

Time series forecasting using ARIMA for modeling of glioma growth in response to radiotherapy

Previsão com séries temporais usando ARIMA para modelagem de crescimento de glioma em resposta à radioterapia

Larissa Miguez da Silva¹; Gustavo Benitez Alvarez²; Eliane da Silva Christo³; Vanessa da Silva Garcia⁴; Gerardo Amado Pelén Sierra⁵

Abstract

In present days, the growing number of people suffering from cancer has been a major cause for concern worldwide. Glioblastoma in particular, are primary tumors in glial cells located in the central nervous system. Because of this sensitive location, mathematical models have been studied and developed as alternative tools for analyzing tumor growth rates, assisting on the decision-making process for treatment dosage, without exposing the patient's life. This paper presents two time series models to estimate the growth rate of glioblastoma in response to ionizing radiotherapy treatment. The results obtained indicate that the proposed time series methods attain predictions with a Mean Absolute Percentual Error (MAPE) of approximately 1% to 4%, and simulations show that the Autoregressive Integrated Moving Average (ARIMA) method surpasses the Holt method based on the Mean Square Error (MSE) and MAPE values obtained. Furthermore, the results show that the time series method is applicable to data from two different mathematical models for glioblastoma growth.

Keywords: Mathematical models. Tumor growth. Glioblastoma. Time series forecast. Radiotherapy.

Resumo

Atualmente, o crescente número de pessoas que sofrem de câncer tem sido um grande motivo de preocupação em todo o mundo. Os glioblastomas, em particular, são tumores primários em células gliais localizadas no sistema nervoso central. Por conta dessa localização sensível, modelos matemáticos têm sido estudados e desenvolvidos como ferramentas alternativas para análise das taxas de crescimento tumoral, auxiliando na tomada de decisão quanto à dosagem do tratamento, sem expor a vida do paciente. Este artigo apresenta dois modelos de séries temporais para estimar a taxa de crescimento do glioblastoma em resposta ao tratamento com radioterapia ionizante. Os resultados obtidos indicam que os métodos de séries temporais propostos obtêm previsões com Mean Absolute Percentual Error (MAPE) de aproximadamente 1% e 4%, e as simulações mostram que o método ARIMA supera o método de Holt com base no Mean Square Error (MSE) e MAPE. Além disso, os resultados mostram que o método das séries temporais é aplicável a dados de dois modelos matemáticos diferentes para o crescimento de glioblastoma.

Palavras-chave: Modelos matemáticos. Crescimento tumoral. Glioblastoma. Séries temporais. Radioterapia.

¹ Ms., Masters Prog. in Comput. Modeling in Science and Tech., UFF, Volta Redonda, RJ, Brazil, E-mail: larissamiguez@id.uff.br

² Prof. Dr., Department of Exact Sciences, UFF, Volta Redonda, RJ, Brazil, E-mail: benitez.gustavo@gmail.com

³ Profa. Dra., Department of Production Engineering, UFF, Volta Redonda, RJ, Brazil, E-mail: elianechristo@id.uff.br

⁴ Profa. Dra., Department of Administration, UFF, Volta Redonda, RJ, Brazil, E-mail: vanessagarcia@id.uff.br

⁵ Ms., Masters Prog. in Comput. Modeling in Science and Tech., UFF, Volta Redonda, RJ, Brazil, E-mail: gapshn@gmail.com

Introduction

On the present world stage, concern with cancer illnesses has intensified due to its undeniable growth tendency. Globally it is estimated that the number of people suffering from cancer has increased to 18.1 million new cases and 9.6 million deaths in 2018 (IARC, 2018). Growth on this index is due to several factors, including population growth and/or aging, as well as the population's life style changes, social and economic development (IARC, 2018; INCA, 2018; CANCER INSTITUTE NSW, 2018).

There are more than one hundred types of cancer, which originate from different parts of the human body. Particularly, this paper focuses on gliomas, which are a collection of tumors originating at the glia or its precursors on the central nervous system (SNC) (HOLLAND, 2000). They are malignant tumors presenting accelerated growth with a high degree of lethality and very short average survival time. Conventional treatment of gliomas involves a combination of different types of therapies, such as surgery, radiotherapy and/or chemotherapy (BARNETT, 2007; BELLOMO; CHAPLAIN; ANGELIS, 2008). Conventional treatment shows little improvement in the patient's survival time and quality of life. There are other treatments like BNCT (Boron Neutron Capture Therapy) and treatments that use stem cells. However, they are not yet available in Brazil. Our research does not intend to make a choice of the best therapy, but to understand and develop mathematical models to optimize treatments and help in understanding this disease.

