

## **A GEOMETRIC APPROACH TO CANCER GROWTH PREDICTION BASED ON COX PROCESSES**

**IULIAN T. VLAD and JORGE MATEU**

Department of Mathematics  
University Jaume I  
Campus Riu Sec, E-12071  
Castellón  
Spain  
e-mail: vlad@uji.es

### **Abstract**

Lévy theory provides a potential mathematical framework to model space and space-time stochastic processes. In addition, spatial point processes define stochastic models for random patterns of points in  $\mathbb{R}^2$ . These processes play a special role in stochastic geometry as the building blocks of more complicated random set models. In this paper, we focus on a family of Lévy-based spatial Cox processes to model and predict tumor growth. We develop a procedure to simulate the growing of tumors. This algorithm can be used to study the evolution in time of any 2- and 3-dimensional geometrical forms such as cancer skin and all type of boundary evolution. We analyze real data and show that the procedure developed works fine and is useful for prediction purposes.

### **1. Introduction**

The evolution in time of some objects is the subject of study of many researchers worldwide. Special attention has been given to cancer, and a way to understand this disease is to know how it evolves over time.

---

2010 Mathematics Subject Classification: 60G55, 62M30, 62M40.

Keywords and phrases: cobweb algorithm, Cox processes, space-time modelling, tumor growth.

Received December 11, 2014

We can find, in reading, many studies about mathematical modelling of tumors; see, for example, Bramson and Griffeath [2]; Cressie [5]; Qi et al. [15]; Lee and Cowan [12]; Kansal et al. [11]; Barndorff-Nielsen and Schmiegel [1]; Jónsdóttir and Vedel Jensen [9]; and Jónsdóttir et al. [10].

In Richardson [16], the growth object in the plane at time  $t$  is a random subset  $Y_t$  of  $\mathbb{Z}^2$  consisting of the “infected sites”, described also by a Markov process. An uninfected site is transferred to an infected site with a rate proportional to the number of infected nearest neighbours. It can be shown that if  $Y_0$  consists of a single site, then  $Y_t/t$  has a non-random shape as  $t \rightarrow \infty$ .

Bramson and Griffeath [2] denote the set of sites occupied by cancer cells  $\xi_t^A$ , at time  $t$  and given that the original cancerous population ( $\xi_t^0$ ) occupies  $A \in S_0$ , and the processes  $(\xi_t^\alpha)_{t \geq 0}$  are Markov, they define for  $\lambda > 0$ , a jump (growth) rate as

$$A \rightarrow A \cup \{x\} (x \notin A) \text{ at rate } \lambda |\{y \in A : \|y - x\| = 1\}|,$$

$$A \rightarrow A - \{x\} (x \in A) \text{ at rate } 1,$$

where  $|\Lambda|$  is the cardinality of  $\Lambda \in S_0$  and  $\|x\|$  is the Euclidean distance from  $x$  to 0. This study is a extension of the stochastic model defined by Williams and Bjerknes [17] to accommodate expansion of cancer cells. That model treats each cell as normal and abnormal (cancerous) independently and assume that both are located in a planar lattice  $\mathbb{Z}^2$ . Starting with a single abnormal cell at the origin and from the hypothesis that abnormal cells reproduce faster than normal cells, they assume that with each cellular division, one daughter cell stays fixed while the other usurps the position of a neighbour.

A related growth model in continuous space has been discussed in Deijfen [7]. For planar objects, the model is constructed from a spatiotemporal Poisson point process on  $\mathbb{R}^3$ ,

$$\Psi = \{(x_i, t_i)\}.$$

The random growing object  $Y_y \subset \mathbb{R}^2$  is a subset of

$$\bigcup_{\{i:t_i \leq t\}} B(x_i, r),$$

constructed such that  $Y_t$  is always connected. Here,  $B(x, r)$  is a circular disc with center  $x$  and radius  $r$ . In this model,  $t_i$  is thought of as a time point of outburst and  $x_i$  as the location of the outburst in the tumor. A closely related discrete-time Markov growth model has been proposed by Cressie [5]. This model can be characterized as a sequence of Boolean models

$$Y_{t+1} = \bigcup \{B(x_i, r) : x_i \in Y_t\},$$

where  $\{x_i\}$  is a homogeneous Poisson point process in  $\mathbb{R}^2$ ; see Cressie and Laslett [4]; Cressie [5]; and Cressie [6].

In all these papers, it is obvious that the form of the cancer in the future depends by the edge and structure of the cancer in the past (function  $f$ ) and also by some external factors like mitosis, nature of cancer (benign or malign), biological tissue density, etc (all these factors can be included in a function  $g$ ). So, a *very general growth model* can be expressed as

$$Y_{t+1} = f(Y_t) + g. \tag{1}$$

For not complicating the model, we should refer to the cancer tumor just to their external edge, not to their internal structure. Then we can tell that the object of study can be a shape with a particular boundary of whose genesis is a single point. Let us study this as a point process and let us make some geometrical interpretations to calculate the rate growth. For this, it is not strictly necessary to have deterministic expressions of the functions  $f$  and  $g$  and make a complex model, which aims to offer the values of the boundary object at any instant time. We can calculate these values with an algorithm by observing the speed of expansion of the tumor, expressed as a constant velocity of vectors in some directions.

The aim of this research is to observe the dynamics of cancer tumors and to develop and implement new methods and algorithms for prediction of tumor growth. We offer some tools to help physicians for a better understanding and treatment of this disease. Using a prediction method and comparing with the real evolution, a physician can note if the prescribed treatment has the desired effect, and according to this, if necessary, to take the decision of surgically intervention.

The plan of the paper is the following. Section 2 presents a short overview of the theory of Lévy bases and integration with respect to such bases is given, we recall standard results about spatial Cox processes, and finally, we propose different types of growth models. In Section 3, a new algorithm, the cobweb, is presented and developed based on the proposed methodology. The implementation in Matlab software comes in Section 4. Section 5 presents some real data analysis. The paper ends with some conclusion.

## 2. Spatial Cox Point Processes

A spatial point pattern is a set of points  $\{x_i \in A : i = 1, \dots, n\}$  for some planar region  $A$ . Very often,  $A$  is a sampling window within a much larger region and it is reasonable to regard the point pattern as a partial realization of a stochastic planar point process, the events consisting of all points of the process which lie within  $A$ .

Let  $N$  be a spatial point process that is defined on  $\mathbb{R}^2$  but is observed on a finite observation window  $W$ . For an arbitrary Borel set  $A \in \mathbb{R}^2$ , let  $|A|$  and  $N(A)$  denote the area of  $A$  and the number of events from  $N$  that are in  $A$ , respectively. In some applications, it is reasonable to think of the spatially varying intensity function,  $\lambda(x)$  of a point process to be itself a realization of an underlying stochastic process  $\Lambda(x)$ .

A point process  $X$  is a *Cox process* if: (a)  $\Lambda(x)$  is a non-negative valued stochastic process; and (b) conditional on the realization of  $\Lambda(x)$ , the point process is an inhomogeneous Poisson process with intensity

function  $\Lambda(x)$  (Cox [3]). In this case, we say that  $X$  is a Cox process driven by  $\Lambda$ . In this context, the resulting point process inherits the properties of the  $\Lambda(x)$  process in a natural way.

Cox processes provide useful and frequently applied models for aggregated spatial point patterns, where the aggregation is due to a stochastic heterogeneity. Indeed,  $\Lambda$  usually models this unobserved random heterogeneity. Shot noise Cox processes, log Gaussian Cox processes and log shot noise Cox processes will appear as natural building blocks in a modelling framework for Cox processes.

