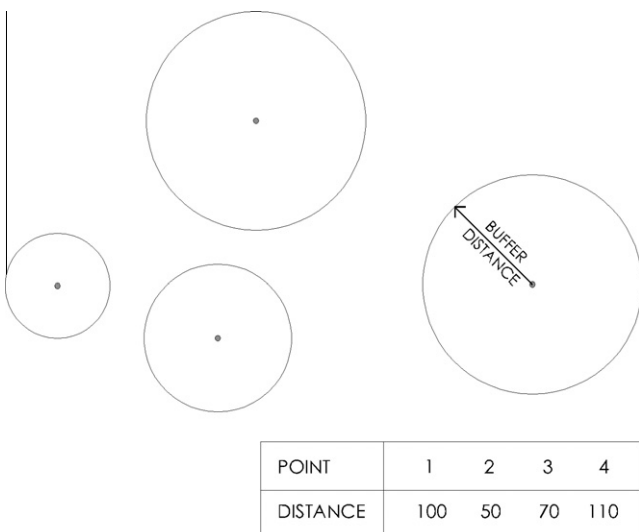


Fig. 7. Area of influence of specific data with layer attributes.

In case that multiple points need to be analyzed in the layer, the system should test the existence of overlaps between the areas of each point (Fig. 8). These overlaps should be deleted in such a way that the result is a polygonal layer that represents the zone covered by the union between all the areas. Hence, this procedure implies the application of two additional operations: intersection (Fig. 8) and dissolution (Fig. 9).

The creation of buffers gives as a result a new polygonal-type layer in the system, which represents the areas of influence generated from both weighted and fixed value distances. The resulting table of polygons will contain the identifiers of the polygons created in the procedure, and a new attribute that indicates whether the polygon is found inside or outside the area of influence (Fig. 10). If the polygon is within the area of influence is represented by 1, otherwise it is represented as 0. In this case it is necessary to model those areas that are not adapted to the circular shape.

In (Emre Celebi et al., 2007) an algorithm is shown, used to simulate those kind of shapes, taking into account not only the point within the circle, but also those that are tangent to the polygon that circumscribes them.

6. System description

In this study, a new *Organization Based System for Forest Fire Forecasting (OBIFS)* is presented. It is formed by an organization of agents that connects some different interface agents, dedicated to interact with the users in different platforms, to an inner CBR system, made by different services, implemented by a series of agents that answer to the different requests of the users. In OBIFS, the data collected from the Gestosa (Gestosa, 2005) experiments is

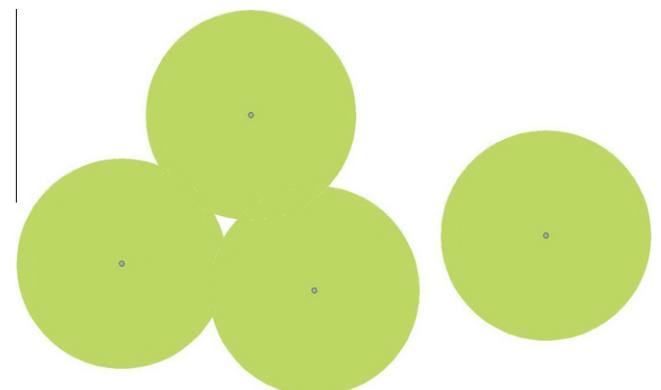


Fig. 9. Calculation of the dissolution of areas of influence of specific data.

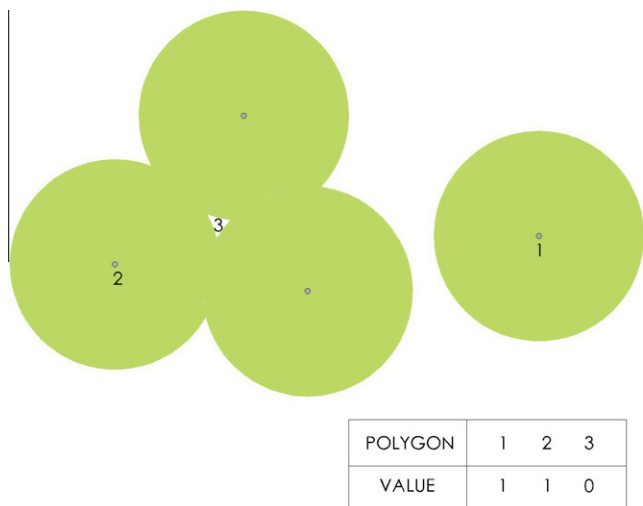


Fig. 10. Calculation of the dissolution of areas of influence of specific data with layer attributes.

processed and structured as cases. Table 2 shows the main variables that defines a case. Cases are the key to obtain solutions to future problems through a CBR system.