Radiotherapy was chosen here because more than 50% of patients undergoing glioma treatment use radiotherapy to fight the disease. This number derives from the fact that radiotherapy is precise in removal of tumoral mass, without causing neurological damage (WALKER; STRIKE; SHELINE, 1979). Meanwhile, it becomes necessary to forecast radiation dosages to be ministered to the patient during treatment, this for the dose-effect relationship to remain harmless. In future works we must address other therapies. Thus, it will be possible to make comparisons between therapies and combinations of them to seek better treatment optimization.

The main goal of cancer treatment therapies is destroying cancer cells preserving the healthy ones. In this context the importance of the interdisciplinary relationship between medical science and mathematics in health research studies stands out, both aiming towards the same goal of developing models and techniques capable of con-

tributing to cancer fighting research (BARNETT, 2007; BELLOMO; CHAPLAIN; ANGELIS, 2008; HOLLAND, 2000; MURRAY, 2012).

Two mathematical models can be found in literature which describe the growth rate of gliomas (DOLGIN, 2014; LEDER *et al.*, 2014; ROCKNE *et al.*, 2008; SWANSON; ALVORD JUNIOR; MURRAY, 2000), both models changing continually in time. The model proposed in Swanson, Alvord Junior and Murray (2000) and Rockne *et al.* (2008) consists of a partial differential equation, and it considers the liquid proliferation and tumor invasion rates, seen as a solid form (macro). On the other hand, the model proposed in Leder *et al.* (2014) and Dolgin (2014) is based on Markov's chains, and considers the dynamics of two different cancer cell populations, and beyond their propagation it establishes a conversion and reversion rate between these populations. Even with all this, in practice it is very difficult to accompany this increase and decrease in the populations of cancer cells continuously. In general, in population growth studies measurements are performed at regular time intervals. In this case, time is said to be discrete, because population size is known point by point. Thus, discrete model utilization becomes viable.

There are several discrete models such as Bernoulli's, binomial and models based on stochastic processes such as time series. Time series has been integrated into several research works, aiding with the inference on the basic properties of the series observations (DOMINGOS; OLIVEIRA; MATTOS NETO, 2019; TEIXEIRA; IKEHARA; BRANDT, 1981; ZHANG, 2003). The application of models based on time series linked in association with computing have shown a degree of precision without significant error. For example, a model utilizing the Holt method with time series is presented as a new alternative for glioma growth models (JESUS *et al.*, 2014). In Jesus *et al.* (2014) it is showed that the Holt's exponential smoothing method is efficient to describe the data from mathematical model based in reactive-diffusive partial differential equation (ROCKNE *et al.*, 2008).

This paper, together with Jesus *et al.* (2014), reaffirms the great potential that exists in the use of statistical approaches to understand glioblastoma growth. The objective of this work is to apply two time series models for tumor growth predictions based on simulated data from model in Leder *et al.* (2014). The two time series models used here are both well known statistical approaches: ARIMA method and Holt method. It is well known that there are more powerful time series methods, however the two choices made proved to be sufficient to repro-

duce the results of mathematical model in Leder *et al.* (2014). The results of the two time series methods are compared with results in Leder *et al.* (2014) and Dolgin (2014), and several analysis are taken into consideration in accordance with the dosage fractioning schedule, where four different scenarios exist. Since the parameters of the two mathematical models can be adjusted for a single individual (LEDER *et al.*, 2014; ROCKNE *et al.*, 2008), and considering that both models can be reproduced satisfactorily by time series, then it is possible to state that the time series is a statistical approach that can also be adjusted to a single individual.

Mathematical models of tumor growth

Population growth dynamics can be observed in countless nature phenomena and they are utilized in several different contexts. In the field of Biology, these models can be used to describe cell proliferation and its dynamics in the environment (MURRAY, 2012).

In recent decades, many mathematical models have been developed in the literature aiming to observe tumoral behavior. For example, models based on partial differential equations such as the reactive-diffusive model (ROCKNE *et al.*, 2008). In addition to those models that consider homogeneity of tumoral cells, a different model was proposed in Dolgin (2014), where Markov's chains are employed and heterogeneity of tumoral cells and dynamically acquired radioresistance to forecast the efficacy of different radiation schemes are taken into consideration.

In addition to these models, studies found that using Holt's time series based method can be effective as a new alternative to tumor growth models (JESUS *et al.*, 2014). However, for the use of time series, a set of observations is required to estimate the growth of glioma over time after therapy. Due to the limitations of data available in the literature and the lack of clinical histories, in the paper mentioned above time series were based on data generated by the mathematical model via reactive-diffusive equation (ROCKNE *et al.*, 2008).