### 2.1. Lévy-based Cox processes

Let  $(\Omega, \mathcal{A})$  be a measurable space. We assume that  $\Omega$  is a Borel subset of  $\mathbb{R}^d$ , and  $\mathcal{A}$  is the  $\delta$ -ring  $\mathcal{B}_b(\Omega)$  of bounded Borel subsets of  $\Omega$ . We consider a collection of real-valued random variables  $L = \{L(A), A \in \mathcal{A}\}$  with the following properties:

- $L(A_1), \dots, L(A_n), \dots$  are independent random variables for every sequence  $\{A_n\}$  of disjoint sets in  $\mathcal{A}$ , and  $L(\cup_n A_n) = \sum_n L(A_n)$  a.s. provided  $\cup_n A_n \in \mathcal{B}_b(\Omega)$ .
- For every  $A$  in  $\mathcal{A}$ ,  $L(A)$  is infinitely divisible.

If  $L$  has these properties,  $L$  is called a Lévy basis. In addition,  $L$  is a non-negative Lévy basis if  $L(A) \geq 0$  for all  $A \in \mathcal{A}$  (cf. Barndorff-Nielsen and Schmiegel [1]).

For a random variable  $X$ , Jónsdóttir et al. [10] denote the cumulant function  $\log \mathbb{E}(e^{i\nu X})$  by  $C(\nu, X)$ . If  $L$  is a Lévy basis, then the cumulant function of  $L(A)$  is expressed as

$$C(\nu, L(A)) = i\nu a(A) - \frac{1}{2} \nu^2 b(A) + \int_{\mathbb{R}} (e^{i\nu r} - 1 - i\nu r 1_{[-1,1]}(r)) U(dr, A), \quad (2)$$

where  $a$  is a  $\sigma$ -additive set function on  $\mathcal{A}$ ,  $b$  is a measure on  $\mathcal{A}$ ,  $U(dr, A)$  is a measure on  $\mathcal{A}$  for fixed  $dr$  and a Lévy measure on  $\mathcal{B}(\mathbb{R})$  for each fixed  $A \in \mathcal{A}$  (i.e.,  $U(\{0\}, A) = 0$  and  $\int_{\mathbb{R}} (1 \wedge r^2) U(dr, A) < \infty$ , where  $\wedge$  denotes minimum).

The measure  $U$  is referred to as the *generalised Lévy measure* and  $L$  is said to have the *characteristic triplet*  $(a, b, U)$ . In addition, (a) if  $b = 0$ , then  $L$  is called a *Lévy jump basis*, and (b) if  $U = 0$ , then  $L$  is a *Gaussian basis*. A general Lévy basis  $L$  can always be written as a sum of a Gaussian basis and an independent Lévy jump basis.

A particular example of a Gaussian Lévy basis is obtained by attaching independent Gaussian random variables  $\{X_i\}$  to a locally finite sequence  $\{\eta_i\}$  of fixed points and defining

$$L(A) = \sum_{\eta_i \in A} X_i, \quad A \in \mathcal{A}.$$

Let  $\mathcal{S}$  be a Borel subset of  $\mathbb{R}^d$ . A point process  $X$  on  $\mathcal{S}$  is called a *Lévy driven Cox process* (LCP), if  $X$  is a Cox process with a driving field

$$\Lambda(\xi) = \int_{\Omega} k(\xi, \eta) L(d\eta), \quad \xi \in \mathcal{S}, \quad (3)$$

where  $L$  is a non-negative Lévy basis on  $\Omega$ .

Furthermore,  $k$  is a non-negative function on  $\mathcal{S} \times \Omega$  such that  $k(\xi, \cdot)$  is integrable with respect to  $L$  for each  $\xi \in \mathcal{S}$  and  $k(\cdot, \eta)$  is integrable with respect to the Lebesgue measure on  $\mathcal{S}$  for each  $\eta \in \Omega$ .

Note that it is always possible for each pair  $(k, L)$  to construct an associated pair  $(\bar{k}, \bar{L})$  generating the same driving field  $\Lambda$ , where now  $\bar{k}(\cdot, \eta)$  is a probability kernel. We may simply let

$$\bar{k}(\xi, \eta) = k(\xi, \eta) / \alpha(\eta),$$

$$\bar{L}(d\eta) = \alpha(\eta)L(d\eta),$$

where

$$\alpha(\eta) = \int_{\mathcal{S}} k(\xi, \eta)d\xi.$$

It is important to note that from the non-negativity of the Lévy basis  $L$ , we get that  $L$  is equivalent to a random measure on  $\Omega$ . Thus, the measurability of  $\Lambda$  follows from the measurability of  $k$  as a function of  $\eta$  and  $\xi$ . Therefore,  $\Lambda$  is a well-defined random field and (under the condition of local integrability) the driving measure  $\int_B \Lambda(\xi)d\xi$ ,  $B \in \mathcal{B}_b(\mathcal{S})$ , is also a well-defined random measure determined by the finite-dimensional distributions of  $L$ .

The function  $k$  and the Lévy basis  $L$  will be chosen such that  $\Lambda$  is almost surely locally integrable, i.e.,  $\int_B \Lambda(\xi)d\xi < \infty$  with probability 1 for  $B \in \mathcal{B}_b(\mathcal{S})$ . A sufficient condition for the last property is that, cf. (Møller [13], Remark 5.1)

$$\int_B \mathbb{E}\Lambda(\xi)d\xi < \infty, \quad B \in \mathcal{B}_b(\mathcal{S}) \tag{4}$$

## 2.2. Lévy-based tumor growth modelling

Let us denote the growing object as a planar object at time  $t$  by  $Y_t \in \mathbb{R}^2$ , and assume that  $Y_t$  is compact and *star-shaped* with respect to a point  $O(x_0, y_0) \in Y_t$  for all  $t$ . We treat here a star-shaped object like a two-dimensional geometric shape and the growth model like a rigid transformation in time of the primary star-shaped object, defining a second star-shaped object that includes the boundary of the initial object.

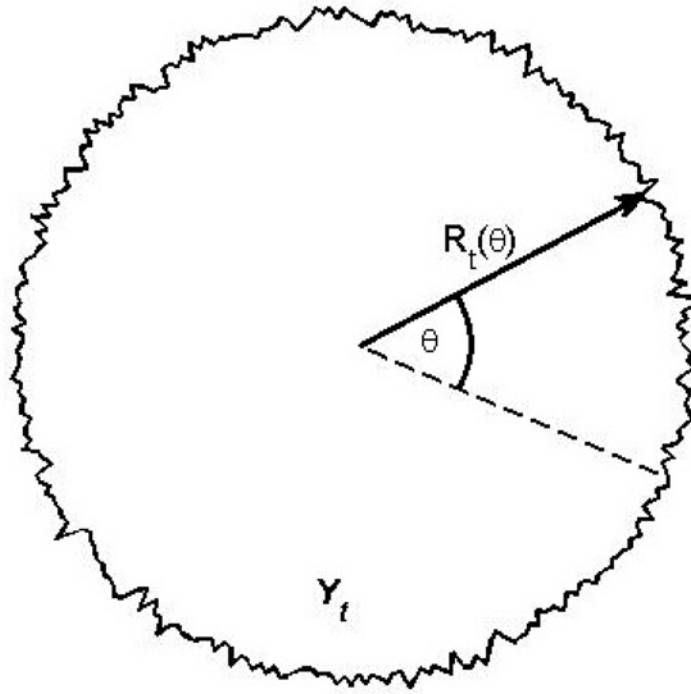
In geometry, two subsets of a Euclidean space have the same shape if one can be transformed into the other by a combination of translations, rotations (together also called *rigid transformations*), and uniform

scalings. Note that we talk about star-shaped that can grow its boundary in a random way. So, we can not say that the transformation of the initial star-shape object in time is a rigid transformation.