The functionalities of OBIFS are accessed using different interfaces executed on PCs or smart phones. Users interact with the system by introducing data, requesting a prediction or revising a solution generated (i.e., prediction).

The interface agents communicate with the services through the agents' platform and vice versa. The interface agents perform all the different functionalities which users make use for interacting with OBIFS. The different phases of the CBR system have been modeled as services, so each phase can be independently requested. For example, one user may only introduce information in the system (e.g. a new case), while another user could request a new prediction.

All information is stored in the case base and OBIFS is ready to predict future situations. A problem situation must be introduced in the system for generating a prediction.

Then, the most similar cases to the current situation are retrieved from the case base. Once a collection of cases are chosen from the case base, they must be used for generating a new solution to the current problem. Growing Radial Basis Functions Networks (Idri et al., 2010; Karayiannis and Mi, 1997) are used in OBIFS for combining the chosen cases in order to obtain the new solution. OBIFS determines future probabilities in a certain area. OBIFS divides the area to be analyzed in squares of approximately half a degree side for generating a new prediction. Then, the system calculates the demanded parameters in each square. The squares are coloured with different gradation depending on the values of the requested parameters.

Table 2
Variables that define a case.

Variable	Definition	Unit
Longitude	Geographical longitude	Degree
Latitude	Geographical latitude	Degree
Date	Day, month and year of the analysis	dd/mm/yyyy
Bottom pressure	Atmospheric pressure of the area	Newton/m ²
Temperature	Celsius temperature in the area	°C
Meridional wind	Meridional direction of the wind	m/s
Zonal wind	Zonal direction of the wind	m/s
Wind strength	Wind strength	m/s
Area of the fire	Surface covered by the fires	km ²

Fig. 11 shows the structure of OBIFS. There are three basic blocks in OBIFS: Interface Agents, Services and Agents' Communication Structure. The system covers from the users interfaces, solved by the interface agents, to the data management, solved by the services that can access the case base. These blocks provide all the system functionalities and are described as follows.

First *Interfaces Organization* represents all the programs that users can use to exploit the system functionalities. Applications are dynamic, reacting differently according to the particular situations and the services invoked. They can be executed locally or remotely, even on mobile devices with limited processing capabilities, because computing tasks are largely delegated to the agents and services.

Then the *CBR Services Organization* represents the activities that the architecture offers. They are the bulk of the functionalities of the system at the processing, delivery and information acquisition levels. Services are designed to be invoked locally or remotely and can be organized as local services, web services, GRID services, or even as individual stand alone services. OBIFS has a flexible and scalable directory of services, so they can be invoked, modified, added, or eliminated dynamically and on demand.

It is absolutely necessary that all services follow a communication protocol to interact with the rest of the components. CBR systems have proved to be quite efficient in prediction tasks in environmental situations (Baruque et al., 2010).

Finally, the *Communication Organization* integrates a set of agents, each one with special characteristics and behaviour. An important feature in this architecture is that the agents act as controllers and administrators for all applications and services, managing the adequate functioning of the system, from services, applications, communication and performance to reasoning and decision making. In OBIFS, services are managed and coordinated by deliberative BDI (Belief Desire Intention) (Mikic Fonte et al., 2012). The agents