In this work, the time series were generated from the model described in Leder *et al.* (2014). For this purpose, MATLAB[®] software was used where a computer code was made for the model solution via Markov's chains, represented by equations (1) and (2). These two equations describe the dynamics between two glioblastoma cell sub-populations in response to radiation therapy. The linear-quadratic model is used for fractionation of therapy, where the i th dose of treatment is denoted by d_i .

Thus, N_{i+1}^S denotes the number of radiation-sensitive cells (RSC) at time t surviving a dose of treatment d_i , and N_{i+1}^R is the number of radiation-resistant cells (RRC). N_i^S and N_i^R represent the respective populations before applying dose d_i . Time t corresponds to the time between treatments i and $i + 1$.

$$\begin{aligned} N_{i+1}^S(t) = & N_i^S e^{-[\alpha_S d_i + \beta_S d_i^2]} \left[(1 - \gamma) e^{r_S(t-L_S)^+} + \gamma e^{-\nu t} + \right. \\ & + a_S \gamma \nu \int_0^t e^{r_S(t-\tau-M_S)^+} \int_0^{(\tau-L_R)^+} e^{-\nu y} e^{r_R(\tau-y-L_R)^+} dy d\tau \left. \right] + \\ & + a_S N_i^R e^{-[\alpha_R d_i + \beta_R d_i^2]} \int_{L_R}^{\max(t, L_R)} e^{r_R(\tau-L_R)} e^{r_S(t-\tau-M_S)^+} d\tau. \end{aligned} \quad (1)$$

$$\begin{aligned} N_{i+1}^R(t) = & N_i^R e^{-[\alpha_R d_i + \beta_R d_i^2]} e^{r_R(t-L_R)^+} + \\ & + N_i^S e^{-[\alpha_S d_i + \beta_S d_i^2]} \gamma \nu \int_0^t e^{-\nu s} e^{r_R(t-\tau-L_R)^+} d\tau. \end{aligned} \quad (2)$$

Concerning the dynamics between these two cell sub-populations, it has been reported that the rate of proliferation of RSC and RRC by the end of the quiescence is given for r_S and r_R respectively. RRC cells convert to RSC cells with a rate of a_S , and ν denotes the rate at which RSC revert to the RRC. The α_S , α_R , β_S and β_R parameters belong to the linear-quadratic model and represent the lesions in the DNA produced from a single radiation beam in the cancer cells. The parameter γ denotes the fraction of RSC that revert to the RRC. The L_S and L_R parameters represent the minimum time that RSC and RRC cells take to return to the cell cycle respectively, and M_S is the minimum time for a newly created RSC to lead to clonal expansion. The notation $()^+$ assumes the value zero if the expression within parentheses is negative, otherwise it assumes the value of the expression. All these parameters were adjusted based on experiments with mice (LEDER *et al.*, 2014).

Finally, let N_1^S and N_1^R be their initial populations just before the first dose of radiotherapy, and consider the following relationship between these two initial populations $R = N_1^S / N_1^R$. Therefore, after completion of radiotherapy fractionation in K doses, the quotient of tumor cell relative to the initial population is defined as

$$x_i = \frac{N_{K+1}^S(t) + N_{K+1}^R(t)}{N_1^S + N_1^R} = \frac{N_{K+1}^S(t) + N_{K+1}^R(t)}{N_1^S [1 + 1/R]}. \quad (3)$$

Four dose fractionation scenarios were simulated, as shown in Table 1. The first scenario, Single, corresponds to the treatment with a single dose given on the first day.

The second scenario, Standard, corresponds to the treatment used frequently in hospitals and clinics, and consists of splitting radiotherapy into five equal doses given considering equal temporal interstices. The third and fourth scenario, Optimum and Optimum-2, corresponds to treatments with dose fractionations different from the Standard, which result in less tumor growth for this mathematical model. It is important to mention that these two “optimal” schemes do not correspond to the optimal mathematical. They are the best fractionation schemes found in Leder *et al.* (2014) using a heuristic optimization method. We adopted the same treatment schemes described in Leder *et al.* (2014) to show that time series modeling can reproduce similar results.

Table 1 – Therapy schedule.

Schedule	Dose	Mon	Tue	Wed	Thu	Fri
Single ($K = 1$)	10 Gy	8am	-	-	-	-
Standard ($K = 5$)	2 Gy	8am	8am	8am	8am	8am
Optimum ($K = 10$)	1 Gy	8am 2pm 5pm	5pm	3pm 5pm	5pm	3pm 4pm 5pm
Optimum-2 ($K = 8$)	1 Gy	-	4pm	-	9am 1pm 5pm	9am 1pm 5pm
	3 Gy	8am	-	-	-	-

Source: Leder *et al.* (2014).