The boundary of  $Y_t$  can be determined from the variation in time  $t$  and direction (from angle)  $\phi \in [0, 2\pi)$  of the vector  $\vec{R}_t(\phi)$  denoted by the radial function,

$$R_t(\phi) = \max\{r : (x_0, y_0) + r(\cos \phi, \sin \phi) \in Y_t\}; \quad \phi \in [0, 2\pi). \quad (5)$$

In Figure 1, we show an example of such star-shape object noted with  $Y_t$  and  $R_t(\phi)$  is the distance from a reference point  $O(x_0, y_0)$  to the boundary of the object.



**Figure 1.** The star-shape object  $Y_t$ .



Jónsdóttir et al. [10] consider this as a random variable  $X_t(\sigma)$  depending on time  $t$  and position  $\sigma$  in space. They assume that  $(\sigma, t) \in \Omega = \mathcal{S} \times \mathbb{R}$ , where  $\mathcal{S} \subseteq \mathbb{R}^n$ . A Lévy-based spatio-temporal model for  $X = \{X_t(\sigma) : (\sigma, t) \in \Omega\}$  is based on the ambit set  $A_t(\sigma)$  associated with each point  $(\sigma, t) \in \Omega$ , which defines the dependency on the past at time  $t$  and position  $\sigma$ , and satisfies the conditions

$$(\sigma, t) \in A_t(\sigma),$$

$$A_t(\sigma) \subseteq \mathbb{S} \times (-\infty, t].$$

The *linear spatio-temporal Lévy model* for  $X = \{X_t(\sigma) : (\sigma, t) \in \Omega\}$  is defined as

$$X_t(\sigma) = \int_{A_t(\sigma)} f_t(\xi, \sigma) L(d\xi), \quad (6)$$

where  $L$  is a Lévy basis and  $f_t(\xi, \sigma)$  is the deterministic weight function.

The process

$$\bar{X} = \{\exp(X_t(\sigma)) : (\sigma, t) \in \Omega\},$$

is said to follow an exponential spatio-temporal Lévy model.

So, the *model who describes the growth of a planar star-shaped object*, using its radial function  $R_t(\phi)$  at time  $t$  and angle  $\phi$ , can start from the time derivative of the radial function equation

$$\frac{\partial}{\partial t} R_t(\phi) = \mu_t(\phi) + \int_{A_t(\sigma)} f_t(\xi, \phi) L(d\xi); \quad \phi \in [0, 2\pi), \quad (7)$$

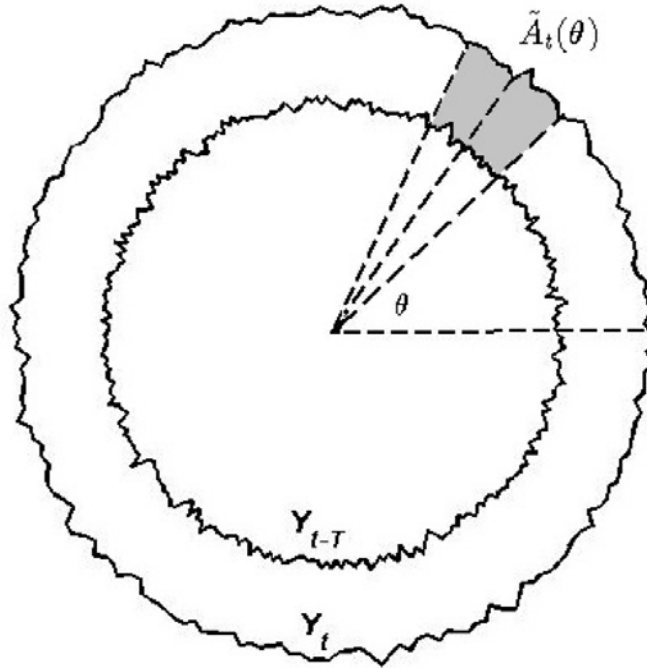
which is the growth rate (cf. Jónsdóttir et al. [10]). Here,  $L$  is the Lévy-basis on  $[0, 2\pi) \times \mathbb{R}$ ;  $A_t(\phi) \subseteq [0, 2\pi) \times (-\infty, t]$  is a subset of the past of time  $t$ , called *ambit set* (Barndorff-Nielsen and Schmiegel [1]);  $f_t(\cdot, \phi) : [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}$  is a deterministic weight function (assumed to be suitable for the integral to exist) and the deterministic function  $\mu_t : [0, 2\pi) \rightarrow \mathbb{R}$  contributes to the overall growth pattern while the stochastic integral determines the dependence structure in the growth process.

The ambit set  $A_t(\phi)$  plays an important role in this modelling approach and affects the degree of dependence on the past. The extent of the dependence on the past may be specified by the minimal time-lag  $T(t)$  such that

$$A_t(\phi) \subseteq [0, 2\pi) \times [t - T(t), t]; \quad \phi \in [0, 2\pi).$$

The form of the ambit set  $A_t(\phi)$  will depend on the specific growth process being modelled. For the interpretation of (7) as a growth model, Jónsdóttir et al. [10] represent the ambit set as a stochastic subset of the growing object. This is possible if the stochastic time transformation  $t \rightarrow R_t(\phi)$  is non-decreasing for each  $\phi \in [0, 2\pi)$ . They represent the ambit set  $A_t(\phi)$  as a subset of  $Y_t$ ,

$$\bar{A}_t(\phi) = \{(R_s(\theta) \cos \theta, R_s(\theta) \sin \theta) : (\theta, s) \in A_t(\phi)\}. \quad (8)$$



**Figure 2.** Stochastic representation of  $\bar{A}_t(\phi)$ .

It follows from the fact that  $A_t(\phi) \subseteq [0, 2\pi) \times (-\infty, t]$  that  $\bar{A}_t(\phi)$  (Figure 2) is actually a subset of  $Y_t$ . Furthermore, since  $(\phi, t) \in A_t(\phi)$ , the set  $\bar{A}_t(\phi)$  touches the boundary of  $Y_t$  at the point  $(R_t(\phi) \cos \phi, R_t(\phi) \sin \phi)$ . It is the “events” in  $\bar{A}_t(\phi)$  that influence the growth rate at time  $t$  in direction  $\phi$ .

In the particular case where  $L$  is a *Poisson basis* and  $\Psi$  is the *associated Poisson point process* on  $[0, 2\pi) \times \mathbb{R}$ , then *the growth model* can be written as

$$\frac{\partial}{\partial t} R_t(\phi) = \mu_t(\phi) + \sum_{\bar{\xi} \in \Psi_t \cap \bar{A}_t(\phi)} \bar{f}_t(\bar{\xi}; \phi), \quad (9)$$

where the parameter  $\bar{\Psi}_t$  of the sum is a subset of  $Y_t$

$$\bar{\Psi}_t = \{(R_{t_i}(\theta_i) \cos \theta_i, R_{t_i}(\theta_i) \sin \theta_i) : t_i \leq t\}. \quad (10)$$

Finally, if

$$\bar{f}_t((s \cos \theta, s \sin \theta); \phi) = f_t((\theta, s); \phi), \quad (11)$$

and according to (9), the growth rate at time  $t$  in the direction  $\phi$  depends on the outbursts at time points before  $t$ , which lie in the stochastic neighbourhood  $\bar{A}_t(\phi)$ . Under (7), the induced model for  $R_t(\phi)$  will be (cf. Jónsdóttir et al. [10]) of the same linear form, since

$$\begin{aligned} R_t(\phi) &= R_0(\phi) + \bar{\mu}_t(\phi) + \int_0^t \int_{A_s} f_s(\xi; \phi) L(d\xi) ds \\ &= R_0(\phi) + \bar{\mu}_t(\phi) + \int_{\bar{A}_t} \bar{f}_t(\xi; \phi) L(d\xi), \end{aligned} \quad (12)$$

where  $R_0$  is the radial function at time  $t = 0$ , and

$$\bar{\mu}_t(\phi) = \int_0^t \mu_s(\phi) ds, \quad (13)$$

$$\bar{A}_t(\phi) = \bigcup_{0 \leq s \leq t} A_s(\phi), \quad (14)$$

$$\bar{f}_t(\xi; \phi) = \int_0^t \mathbf{1}_{A_s(\phi)}(\xi) f_s(\xi; \phi) ds. \quad (15)$$