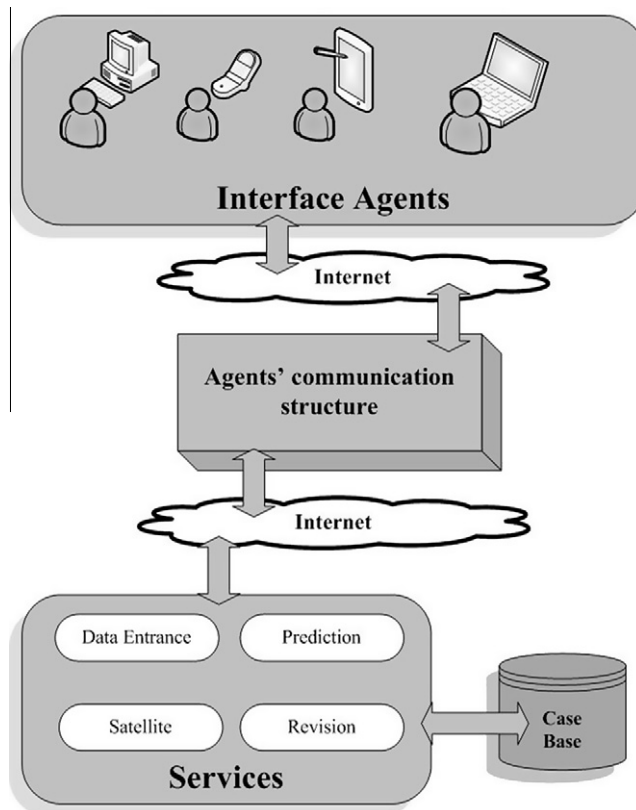


Fig. 11. OBIFS structure.

modify their behaviour according to the users' preferences, the knowledge acquired from previous interactions, as well as the choices available to respond to a given situation. The communication protocol allows applications and services to communicate directly with the Agents Platform. This protocol is based on SOAP specification to capture all messages between the platform and the services and applications (Cerami, 2002). Services and applications communicate with the Agents Platform via SOAP messages. A response is sent back to the specific service or application that made the request. All external communications follow the same protocol, while the communication among agents in the platform follows the FIPA Agent Communication Language (ACL) (Boella et al., 2011) specification. Agents, applications and services in OBIFS can communicate in a distributed way, even from mobile devices. This makes it possible to use resources no matter its location. It also allows the starting or stopping of agents, applications, services or devices separately, without affecting the rest of resources, so the system has an elevated adaptability and capacity for error recovery. Users can access to OBIFS functionalities through distributed applications which run on different types of devices and interfaces (e.g. computers, smart phones). Next, the disaggregation of the CBR methodology into services is explained. The four main phases of the CBR cycle (Recio-García et al., 2010) are divided into different services covering the entrance of data, the recovery of information from the case base and the solution generation and validation. The data used to check the OBIFS system were a subset of the available data that had not been previously used in the training phase. The predicted situation was contrasted with the actual future situation as it was known (historical data was used to train the system and also to test its correction). The proposed solution, in most of the variables, for the real data set (OBIFS) tested in this study had a near 90% accuracy rate (see Section 8).

7. System dynamics

In the previous section, the main elements in the different organizations of agents that form this system have been explained, detailing the differences between those organizations, and how the information flows from one to another.

In this section, the internal working of the *Services Organization* is explained, making a special point in the data dynamics and how the predictions produced by the system are generated depending on both the data available and the problem to be resolved.

7.1. Data introduction and case base creation

When data about forest fire is introduced, OBIFS must complete the information about the area including atmospheric and weather information. OBIFS uses Fast Iterative Kernel PCA (FIK-PCA) which is an evolution of PCA (Corchado et al., 2008; Gunter et al., 2007). This technique reduces the number of variables in a set by eliminating those that are linearly dependent, and it is quite faster than the traditional PCA.

The variables shown in Tables 1 and 2 represent the necessary information to eventually predict the evolution of forest fire. The *area of the fires* parameter in the case base represents the parameter that is going to be predicted. That prediction will be based on the values of the rest of the parameters (e.g. latitude, longitude, meridional wind, zonal wind...) which are the inputs of the neural network that will generate the future value of the *area of the fires* parameter, that is, the solution to the problem. The future situation of the *area of the fire* parameter will be generated from the rest of the parameters of the problem introduced in the system and also of the cases retrieved from the case base.

To calculate the areas to take into account when storing the case information in the case base, an isotropic buffer has been used

to divide the whole analyzed area in smaller pieces that make it easier and more accurate to calculate the relative dates and, thus, to improve the quality of the stored information, what will generate better and more accurate predictions.

To improve the convergence of the Kernel Hebbian Algorithm (KHA) used by Kernel PCA, FIK-PCA set η_t proportional to the reciprocal of the estimated values. Let $\lambda_t \in R_+^r$ denote the vector of eigenvalues associated with the current estimate of the first r eigenvectors. The new KHA algorithm sets the i th component of η_t to:

$$[\eta_t]_i = 1/[\lambda_t]_i \tau / t + \tau \eta_0 \quad (18)$$

The final variables are, obviously, linearly independent and are formed by combination of the previous variables. The values of the original variables can be recovered by doing the inverse calculation to the one produced to obtain the new variables. The variables that are less used in the final stored variables are those whose values suffer less changes during the periods of time analysed (salinity, temperature and pressure do not change from one day to another, then, they can be ignored considering that the final result does not depend on them).