The Table 2 shows the numerical values used for each parameter in the simulations. For more details on the model consult the reference Leder *et al.* (2014).

Table 2 – Parameters used in the simulations.

Parameter	Value/Units
R	20
α_S	0.0987 1/Gy
β_S	$1.14 \cdot 10^{-7}$ 1/Gy ²
α_R	0.0395 1/Gy
β_R	$4.58 \cdot 10^{-8}$ 1/Gy ²
γ	0.15
v	1.15 1/hour
r_S	0.0088 1/hour
a_S	0.0001 1/hour
L_S	24 hour
L_R	36 hour
M_S	24 hour
r_R	0.0001 1/hour

Source: Leder *et al.* (2014).

Two time series

The models used to describe time series are stochastic processes, that is, sets of random variables that vary with time where probabilistic laws apply (MORETTIN; TOLOI, 1981). Moreover, a time series entails serial dependence. Being a time series a set of observations, measured sequentially over time, these may be related to time series of continuous or discrete-time. Thus, a time series is said to be discrete when the set of observations is finite or countably infinite, and it is said to be continuous when the set is non-countably infinite (MORETTIN; TOLOI, 1981).

Then, to model tumor growth as a univariate discrete time series, the variable X to be considered, will represent the relative number of cancer cells in a given time interval after the application of radiation therapy, as given by the equation (3). Since x_t can be considered a random variable, it is inferred that the time series x_t can be defined as a sampling of a stochastic process over X . Therefore, the use of time series, as a forecasting tool, can be an alternative method for simulating the most suitable treatment for the patient in the cases studied of radiotherapy dose fractionation. In addition, time series can be used as a comparison tool between two glioma growth models. For instance, the model proposed in Rockne *et al.* (2008) and the model proposed in Leder *et al.* (2014).

R software was used for statistical computing and graphics. After performing the simulation of mathematical model in Leder *et al.* (2014), charts were produced for a better evaluation of the existing components present in the time series in order to determine the best methods to be utilized. For this purpose, the relative number of cancer cell error adjustments were calculated utilizing the MAPE - Mean Absolute Percent Error - for Holt’s exponential smoothing methods with additive and multiplicative error, in order to monitor the calculation and verify if the adjustments applied remain as close as possible to the simulated data (HYNDMAN *et al.*, 2008). Such error is calculated using the following equation

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|x_t - \tilde{x}_t|}{x_t}, \quad (4)$$

where x_t are the actual values of the time series in period t , \tilde{x}_t are the predicted values for the time series in period t and N is the number data of the time series.

Holt method

From the graphical analysis of the charts, it was observed that the data present just an exponential trend without the presence of a seasonality component. When a given series has an additive (linear) trend and does not exhibit seasonality, Holt's double exponential smoothing model can be used for forecasting (HYNDMAN, 2002; HYNDMAN *et al.*, 2008; HYNDMAN; AKRAM; ARCHIBALD, 2008). This model, unlike the former, has two smoothing components, α and β ranging between 0 and 1. The prediction is given by

$$Z_{t+k} = L_t + kT_t. \quad (5)$$

The update equations for level and trend parameters are

$$L_t = \alpha x_t + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (6)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (7)$$

where Z_{t+k} is the relative number of cancer cells prediction, h is the prediction horizon with $k = 1, 2, \dots, h$. Being L_t the level component of the series and T_t the trend component of the series, and x_t is the relative number of cancer cells observed at time t . This method, in turn, can be analyzed with additive and multiplicative error, being represented by (A,A,N) and (M,A,N), respectively.

ARIMA method

An alternative method which can also be fitted to the dataset generated by equation (3) is the Autoregressive Integrated Moving Average (ARIMA) method, a more robust and flexible model that, according to the Box-Jenkins methodology (BOX *et al.*, 2015), yields predictions based on the current and past values of these series. It is important to note that the series should be composed of more than 30 histories datas for best results. In time series analysis an autoregressive integrated moving average model is a generalization of an autoregressive moving average model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series.

When considering a non-stationary trend-driven process, a transformation applying a difference operator can make the series stationary. Hence, an ARIMA (p, d, q) model is an autoregressive integrated moving average process of d order suited to represent the portrayed series in this case.