Note that the ambit sets associated with the radial function itself are increasing, that is,

$$t_1 \leq t_2 \Rightarrow \bar{A}_{t_1}(\phi) \subseteq \bar{A}_{t_2}(\phi).$$

Another model proposed by Jónsdóttir et al. [10] is expressed in terms of the time derivative of  $\ln(R_t(\phi))$

$$\frac{\partial}{\partial t} (\ln(R_t(\phi))) = \mu_t(\phi) + \int_{A_t(\phi)} f_t(\xi; \phi) L(d\xi), \quad (16)$$

and the induced model is an exponential spatio-temporal Lévy model

$$R_t(\phi) = R_0(\phi) \exp\left(\bar{\mu}_t(\phi) + \int_{\bar{A}_t(\phi)} \bar{f}_t(\xi; \phi) L(d\xi)\right). \quad (17)$$

The choices of Lévy basis  $L$ , ambit sets  $A_t(\phi)$ , weight functions  $f_t(\xi; \phi)$  and  $\mu_t(\phi)$  completely determine the growth dynamics. These four ingredients can be chosen arbitrarily and independently, which results in a great variety of different growth dynamics.

Finally, Jónsdóttir and Vedel Jensen [9] propose a *Gaussian radial model for star-shaped objects*. The object at time  $t + 1$  is a stochastic transformation of the object at time  $t$  such that the radius vector function of the object fulfils

$$R_{t+1}(\phi) = R_t(\phi) + Z_t(\phi), \quad \phi \in [0, 2\pi), \quad (18)$$

where  $Z_t$  is a cyclical Gaussian process

$$Z_t(\phi) = \mu_t + \sum_{k=1}^{\infty} [A_{t,k} \cos(k\phi) + B_{t,k} \sin(k\phi)],$$

and assumed that the coefficients  $\mu_t$ ,  $A_{t,k}$ , and  $B_{t,k}$  are the coefficients of the Fourier series of  $Z_t$ , which has an important geometric interpretations relating to the growth process.

In the next section, we offer a geometric interpretation of the growth tumor determined by the radius  $R_t$  and angle  $\phi$  values at different instants of time  $t$  and we propose new algorithm to calculate the growing rate.

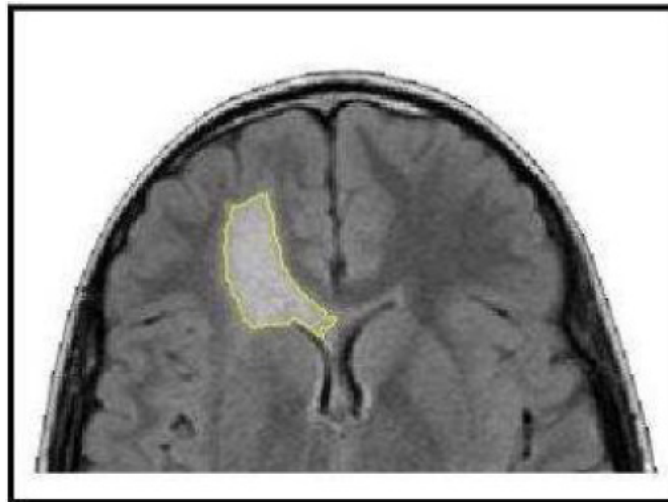
### 3. Modelling Tumor Growth: A New Algorithm

Principal questions for this subject are how fast the tumor grows, how rapidly does it invade and replace brain tissue, and what is the life expectancy of the patient? In general, tumor growth depends on the rate of mitosis (birth of new cells) and the rate of apoptosis (cell death). A tumor in which the rate of mitosis is equal to the rate of apoptosis does not appear to grow - it stays the same size as new tumor cells and simply replaces cells which die and the number of tumor cells stays the same.

To make a prediction about tumor growth we first need data (at least two images) at a predetermined interval time, to see if the tumor is growing and what is the relative velocity. Once we have located the tumor, the next step is to make a second tomography and compare the boundary of the tumor. In Figure 3, we show a tomography of a tumor and its location within the brain.



(a)



(b)

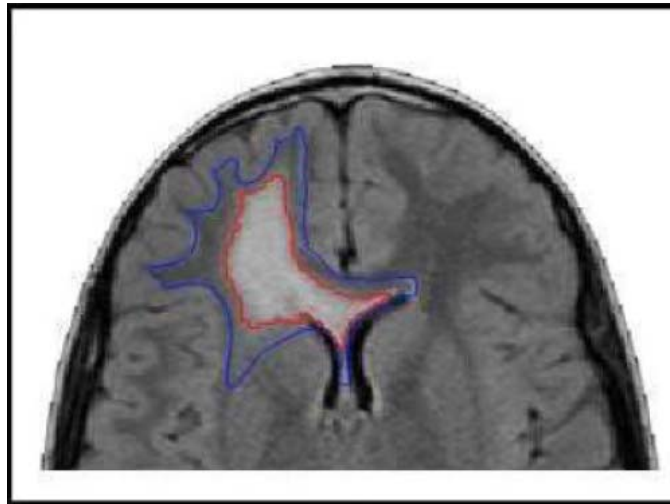
**Figure 3.** Brain tumor (a) and its location (b).

The tomography should be done in the same conditions as the previous image. The position of the patient must be identical so that the image is collected with the same cartesian coordinates as the previous image. The time period between two consecutive analysis is set as a parameter resolution  $k$ . Prediction of tumor growth directly depends on a

good image acquisition. In Figure 4, we have the second image after one month from when the disease was discovered. With the red line the tumor is marked and in the right-hand picture, we consider a sample space like a “star-shape” delimited by the blue line.



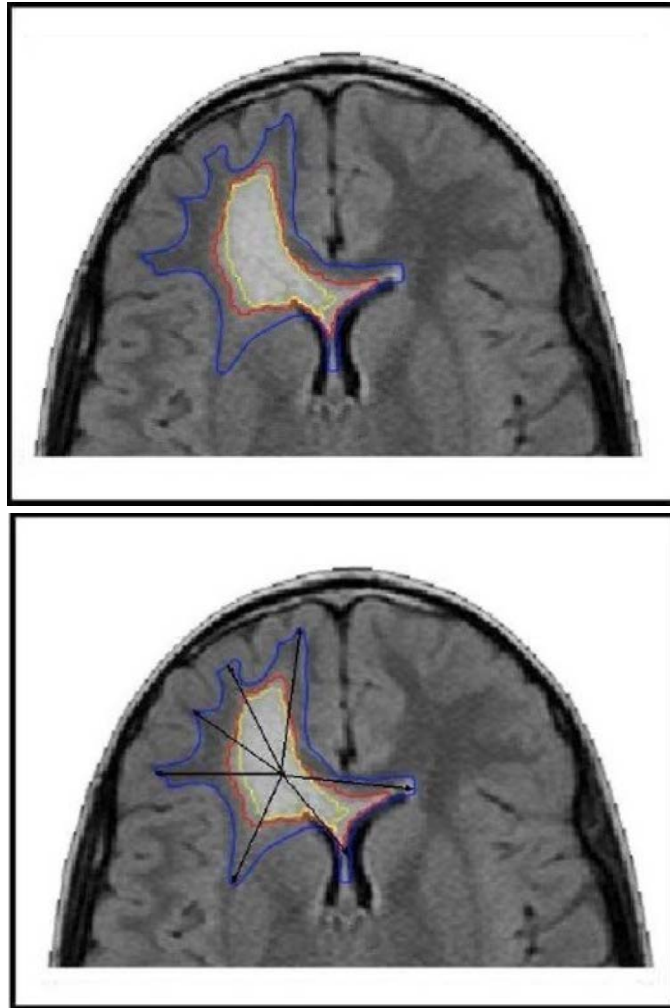
(a)



(b)

**Figure 4.** Second image acquisition after one month: (a) original image acquisition; (b) boundary of tumor (red) within the sample space (blue).