When introducing the data into the case base, Growing Cell Structures (GCS) (Fritzke, 1994) are used. GCS can create a model from a situation organizing the different cases by their similarity. If a 2D representation is chosen to explain this technique, the most similar cells (i.e., cases) are near one of the other. If there is a relationship between the cells, they are grouped together, and this grouping characteristic helps the CBR system to recover the similar cases in the next phase. When a new cell is introduced in the structure, the closest cells move towards the new one, changing the overall structure of the system.

The weights of the winning cell, ω_c , and its neighbours, ω_n , are changed. The terms ε_c and ε_n represent the learning rates for the winner and its neighbours, respectively, x represents the value of the input vector and the terms ε_c and ε_n represent the learning rates for the winner and its neighbors, respectively.

$$\omega_c(t+1) = \omega_c(t) + \varepsilon_c(x - \omega_c) \quad (19)$$

$$\omega_n(t+1) = \omega_n(t) + \varepsilon_n(x - \omega_n) \quad (20)$$

Once the case base has stored the historical data, and the GCS has learned from the original distribution of the variables, the system is ready to receive a new problem.

When a new problem comes to the system, GCS are used once again. The stored GCS behaves as if the new problem would be stored in the structure and finds the most similar cases to the problem introduced in the system. In this case, the GCS does not change its structure because it has been used to obtain the most similar cases to the introduced problem. Only in the retain phase the GCS changes again, introducing the proposed solution if it is correct.

7.2. Prediction generation after a requests

When a prediction is requested by a user, the system starts recovering from the case base the most similar cases to the problem proposed. Then, it creates a prediction using Growing Radial Basis Function networks (GRBF) (Ros et al., 2007). Once the most similar cases are recovered from the case base, they are used to generate the solution.

GRBF networks are used to obtain the predicted future values corresponding to the proposed problem. This adaptation of the RBF networks allows the system to grow during training gradually increasing the number of elements (prototypes) which play the role of the centres of the radial basis functions. The creation of the Growing RBF must be made automatically which implies an

adaptation of the original GRBF system. The error for every pattern is defined by (21).

$$e_i = l/p \sum_{k=1}^p \|t_{ik} - y_{ik}\| \quad (21)$$

where t_{ik} is the desired value of the k th output unit of the i th training pattern, y_{ik} the actual values of the k th output unit of the i th training pattern. Once the GRBF network is created, it is used to generate the solution to the proposed problem. The solution proposed is the output of the GRBF network created with the retrieved cases. The input to the GRBF network, in order to generate the solution, is the data related with the problem to be solved, the values of the variables stored in the case base.

8. Evaluation of the results

OBIFS uses different artificial intelligence techniques to cover and solve all the phases of the CBR cycle. FIK-PCA is used to reduce the number of variables stored in the system, getting about a 60% of reduction in the size of the case base if the initial number of elements of the case base is over 3000, which is the point when the algorithm may reduce redundancy in a consistent way. This adaptation of the PCA also implies a faster recovery of cases from the case base (more than 7% than storing the original variables).

The predicted situation was contrasted with the actual future situation. The future situation was known, as long as historical data was used to develop the system and also to test the correction of it. The proposed solution was, in most of the variables, more than a 90% accurate when having at least 4000 different cases stored in the case base. For every problem defined by an area and its variables, the system offers nine solutions (i.e., the same area with its proposed variables and the eight closest neighbours). This way of prediction is used in order to clearly observe the direction of the forest fire which is useful in order to determine the areas that will be affected by the forest fire.

Table 3 shows a summary of the results obtained after comparing different techniques with the results obtained using OBIFS. The table shows the evolution of the results along with the increase of the number of cases stored in the case base. All the techniques analyzed improve its results while increasing the number of cases stored. Having more cases in the case base, makes easier to find similar cases to the proposed problem and then, the solution is more accurate. The “RBF” column represents a simple Radial Basis Function Network that is trained with all the data available. The network gives an output that is considered a solution to the problem. The “CBR” column represents a pure CBR system, with no other techniques included; the cases are stored in the case base and recovered considering the Euclidean distance. The most similar cases are selected and after applying a weighted mean depending on the similarity of the selected cases with the inserted problem, a solution is proposed. The “RBF + CBR” column corresponds to the possibility of using a RBF system combined with CBR.