In the case of a non-stationary variable, the following equation is used (BOX *et al.*, 2015; MORETTIN; TOLOI, 1981; MONTGOMERY *et al.*, 2015)

$$Z_t = x_t - x_{t-1} = (1 - B)^1 \Delta^1 x_t, \quad (8)$$

where B is the "Backward Shift Operator" widely used by Box *et al.* (2015) when describing BJ models, and is defined as $B^k = Z_t - Z_{t-k}$. Considering that Z_t is stationary, it follows that the x_t variable is integrated of order 1. Taking d differences, the above equation can be rewritten as

$$Z_t = \Delta^d x_t. \quad (9)$$

Hence, the variable Z_t adheres to an ARIMA (p, d, q) process as follows

$$Z_t = \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \dots + \beta_p Z_{t-p} + \dots + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}. \quad (10)$$

Thus, substituting equation (8) to (10), we obtain

$$\Delta x_t = \beta_1 \Delta x_{t-1} + \beta_2 \Delta x_{t-2} + \dots + \beta_p \Delta x_{t-p} + \dots + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}. \quad (11)$$

ARIMA is actually a class of models that explains a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values.

Results

Exponential smoothing model

The evolution of glioma in response to several cases of dose fractionation was simulated as previously presented. The data generated by such simulation provided the basis for the series utilized in this paper. The time series aforementioned feature 100 observations over a 50 day period, being collected in 12-hour intervals. However, only 90 simulated data were used for forecasting and the last 10 simulated data will be predicted. When using such data, the α and β parameters found for the Holt's exponential smoothing model are shown in the Table 3. It is possible to verify, through the AIC - Akaike Information Criterion, that the most appropriate model for all cases was the (A,A,N) model. In addition, it is also observed that as the treatments improved the β values decrease.

Table 3 – Exponential smoothing model results for different types of fractionation.

Cases/Models	(A,A,N)			(M,A,N)		
	α	β	AIC	α	β	AIC
Single	0.60	0.99	361.916	0.50	0.33	179.235
Standard	0.35	0.99	-244.128	0.51	0.99	-102.952
Optimum-2	0.75	0.47	-200.615	0.64	0.65	-114.401
Optimum	0.99	0.27	-266.981	0.99	0.23	-82.587

Source: The authors.

This last observation may be important for treatment optimization (FERNÁNDEZ-CARA; PROUVÉE, 2018). That is, the results in Table 3 suggest that there is a functional dependence between dose fractionation and β , such that the more effective the treatment, the lower the parameter value. This dependency, if known, can be useful for finding the best dose fractionation, as a computational algorithm that minimizes the parameter β could be constructed.

ARIMA model

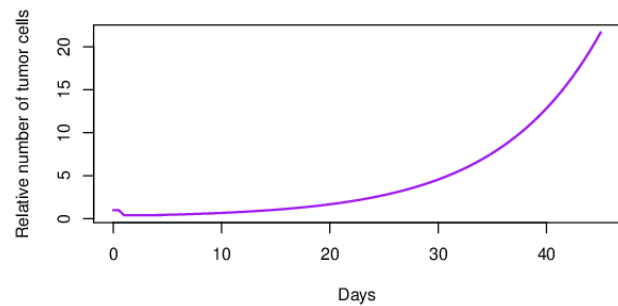
In addition to using Holt’s exponential smoothing method for the treatment of the simulated series, an alternative forecasting method was utilized: the ARIMA method. The time series used here were the same as in the previous section, they consisted of 100 observations over a 50 day period, collected on 12-hour intervals. However, only 90 simulated data elements were used for the forecast, and they were organized by cases to better explain how the models were found.

Single Case:

In this case, the series is shown in Figure 1. It is known that all work for model parameter estimate calculation revolves around the autocorrelation (ACF) and partial autocorrelation (PACF) functions. To show that the method is well adjusted, all resulting values must lie within the 95% confidence interval, bounded by the dotted blue line. However, this is not always the case, so alternatively the best fit scenario is searched for Figure 2 shows ACF and PACF of the residuals, respectively. It can be observed that there are many "lags" above the region’s confidence interval, represented by the blue dotted line. This fact indicates the presence of non-stationarity.