In Figure 5, we note the discovered tumor at time  $t$  with the yellow boundary, the growth tumor at time  $t + \Delta T$  ( $\Delta T = 30$  days) with red boundary, and the star-shape representing the sample space with a blue line. In the right-hand side, we show in black lines the vectors from the center of tumor to the limits of the sample space. The intersections of these vectors with each boundary give the value of the vector at time  $t$ , respectively,  $t + \Delta T$ .  $\Delta T$  represents the time period between analysis (here one month).

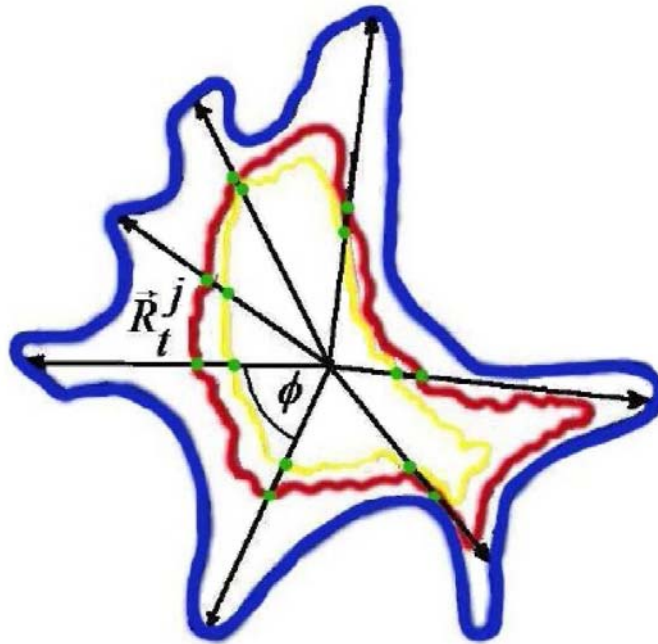


**Figure 5.** Superimposed images of tumor.



The sample space must be as large as possible, in the worst case, it could be the entire cranial box. But we can also assume that in a period of time, the tumor cannot grow over certain limits which depend on the structure of the brain.

We can easily construct a sample space which includes the boundary of the future tumor. The form of this sample space is directly influenced by the shape and positioning of the brain bulbs. It is a fact that the tumors grow more easily in some directions, which depend on the density and the nature of nearby biological material. Now, we can compare the images to see the evolution in time of the tumor. As we can see the tumor tends to grow more in certain directions which can be defined with vectors of growth of the tumor.



**Figure 6.** Tumor growth in star-shape.

In Figure 6, we graphically explain how we use the vectors  $\bar{R}_t^j$  to predict the growth of the tumor. The direction of each vector  $\bar{R}_t^j$ ,  $j = 1, \dots, n$  is given by the line from the center  $O(x_0, y_0)$  to each  $n$  inflexion of the star-shape (in this case, we consider only  $n = 7$ ). We denote the starting point time when we make the first tomography with  $t_0$  as the initial time for our computation.

The area of the initial tumor  $A_{t_0}$  for a time  $t_0$  is given by the boundary of the  $Y_{t_0}$  (yellow line in Figure 6), and the area of the tumor  $A_{t_0+\Delta T}$  at time  $t_0 + \Delta T$  (after one month) is given by the boundary of  $Y_{t_0+\Delta T}$  (red line in Figure 6).

Suppose that the velocity of growth is constant in time, then the value of the vector  $\bar{R}_{t_i}^j$  at the moment  $t_i = t_0 + \alpha\Delta T$ , represented by the radius  $r_i^j$ , is given by

$$r_i^j = \|\bar{R}_{t_0+\alpha\Delta T}^j\| = \|\bar{R}_{t_{i-1}+\Delta T}^j\| + k^{-1}(\|\bar{R}_{t_0+\Delta T}^j\| - \|\bar{R}_{t_0}^j\|), \quad (19)$$

for  $i = 1, \dots, k$  and  $j = 1, \dots, n$ .

The parameter  $\alpha$  is the period of time where upon we wish to make the estimation and this is given by the linear resolution

$$k = \frac{\Delta T}{\alpha}. \quad (20)$$

The angular resolution represents the circle divided by the number of vectors, which in cylindrical coordinates, means

$$n = \frac{2\pi}{\phi}. \quad (21)$$

In Equation (19), the final term of the right-hand side, is a constant defining the “length step”

$$l_k = k^{-1}(\|\bar{R}_{t_0+\Delta T}^j\| - \|\bar{R}_{t_0}^j\|),$$

and if we write  $\|\bar{R}_{t_{i-1}+\Delta T}^j\| = r_{i-1}^j$ , then (19) becomes

$$r_i^j = r_{i-1}^j + l_k. \quad (22)$$

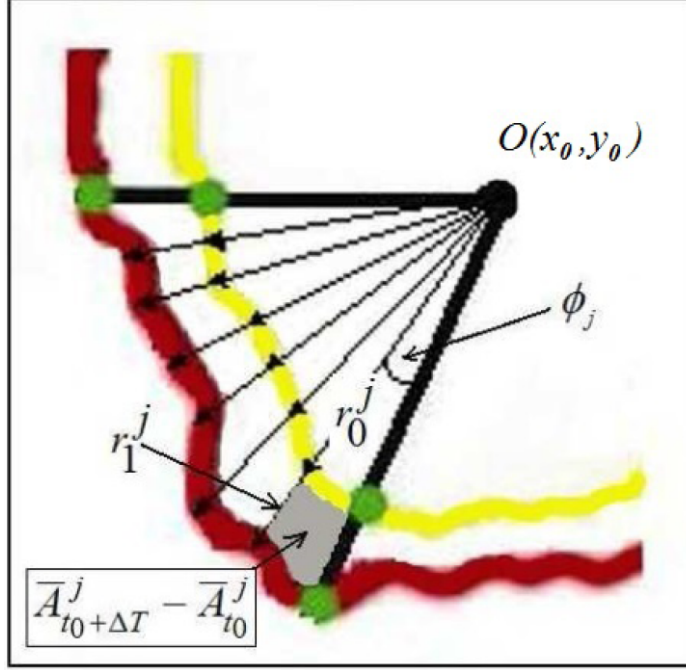
We use the cylindrical coordinates to calculate the area of  $\bar{A}_t(\phi)$  (see Figure 2), which help to calculate the predicted area  $\bar{A}_{t+\Delta T}(\phi)$  of the tumor after a period of time. To calculate this area, we make a discretization of the area region  $\bar{A}_t(\phi)$ , in  $n$  angles  $\phi_j$ ,  $j = 1, \dots, n$  and  $k$  radius  $r_i$ ,  $i = 1, \dots, k$ . So, we split this region into  $n \times k$  surfaces.