Table 3
Percentage of good predictions obtained with different techniques.

Number of stored cases	RBF (%)	CBR (%)	RBF + CBR (%)	OBIFS (%)
100	45	39	42	46
500	48	43	46	55
1000	51	47	58	66
2000	56	55	65	74
3000	59	58	68	86
4000	60	63	69	89
5000	63	64	72	93

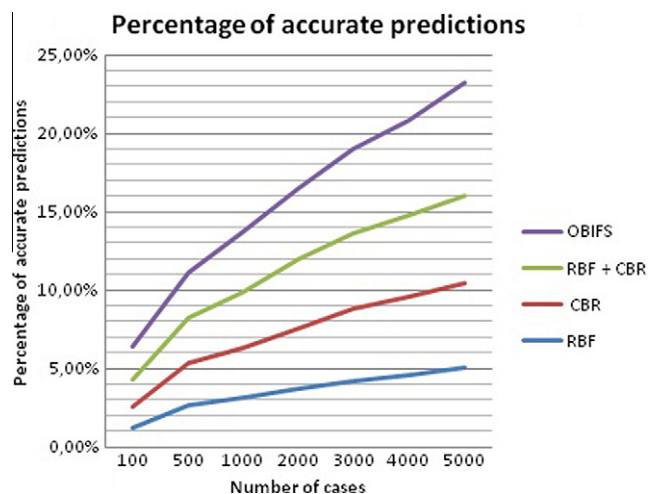


Fig. 12. Percentage of accurate predictions comparing OBIFS with the rest of the used techniques.

The recovery from the CBR is done by the Manhattan distance and the RBF network works in the reuse phase, adapting the selected cases to obtain the new solution.

The results of the “RBF + CBR” column are, normally, better than those of the “CBR”, mainly because of the elimination of useless data to generate the solution. Finally, the “OBIFS” column shows the results obtained by OBIFS, obtaining better results than the three previous analyzed solutions. Fig. 12 shows a graphical interpretation of this results, clearly indicating that the results are better when the amount of stored information is increased, and that the proposed system improves the results obtained by the rest of the compared techniques.

Table 4 shows the evolution of the results along with the increase in the number of cases stored in the case base. The numerical results showed in Table 4 represent the average of a series of tests completed with the available information. The number of tests performed in each iteration (every time the case base grows) represents ten percent of the size of the case base (if the case base contains 1000 elements, then 100 tests will be performed, and so on). When the amount of available information is increased by adding new data to the system (new satellite data, new direct observations), the results are validated again considering the new conditions offered by the new data. The more data available, the better results the systems usually generate. The elements that make part of the ten percent of the data used to test the correction of the system has not previously been used in its training. Those cases also come to form part of the historical information and may be used to check the correction of the predictions generated by the system. For every case stored in the case base there is a future situation corresponding to the solution of that situation. The cases used to test the system are randomly chosen from the overall amount of cases. The results for each of the techniques under analysis improved when the number of cases stored was increased.

Table 4
Improvement obtained using the different techniques after applying the buffer operator.

Number of cases	RBF (%)	CBR (%)	RBF + CBR (%)	OBIFS (%)
100	1.2	1.4	1.7	2.1
500	2.7	2.6	2.9	2.9
1000	3.1	3.2	3.6	3.8
2000	3.7	3.9	4.4	4.5
3000	4.2	4.6	4.8	5.4
4000	4.6	5.0	5.2	6.0
5000	5.1	5.3	5.6	7.2

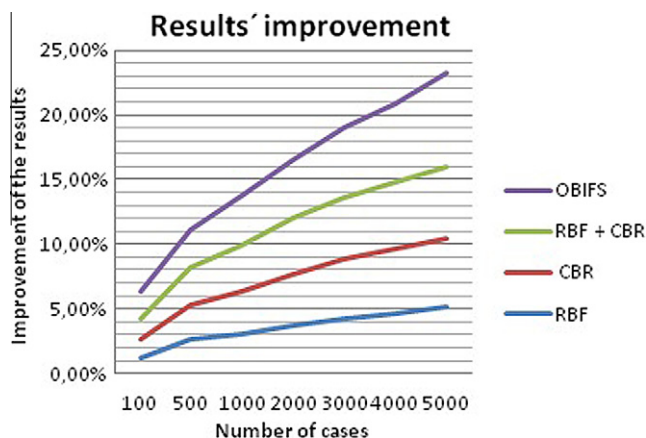


Fig. 13. Improvement in the results of OBIFS predictions compared with the rest of the used techniques.