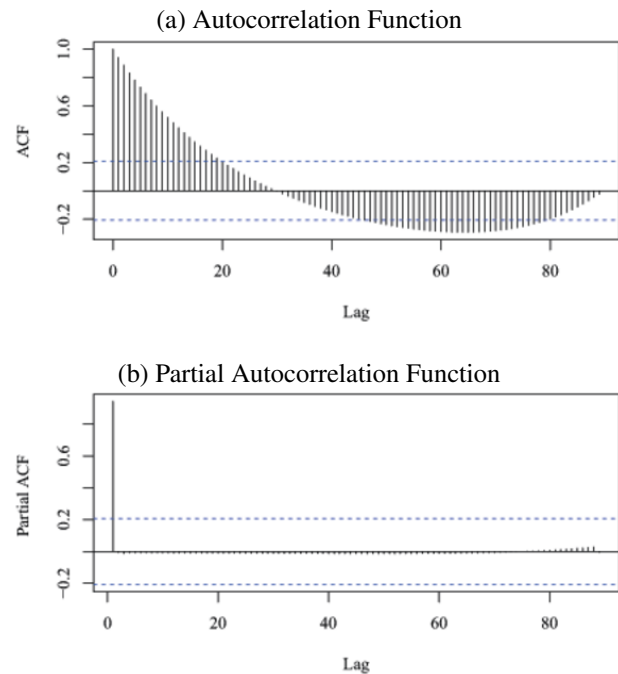
To make the series stationary, the difference operator is applied until the series passes the test. Thus, after two differentiations a better result was obtained. Finally, the ARIMA (2,2,2) model is obtained. After applying the Box-Pierce test the results found

Figure 1 – Simulated model that predicts tumor response and growth following single dose.



Source: The authors.

Figure 2 – Autocorrelation and Partial Autocorrelation Functions of the residuals for Single Case.



Source: The authors.

were: $(X - squared) = 6.3352, df = 1$ and $(p - value) = 0.01184$.

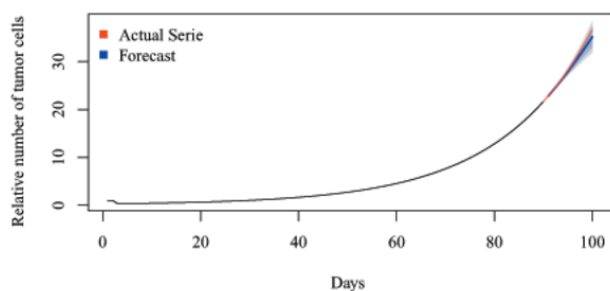
This test verifies the autocorrelations for the estimated residuals, and through its application, it is observed whether the null hypothesis will be accepted or not. The case of a null hypothesis is met when the residuals are

independent and identically distributed (i.d.d.), and in contrast, there exists the h_1 hypothesis which represents the opposite of the null hypothesis.

A method of verifying this hypothesis is whether the estimated model residuals are not correlated. In this case, the value of p is a probability that measures the evidence against the null hypothesis, having a value less than the 5 % significance level. Therefore, it is concluded that the ARIMA (2,2,2) is an acceptable model for this data set.

The next stage is the forecasting phase, where the ARIMA model chosen is used for this case. Predictions were made for 10 steps forward, and the results are shown in the Figure 3. The values predicted by the ARIMA model (2,2,2), shown in blue, and the simulated actual values, shown in red. It can be observed that the predicted results were satisfactory within a confidence interval of 95 %, shown in the blue shadow, and presented a MAPE of only 1 %.

Figure 3 – Series observed with the prediction for the ARIMA model (2,2,2) of the Single Case.



Source: The authors.

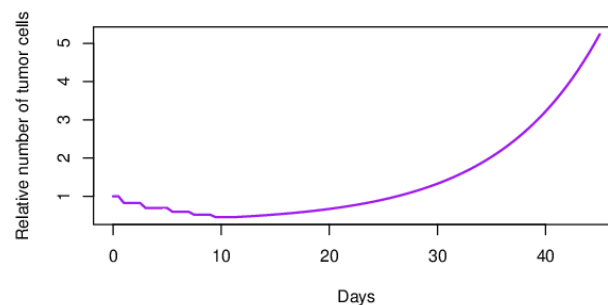
Standard Case:

In this case, the series can be seen in Figure 4. Similarly, the autocorrelation charts were verified for the standard case. For this case, two differentiations were necessary. After the second one, it was obtained that the best model found was the ARIMA (2,2,6), since it showed the lowest error in the predictions. The result is shown in Figure 5. Although the result was not as good as the previous one, this prediction method achieved a MAPE of just 4 %.

Optimum-2 Case:

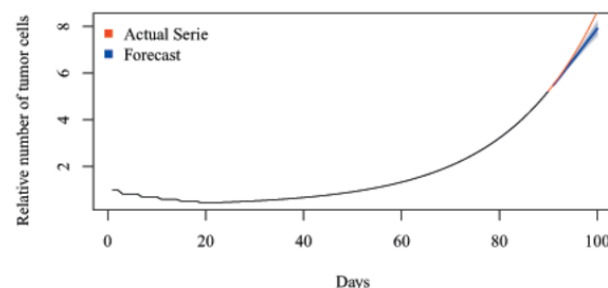
In this case, the series can be seen in Figure 6. It was concluded that the best prediction model is ARIMA (2,2,2) model. This model yields the prediction shown in Figure 7. In this case, a satisfactory result was obtained where the MAPE showed an error of only 4%.