In Figure 7, the gray region represents the  $j$ -th rate of growing tumor in time  $\Delta T$ , noted here as  $\bar{A}_{t_0+\Delta T}^j(\phi) - \bar{A}_{t_0}^j(\phi)$  with  $j = 1, \dots, n$  ( $n =$  angular resolution). In this case, the linear resolution is  $k = 1$ . Therefore, the  $j$ -th portion of area at time  $t_0$  is

$$\bar{A}_{t_0}^j(\phi) = \int_{\phi_j} R_{t_0}^j(\sin(\phi_j) + \cos(\phi_j))d\phi_j, \quad (23)$$

and the  $j$ -th portion of area at time  $t_0 + \Delta T$  is

$$\bar{A}_{t_0+\Delta T}^j(\phi) = \int_{\phi_j} R_{t_0+\Delta T}^j(\sin(\phi_j) + \cos(\phi_j))d\phi_j. \quad (24)$$



**Figure 7.** Calculation of the growth tumor at time  $t + \Delta T$ .

Now, we have the area of the initial tumor

$$\bar{A}_{t_0}(\phi) = \varepsilon + \sum_{j=1}^n \bar{A}_{t_0}^j(\phi), \quad (25)$$

and the area of the tumor from the second tomography is given by

$$\bar{A}_{t_0+\Delta T}(\phi) = \varepsilon + \sum_{j=1}^n \bar{A}_{t_0+\Delta T}^j(\phi). \quad (26)$$

The area of the growth tumor after a time period is composed from  $n \times k$  elementary areas  $\bar{A}_i^j(\phi_j)$ , and is given by

$$A_{t_0+\alpha\Delta T}(\phi) = \varepsilon + \sum_{i=1}^k \bar{A}_{t_i}^j(\phi), \quad (27)$$

where  $\bar{A}_{t_i}^j(\phi)$  is the  $j$ -th predicted portion of area at time  $t_i + \alpha\Delta T$ , given by

$$\bar{A}_{t_i}^j(\phi) = \varepsilon + \sum_{j=1}^n \bar{A}_i^j(\phi), \quad (28)$$

and  $\bar{A}_i^j(\phi)$  is the  $i$ -th and  $j$ -th predicted portion of area

$$\bar{A}_i^j(\phi_j) = \int_{\phi_j} r_i^j(\sin(\phi_j) + \cos(\phi_j))d\phi_j. \quad (29)$$

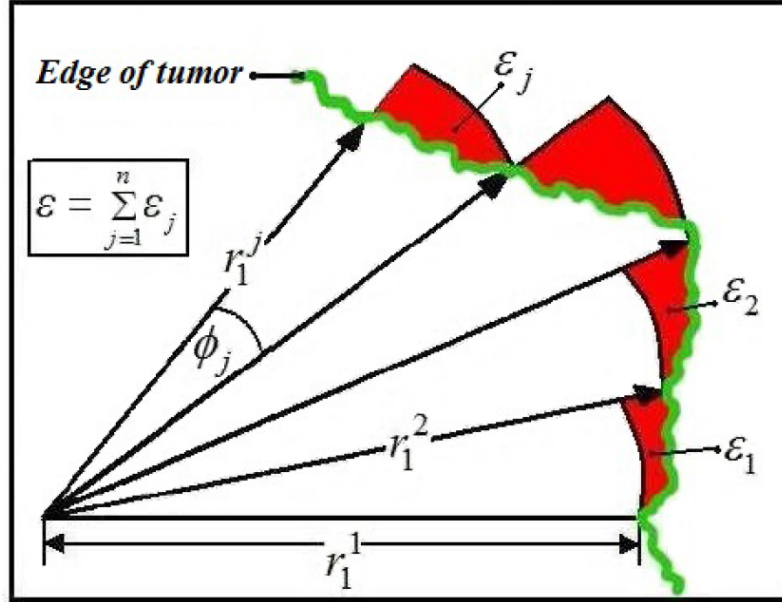
Let us now formulate our proposed algorithm. We name it cobweb, see below in Step 3 for further explanations. We need to choose the number of vectors (angular resolution) and decide for the period of prediction (linear resolution). The precision of the prediction depends on these choices. Once we have this and the values of the vectors at time  $t_0$  and  $t_0 + \Delta T$ , we can start the procedure:

**Step 1.** We compute the value of the vector  $\vec{R}_{t_{i+1}}^j = r_{i+1}^j$  using the current value of this vector (at time  $t_i$ ) and we add the difference of the value of this vector at the current time with the value of the immediate past time, with expression (22). This is represented by a point of the future bounded tumor in the direction of this vector. The union of all these points gives the entire tumor.

**Step 2.** We calculate the area  $\bar{A}_i^j(\phi)$  with expression (29).

**Step 3.** The general formula for computing the growing rate after a period  $t_i$  of time is given from the fact that a portion of area calculated in the step before is used in the current step, and in turn it is used on the following step (like a spider who is building its cobweb):

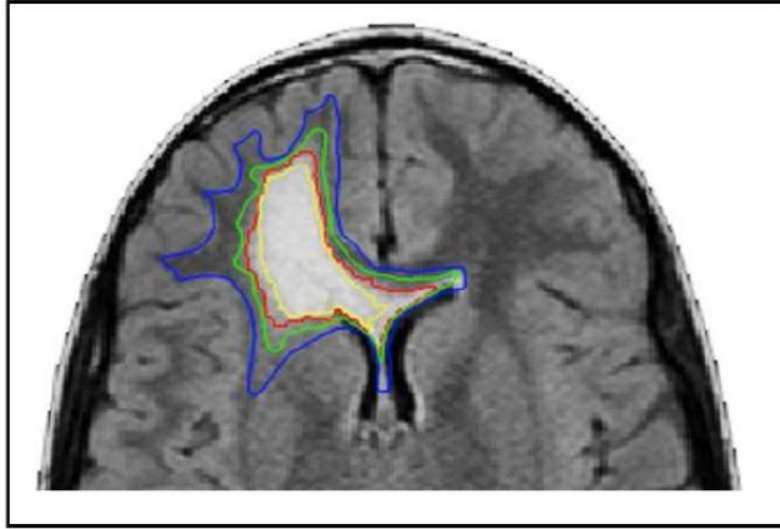
$$A_{t_{i+1}}^j = \varepsilon + \sum_{i=1}^k \sum_{j=1}^n (\bar{A}_i^j + n^{-1}k^{-1}(\bar{A}_i^j - \bar{A}_{i-1}^j)). \quad (30)$$



**Figure 8.** Error propagation.

The error  $\epsilon$  (red region in Figure 8) is given by the fact that we compute an approximated area with this method. This error can be diminished by using a numerical method such as least squares, the trapezium method or just by increasing the resolution.

In Figure 9, we provide an estimation of the growth tumor after two months and the result comes in the green region. In this case, the time period is  $2\Delta T$ . This algorithm can be extended to a three-dimensional space by replacing the vector  $\vec{R}_t^j$  with a surface generated from this vector with spherical coordinates. In this case, we should define a new resolution (radial resolution).



**Figure 9.** Predicted tumor after two months.

#### 4. Software

We implemented a set of new functions in Matlab software to proceed with our approach based on images coming from magnetic resonance imaging (MRI), computed tomography (CT) or another such techniques.

##### 4.1. Input data

As input data we need at least two images (ideally three) of the same brain tumor, taken at predetermined or known time intervals. The precision of our method of prediction is given mainly by the number of vectors in which the direction of the tumor growth is forecast; they divide the circle counterclockwise in a number of angles equal to the number of vectors, so that we can say that they represent the angular resolution. We also need to input the time between any two images, and the elapsed time since the second image.

In our case, these last two input data are days but can also be months or hours (entire unit time). The accuracy of prediction also depends on these elections based on longitudinal resolution, which gives the unit time of growth rate per vector.

#### 4.2. Procedures that must be fulfilled

The main objective of this code is to implement the algorithm presented in the previous section, and obtain a prediction of the tumor boundary. To perform this, we need to follow some stages:

(a) Comparison of the two (or three) images taken as input data to determine the rate of increase or decrease of the tumor from each chosen vector (direction).

(b) To obtain the most accurate growth rate in each direction, it is absolutely necessary that the images are taken under the same conditions, observing a single point of reference, respecting and preserving the same cartesian reference for all future tests of the same patient.

(c) In a two-dimensional situation, we have to determine the approximate tumor center for the first image, and it will be preserved for all other images (second and/or third image if necessary).