Fig. 13 shows a graphical representation of the results in Table 4, clearly showing that the use of the buffer operator in the OBIFS leads to an improvement in the results. Table 4 shows the improvement obtained in the different systems previously explained after applying the buffer operator to them.

9. Conclusions

OBIFS is a new solution for predicting the future situation of the forest fires, analyzing different parameters. This system presents a distributed multi-agent architecture which allows the interaction of multiple users at the same time. Distributing resources also allows users to interact with the system in different ways depending on their specific needs for each situation (e.g. introducing data or requesting a prediction). This architecture becomes an improvement with previous tools where the information must be centralized and where local interfaces were used.

With the vision introduced by OBIFS, all the different people that may interact with a contingency response system collaborate in a distributed way, being physically located in different places but interchanging information in a collaborative mode. (Aamodt and Plaza, 1994; Althoff et al., 2005) OBIFS makes use of the CBR methodology to create new solutions and predictions using past solutions given to past problems. The CBR structure has been divided into services in order to optimize the overall performance of OBIFS. Generalization must be done in order to improve the system. Applying the methodology explained before to diverse geographical areas will make the results even better, being able to generate good solutions in more different situations. The current system has been mainly developed using data from the Gestosa experiments in central Portugal. With that information, OBIFS has been able to generate solutions to new situations, based on the available cases. If the amount and variety of cases stored in the case base is increased, the quality of the results will also be boosted. Although the performed tests have provided us very useful data, it is necessary to continue developing and enhancing OBIFS. The number of possible interfaces can be augmented, including independent sensors that may send information to the system in real-time. The data received by the system must be analyzed in order to detect new useful information and to generate fast and accurate solutions to existing problems without the direct intervention of the users. Then, the system will not only be a contingency response but also a kind of supervising system especially in dangerous geographical areas. The next steps in the development of the buffer operator applied to this organization-based multiagent system will be the application of the system to other

case studies and improvements to image analysis and the use of the buffer operator; thereby introducing novel techniques that may generate better results in this and other fields of knowledge.

Acknowledgements

This research is partially supported by projects TIN2010–21272-C02–01 from the Spanish Ministry of Science and Innovation and SA405A12–2 from Junta de Castilla y Leon. This work was also supported in the framework of the IT4 Innovations Centre of Excellence project, Reg. No. CZ.1.05/1.1.00/02.0070 by operational programme' Research and Development for Innovations' funded by the Structural Funds of the European Union and state budget of the Czech Republic, EU.