Figure 4 – Simulated model that predicts tumor response and growth following Standard Case.



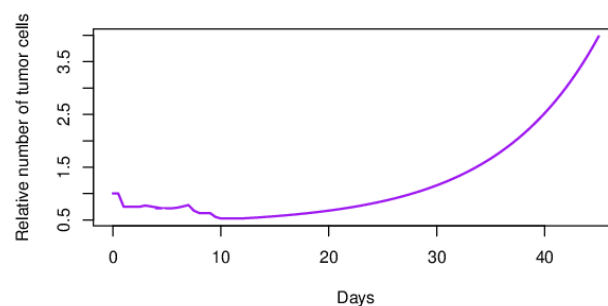
Source: The authors.

Figure 5 – Series observed with the prediction for the ARIMA model (2,2,6) of the Standard Case.



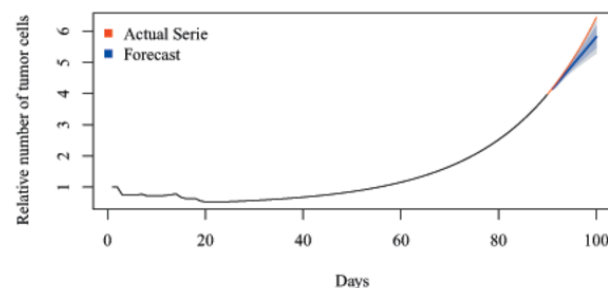
Source: The authors.

Figure 6 – Simulated model that predicts tumor response and growth following Optimum-2 Case.



Source: The authors.

Figure 7 – Series observed with the prediction for the ARIMA model (2,2,2) of the Optimum-2 Case.

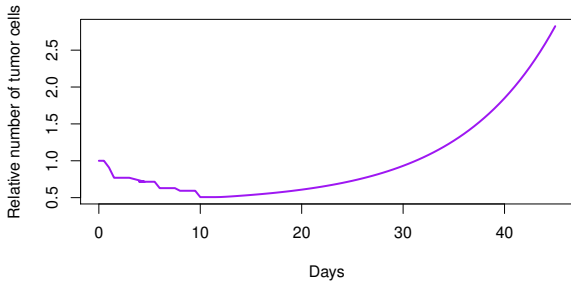


Source: The authors.

Optimum Case:

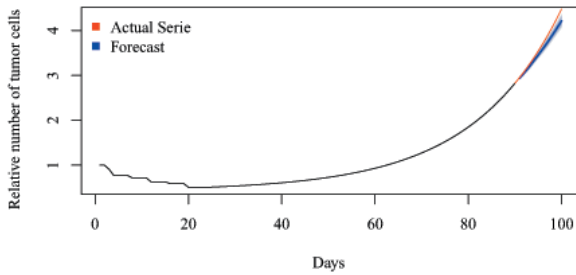
The series can be seen in Figure 8. The best model found was the ARIMA (4,2,5) model, and its prediction can be observed in Figure 9. In this case, estimates close to the 95% confidence interval were obtained, with a MAPE of only 1%. Therefore, the ARIMA (4,2,5) model was effective for predictions in this case.

Figure 8 – Simulated model that predicts tumor response and growth following Optimum Case.



Source: The authors.

Figure 9 – Series observed with the prediction for the ARIMA model (4,2,5) of the Optimum Case.



Source: The authors.

Discussions

Mathematical models assist in decision making both in diagnosis and treatment. In particular, Markov’s chain mathematical model (LEDER *et al.*, 2014) proves helpful for a detailed understanding of the dynamics between two types of cancer cells. In addition, this model suggests as a better treatment in radiotherapy a fractional dose different from the standard. This result is characteristic of models that use more than one type of different cells to describe tumor growth. The model based on the partial differential equation uses a single cell type, and presents standards as the best treatment (ROCKNE *et al.*, 2008). The standard scheme as the best treatment using this type of model was confirmed in Junior (2016), Souza, Neves and Alvarez (2015) and Barbosa *et al.* (2019). These different predictions deserve further study of both models (LEDER *et al.*,

2014; ROCKNE *et al.*, 2008). However, time series modeling reproduces the predictions of both models (LEDER *et al.*, 2014; ROCKNE *et al.*, 2008), since similar results to those described in Rockne *et al.* (2008) were reproduced in Jesus *et al.* (2014). Here, similar results to those described in Leder *et al.* (2014) were reproduced.