(d) A good precision is obtained if the images to be compared are taken on the same plane and do overlap. If this is not the case, this means that the patient does not have the same position and then we have to apply some transformations (translations and/or rotations).

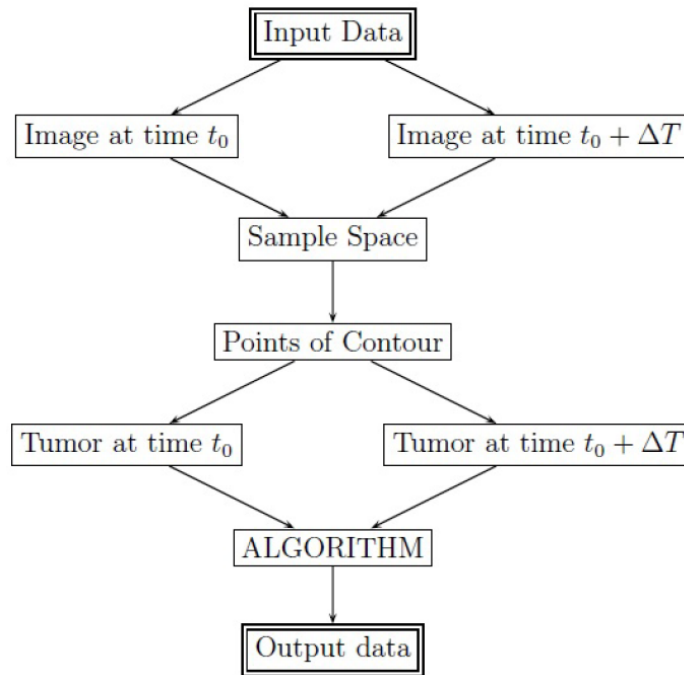
(e) To get a more precise outline of the tumor prediction, we need to compare the different stages of the tumor development. This will be done by entering the coordinate points of the contour for each instance of time. By default the number of points is set to 20 but the user can modify this to increase the precision in detriment of computing time. The user must choose the coordinates of designated points (number of contour points) exactly at the intersection of the line vector with the contour line of the edge of the tumor,  $P_i(x_i; y_i) = \overrightarrow{R}_i^j \cap Y_t$  (see Figure 2 and Figure 7).



(f) Besides the angular resolution (the number of vectors that will calculate the prediction), accuracy of calculation will be influenced by how precise and accurate the contour points are chosen. To ease the task, we use the image segmentation function (“*imcontour*” [8] in Matlab) that will depict more clearly the outline of the tumor.

(g) We can interpolate the outline by using splines over the existing points to get a more precise outline.

The output of these functions represent the contour of the predicted tumor after the time designated by the user. This will be plotted in the same two-dimensional plane with the last tumor, together with all stages of the tumor development in time. In Figure 10, we show a diagram of the built software.



**Figure 10.** Diagram of the script.

## 5. Real Data Analysis

We study a particular tumor located in the Central Nervous System and called *glioblastoma multiform*. In conformity with the World Health Organization, this tumor is the most aggressive tumor with type and grade according to the IV-th classification. Observing the scanned images, it is clear that there is presence of multiple tumors in the body, a fact called metastasis. These metastatic tumors are children of primary tumors from breast, lung, colon, stomach, and skin (melanoma), but in our case, the first one was the brain tumor.

A patient with glioblastoma multiform has an average life span of one year, receiving radiation therapy, steroids, and anti-convulsants. Otherwise the patient dies long before one year. For the patients affected by this type of tumor, a neurological deterioration is noted producing difficulty in organizing and coherently expressing ideas, and then losing the mobility function, all depending on the order in which the tumor affects the brain and areas focused on memory, speech, motor function, etc.

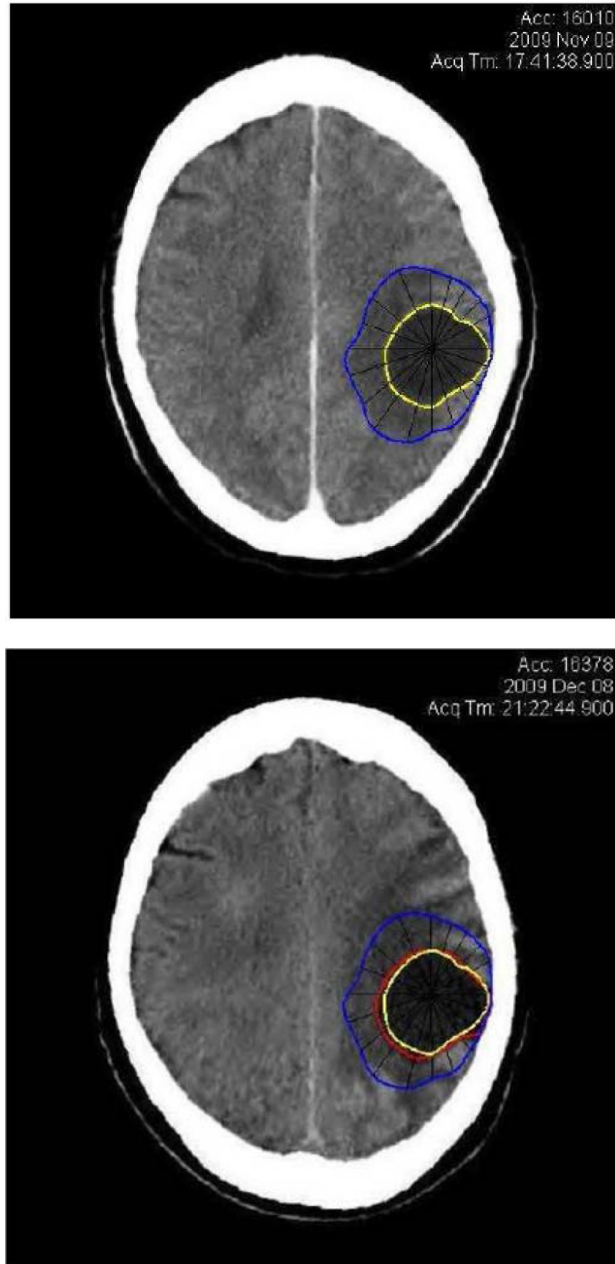
Here we selected three images taken in the same plane: two from November 9 and December 8, 2009 and one from January 10, 2010. Using the first two images, we can make a prediction of the tumor growth for the next temporal instant, and we can thus compare the prediction with the original third image.

Using the Matlab function “*linie.m*” [8], we can determine the approximate center point of the tumor by choosing 4 points in the contour of the tumor, and opposite two by two. This will plot two lines, whose intersection provides the coordinate of the center point. We then determine the sample space (blue contour) and the growing directions (vectors, black lines). By default the angular resolution is 20, meaning that the user should choose 20 points to design the “sample space”. The choice of these points must take into account physical considerations

such as: (a) the coordinates of each chosen point can be at the limit of the tumor after a certain time but not too far away; (b) if the tumor grows very quickly in one direction and the coordinate value that it can take in that direction is not physically possible (for example, it can get out of the head box), then the prediction takes as value the limit point of the sample space, which represents the maximum allowed values in this direction; (c) the designation of the sample space points can be seen as an outline of the tumor when the time is very large; and (d) the last point in the sample space must have the same coordinates as the first point, and if the user misses this, the Matlab program will do this automatically.