References

- Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1), 39–59.
- Alexandridis, A., Vakalis, D., Siettos, C., & Bafas, G. (2008). cellular automata model for forest fire spread prediction: The case of the wildfire that swept through spetses island in 1990. *Applied Mathematics and Computation*, 204(1), 191–201.
- Alonso-Betanzos, A., Guijarro-Berdiñas, B., Hernández-Pereira, E., Paz Andrade, M. I., Jiménez, J., Legido Soto, J. L., et al. (2003). An intelligent system for forest fire risk prediction and fire fighting management in Galicia. *Expert Systems with Applications*, 25(4), 545–554.
- Althoff, K. D., Manz, J., & Nick, M. (2005). Maintaining experience to learn: Case studies on case-based reasoning and experience factory. In *Proceedings on 6th Workshop Days of the German Computer Science Society (GI) on Learning, Knowledge, and Adaptivity (LWA 2005)*, Saarland University, Germany.
- Atsuyuki, O., Barry, B., Kokichi, S., et al. (1992). *Spatial tessellations concepts and applications of voronoi diagrams*. John Wiley & Sons.
- Baruque, E., Corchado, E., Mata, A., & Corchado, J. M. (2010). A forecasting solution to the oil spill problem based on a hybrid intelligent system. *Information Sciences*, 108(10), 2020–2043.
- Batschelet, E. (1981). *Circular Statistics in Biology* (Vol. 24). Academic Press.
- Boella, G., Governatori, G., Hulstijn, J., Riveret, R., Rotolo, A., & Van Der Torre, L. (2011). Time and defeasibility in fipa acl semantics. *Journal of Applied Logic*.
- Bretin, E., Lachaud, J., & Oudet, É. (2011). Regularization of discrete contour by willmore energy. *Journal of Mathematical Imaging and Vision*, 1–16.
- Carrascosa, C., Bajo, J., Julian, V., Corchado, J. M., & Botti, V. (2007). Hybrid multi-agent architecture as a real-time problem-solving model. *Expert Systems with Applications*, 34(1), 2–17.
- Cerami, E. (2002). *Web Services Essentials Distributed Applications with XML-RPC, SOAP, UDDI WSDL*, O'Reilly Associates, Inc.
- Cetin, E., & Yusuf, S. (2011). A data fusion framework with novel hybrid algorithm for multi-agent Decision Support System for Forest Fire. *Expert Systems with Applications*, 38(8), 9225–9236.
- Cheng, T., & Wang, J. (2008). Integrated spatio-temporal data mining for forest fire prediction. *Transactions in GIS*, 12(5), 591–611.
- Chou, Y. H. (1997). *Exploring spatial analysis in geographic information systems*. Onward Press.
- Corchado, J. M., Mata, A., Paz, F., & Pozo, D. (2008). Case-based reasoning systems to forecast the presence of oil slicks, IADIS 2008. *European Conference Data Mining:3–10*.
- Denham, M., Cortés, A., Margalef, T., & Luque, E. (2008). Applying a dynamic data driven genetic algorithm to improve forest fire spread prediction. *Computational Science – ICCS 2008*, 36–45.
- Dignum, V., & Dignum, F. (2007) A logic for agent organizations. *FAMAS@ Agents 3–7*.
- Downs, T. D., & Mardia, K. V. (2002). Circular regression. *Biometrika*, 89, 683–697.
- Dun, C. E., Bhopal, R. S., Cockings, S., Walker, D., Rowlingston, B., & Diggle, P. (2007). Advancing insights into methods for studying environment–health relationships: A multidisciplinary approach to understanding Legionnaires' disease. *Health & Place*, 13(3), 677–690.
- Emre Celebi, M., Alp Aslandogan, Y., Stoecker, W., Iyatomi, H., Oka, H., & Chen, X. (2007). Unsupervised border detection in dermoscopy images. *Skin Research and Technology*, 13(4), 454–462.
- Esri, (2012). The GIS software leader <<http://www.esri.com>>.
- Ferber, J., Gutknecht, O., & Michel, F. (2004). From agents to organizations: an organizational view of multi-agent systems. *Lecture Notes in Computer Science*, 214–230.
- Fritzke, B. (1994). Growing cell structure—a self-organizing network for unsupervised and supervised learning. *Neural Networks*, 7(9), 1441–1460.
- García, E., Argente, E., Giret, A., & Botti, V. (2008). Issues for organizational multiagent systems development. In *Sixth International Workshop From Agent Theory to Agent Implementation (AT2AI-6)*, Citeseer (pp. 59–65).
- Gasser, L. (2001). Perspectives on organizations in multi-agent systems. In M. LUCK, V. MARIK, O. STEPANKOVA, & R. TRAPPL (Eds.), *Multi-agent Systems and Applications. 9th ECCAI Advanced Course, EASSS 2001*. Springer.

- Gestosa (2000). Gestosa 2000 experimental fires in shrub vegetation in central Portugal, *International Forest Fire News* (23): 103–105.
- Gestosa (2005) <<http://www.adai.pt/ceif/Gestosa/>>. A DAI-CEIF (Centre of Forest Fire Studies).
- González, P. (2000). Una alternativa al cálculo de la zona de influencia 2d en sistemas de información geográfica, CEIG 2000: X Congreso Español de Informática Gráfica, Vol. 3, Universidad Jaume I Servicio De, 217.
- Gunter, S., Schraudolph, N. N., & Vishwanathan, S. V. N. (2007). Fast iterative kernel principal component analysis. *Journal of Machine Learning Research*, 8, 1893–1918.
- Herrero-Jiménez, C. M. (2012). An expert system for the identification of environmental impact based on a geographic information system. *Expert Systems with Applications*, 39(8), 6672–6682.
- Huang, H., & Li, Y. (2009). Load balancing spatial analysis in xml/gml/svg based webgis. In *Environmental Science and Information Application Technology. International Conference on ESIAT 2009* (Vol. 2, pp. 538–541). IEEE.
- Idri, A., Zakrani, A., & Zahi, A. (2010). Design of radial basis function neural networks for software effort estimation. *International Journal of Computer Science*, 7(4), N° 3.
- Iliadis, L. S. (2005). A decision support system applying an integrated fuzzy model for long-term forest fire risk estimation. *Environmental Modelling and Software*, 20(5), 613–621.
- Jammalamadaka, S., & Sengupta, A. (2001). *Topics in circular statistics* (Vol. 5). World Scientific Pub Co Inc.
- Josuttis, N. (2007). *SOA in Practice*. O'reilly.
- Karayannis, N. B., & Mi, G. W. (1997). Growing radial basis neural networks: merging supervised and unsupervised learning with network growth techniques. *IEEE Transactions on Neural Networks*, 8, 1492–1506.
- Kolodner, J. L. (1993). *Case-based Reasoning*. Morgan Kaufman.
- Ladrón de Guevara, I., Muñoz, J., de Cózar, O., & Blázquez, E. (2011). fitting of circle arcs. *Journal of Mathematical Imaging and Vision*, 1–15.
- Liu, S. (1992). Buffer operator and its application. *Theories and Practices of Grey System*, 2(1), 45–50.
- Long, D. G. (2001). Mapping fire regimes across time and space: Understanding coarse and fine-scale fire patterns. *International Journal of Wildland Fire*, 10, 329–342.
- Longley, P. (2005). *Geographic information systems and science*. John Wiley & Sons Inc.
- Mardia, K., & Jupp, P. (2000). *Directional statistics*. John Wiley & Sons Inc.
- Martínez-de Dios, J. R., Arrue, B. C., Ollero, A., Merino, L., & Gómez-Rodríguez, F. (2008). Computer Vision Techniques for forest fire perception. *Imagen and Vision Computing*, 26(4), 550–562.
- Mazzeo, G., Marchese, F., Filizzola, C., Pergola, N., & Tramutoli, V. (2007). A multi-temporal robust satellite technique (rst) for forest fire detection. *Analysis of Multi-temporal Remote Sensing Images*, 1–6.
- Mikic Fonte, F. A., Burguillo, J. C., & Llamas istal, M. (2012). An intelligent tutoring module controlled by BDI agents for an e-learning platform. *Expert Systems with Applications*, 39(8), 7546–7554.
- Nishino, K. (2009). Directional statistics brdf model. In *Computer Vision. 2009 IEEE 12th International Conference on IEEE* (pp. 476–483).
- Nuxoll, A. & Laird, J. (2011). Enhancing intelligent agents with episodic memory. *Cognitive Systems Research*.
- Pausas, J., Llovet, J., Rodrigo, A., & Vallejo, R. (2009). Are wildfires a disaster in the mediterranean basin?—a review. *International Journal of Wildland Fire*, 17(6), 713–723.
- Recio-García, J., Casado-Hernández, M., & Díaz-Agudo, B. (2010). Extending cbr with multiple knowledge sources from web. *Case-Based Reasoning. Research and Development*, 287–301.
- Ros, F., Pintore, M., & Chrétien, J. R. (2007). Automatic design of growing radial basis function neural networks based on neighborhood concepts. *Chemometrics and Intelligent Laboratory Systems*, 87(2), 231–240.
- Schou, G. (1978). Estimation of the concentration parameter in von mises-fisher distributions. *Biometrika*, 65(2), 369.
- Serón, F. J., Gutiérrez, D., Magallón, J., Ferragut, L., & Asensio, M. I. (2005). The evolution of a wildland forest fire front. *The Visual Computer*, 21(3), 152–169.
- Singh, H., Hnizdo, V., & Demchuk, E. (2002). Probabilistic model for two dependent circular variables. *Biometrika*, 89(3), 719–723.
- Sun, F., Choi, Y. K., Wang, W., Yang, D. M., Liu, Y., et al. (2011). Obtuse triangle suppression in anisotropic meshes. *Computer Aided Geometric Design*, 28(9), 537–548.
- Taylor, C. C. (2008). Automatic bandwidth selection for circular density estimation. *Computational Statistics and Data Analysis*, 52(7), 3493–3500.
- Tung, Y. H., Tseng, S. S., Weng, J. F., Lee, T. P., Liao, A. Y. H., & Tsai, W. N. (2010). A rule-based CBR approach for expert finding and problem diagnosis. *Experts Systems with Applications*, 37(3), 2427–2438.
- Ye, D., Zhang, M., & Sutanto, D. (2012). Self-organization in agent network: a mechanism and a potential application. *Decision Support Systems*, 53(3), 406–417.
- Zar, J. (1999). *Biostatistical analysis. Vol. 564*. Prentice Hall, Upper Saddle River.