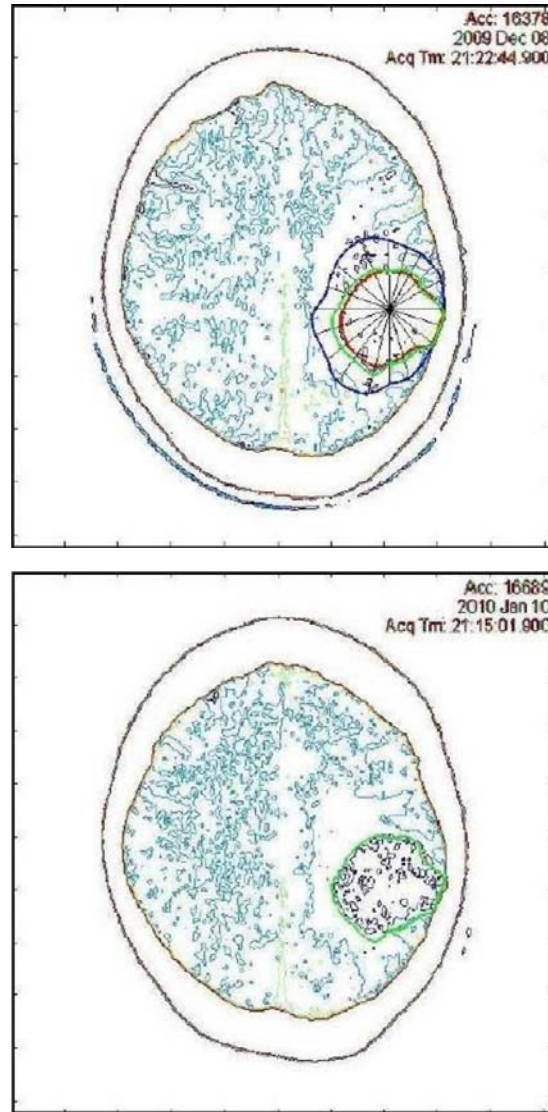
The sample space must include the tumor contour at time  $t$  and  $t + \Delta T$  with which we can make the prediction. The user must select the same number of points, but this time, at the intersection of the vectors (black lines) with the boundary of the tumor. This action will draw the outline of the first tumor with a yellow line (see Figure 11).

Then, we upload the next magnetic resonance image (the second analysis) and follow the same procedure to enter the contour points for the second tumor. This set of points should belong to the vectors and also to the boundary of the tumor, and so each point  $P_i$  must be chosen at the intersection of each vector  $R_{t+\Delta t}^j$  with the boundary tumor  $Y_{t+\Delta T}$ . This will draw a red spline curve corresponding to the boundary of the tumor at time  $t + \Delta T$  (see Figure 11).



**Figure 11.** Real data analysis. A yellow line represents the tumor at time  $t$  (time when it was discovered), and a red line represents the tumor after time  $t + \Delta T$ .

In the next step, we apply our algorithm to predict the growing tumor after 33 days. The prediction (green line in Figure 12) will be drawn and calculated as a spline curve. If we have a third image that represents the tumor stage at time  $t + \Delta T + \alpha$ , we can directly compare the predictive results with the original ones.



**Figure 12.** Predicted tumor.

In Figure 12, we can see the result of the prediction. We use a segmented picture for a better observation of the contour of the tumor. The boundary of the predicted tumor is plotted in green on the background image from December. We can note the evolution of the tumor in time: the first stage comes in a yellow line, and second stage in a red line. In the right-hand side window, we show the predicted tumor over the background image of the tumor from January 10.

## 6. Conclusion

The double stochastic process theory offers a mathematical background to study some natural and physical phenomena in the real world and it takes some conclusions and supplementary information about understanding what is happening in these complex systems.

Mathematical modelling always tries to find a compromise between simplicity of analysis and requirements of realism. On the one hand, we have extremely complex natural and biological systems; on the other hand, we need to formally address some quantitative issues about these systems, which can be often done only through the use of mathematical models that may rest on grossly over-simplified assumptions.

On some occasions, a particular mathematical formalism seems to be pre-adapted to a variety of natural and biological systems and can be profitably used to model a diverse set of processes. Double stochastic Cox processes are one class of such models, used here to solve real problems in the field of medicine.

For most of the realistic problems, the solution of the corresponding exact equation is in practice impossible, so we need to make approximations. Making approximations to solve difficult problems is not a new idea. Appropriate models enable accurate prediction of future behaviour, which can be used to control and optimize various aspects of the system in question. However, these approximations are associated with noise induced upon the real problem. The aim is to keep to a minimum this added noise, as this will increase the prediction quality.

We have presented here a mathematical-statistical approach to analyze the spatiotemporal dynamics of brain tumors. They come in form of processed computer tomography images. We interpret them as collections of image pixels with varying degrees of colour intensity levels. As such, they can be considered as a stochastic process, and we make use of spatiotemporal stochastic processes as the right statistical framework. Using this framework, we are able to predict cancer growth in space and time, and show real data analysis. The results are shown to be satisfactory, as noted in the prediction shown in Figure 12. In addition, we have implemented a Matlab software. The code is available upon contacting the authors. We should also note that we have assumed a constant growth, and in some cases, this growth can not be assumed constant, and we should adapt our modelling strategy to the case of acceleration motion. This is clearly subject of a further research.

### References

- [1] O. Barndorff-Nielsen and J. Schmiegel, Lévy based spatial temporal modelling, with applications to turbulence, *Russian Math. Surveys* 59 (2004), 63-90.
- [2] M. Bramson and D. Griffeath, On the Williams-Bjerknes tumor growth model, *Ann. Probab.* (1981), 173-185.
- [3] D. R. Cox, *The Theory of Stochastic Processes*, Chapman and Hall, London, 1994, 1955.
- [4] N. Cressie and G. M. Laslett, Random set theory and problems of modelling, *SIAM Rev.* 29 (1987), 557-574.
- [5] N. Cressie, *Modelling Growth with Random Sets*, In *Spatial Statistics and Imaging* (A. Possolo and C. A. Hayward, eds), IMS Lecture Notes Monogr. Ser. 20, IMS, Hayward, CA, 1991.
- [6] N. Cressie, *Statistics for Spatial Data*, John Wiley and Sons, New York, 1993.
- [7] M. Deijfen, Asymptotic shape in a continuum growth model, *Adv. In Appl. Probab.* 35 (2003), 303-318.
- [8] Image Processing Toolbox, Matlab user guide. Available at  
[www.mathworks.co.uk/help/pdf doc=images=images tb:pdf](http://www.mathworks.co.uk/help/pdf_doc/images/images_tb.pdf)
- [9] Ýr. Jónsdóttir Kritjana and Eva B. Vedel Jensen, Gaussian radial growth, *Image Anal. Sterol.* 24 (2005), 117-126.

- [10] Ýr. Jónsdóttir Kritjana, J. Schmiegel and Eva B. Vedel Jensen, Lévy-based growth models, *Bernoulli* 14(1) (2008), 62-90.
- [11] A. R. Kansal, S. Torquato, G. R. Harsh, E. A. Chiocca and T. S. Deisboeck, Simulated brain tumor growth dynamics using a three-dimensional cellular automaton, *J. Theor. Biol* 203 (2000), 367-382.
- [12] T. Lee and R. Cowan, A Stochastic Tessellation of Digital Space, In *Mathematical Morphology and its Applications to Image Processing* (J. Serra ed.), 217-224. Dordrecht: Kluwer, 1994.
- [13] J. Møller, Shot noise Cox processes, *Adv. Appl. Prob. (SGSA)* 35(3) (2003a), 614-640.
- [14] J. Møller and R. P. Waagepetersen, *Statistical Inference and Simulation for Spatial Point Processes*, Chapman & Hall/CRC, Boca Raton, 2004.
- [15] A. S. Qi, X. Zheng, C. Y. Du and B. S. An, A cellular automaton model of cancerous growth, *J. Theor. Biol.* 161 (1993), 1-12.
- [16] D. Richardson, Random growth in a tessellation, *Proc. Cambridge Philos. Soc.* 74 (1973), 515-528.
- [17] T. Williams and R. Bjerknes, Stochastic model for abnormal clone spread through epithelial basal layer, *Nature J.* (1972), 236.