Just as well as the model proposed in Leder *et al.* (2014), the time series revealed being effective in tumor growth behavior prediction, presenting errors of about 1 to 4 %. The ARIMA method behaved better in all situations when compared to Holt’s exponential smoothing forecast method since it presents greater complexity in its formulation.

The Tables 4-7 show a comparison between these two time series models for all simulated dose fractionation cases. It can be seen that the ARIMA model behaved better in all cases, presenting a lesser MAPE and AIC. Such behavior was expected, as the ARIMA model is based on the integration of moving averages and the autocorrelation between deviations and lagged observations. For the case of Holt’s exponential smoothing model, as its name suggests, it makes an exponential smoothing of the data and distributes the largest weights for the initial values. In summary, the ARIMA method is broader and distributes equal weights for the whole series, thus guaranteeing better results for the models portrayed here.

Table 4 – Comparison between ARIMA and Holt’s exponential smoothing models for Single Case.

Single Case	AIC	MAPE
Holt (A,A,N)	361.916	0.047828135
ARIMA (2,2,2)	-219.17	0.015342817

Source: The authors.

Table 5 – Comparison between ARIMA and Holt’s exponential smoothing models for Standard Case

Standard Case	AIC	MAPE
Holt (A,A,N)	-244.128	0.048848229
ARIMA (2,2,6)	-474.75	0.040047561

Source: The authors.

There is one open question with regard to the time series for the application in this work. In the data set used, it is noted that the error during treatment with ionizing radiation is greater than after treatment. This may be due to the presence of volatility in the data, caused by the quiescent period in the cell dynamics considered by the

Table 6 – Comparison between ARIMA and Holt’s exponential smoothing models for Optimum-2 Case

Optimum-2 Case	AIC	MAPE
Holt (A,A,N)	-200.615	0.054364309
ARIMA (2,2,2)	-351.5500	0.042270122

Source: The authors.

Table 7 – Comparison between ARIMA and Holt’s exponential smoothing models for Optimum Case

Optimum Case	AIC	MAPE
Holt (A,A,N)	-266.981	0.061801895
ARIMA (2,2,5)	-455.56	0.011628106

Source: The authors.

model (LEDER *et al.*, 2014). That is, the processed data indicates the need to use two different time series. Each of these two time series would be dominant at different time intervals. The first time series, to be determined, would be dominant until the end of radiotherapy application, and the second time series, used in this article, would be dominant from the end of radiotherapy. As an alternative to this objection, we propose as future work a study of ARIMA models in conjunction with machine learning models to make better forecasting during treatment with ionizing radiation (DOMINGOS; OLIVEIRA; MATTOS NETO, 2019; ZHANG, 2003).

Conclusions

In this work, the time series methodology was applied to predict the glioblastoma growth rate in response to radiotherapy. Additionally, the Markov’s chain mathematical model for glioblastoma described in Leder *et al.* (2014) was presented briefly.

Initially, simulations were performed using the Markov’s chain model (LEDER *et al.*, 2014), due to the lack of clinical data in the literature. Then, time series were applied to predict tumor growth, a few steps forward, the theoretical results established were put into practice through the use of these same series. On this occasion, two prediction methods were introduced as well as their behavior shown for the data herein processed. It can be observed that the time series modeling that produced the best performance among the (A,A,N) and (M,A,N) methods was the Holt’s exponential smoothing method (A,A,N), having a MAPE of around 4% to 6%. Subsequently, another method via time series was used: the ARIMA method. Yet, as was to be expected, the ARIMA model produced

better predictions when compared with Holt’s exponential smoothing model, displaying a MAPE of 1% to 4%.

Furthermore, the results obtained here together with those obtained in previous article (JESUS *et al.*, 2014) show that the time series method is applicable to data from two different mathematical models for glioblastoma growth. This is because in previous article (JESUS *et al.*, 2014), the Holt’s exponential smoothing method is efficient to model the simulated data via a reactive-diffusive partial differential equation (ROCKNE *et al.*, 2008). On the other hand, in the present study the ARIMA and Holt’s exponential smoothing methods are efficient to model the simulated data via Markov’s chain mathematical model (LEDER *et al.*, 2014). Therefore, these results show that the time series method can be considered a new alternative model to describe glioblastoma growth. The time series analysis is based on a stochastic process, this allows to work with a confidence level of a predetermined statistical probability. This feature highlights the importance of this analysis alternative for decision making in the fractionation of radiotherapy applied to the patient. In addition, the time series model can be adjusted for a single individual if individual patient data is available.

Acknowledgments

The authors thank the Federal Fluminense University and CAPES financial support. They are also grateful to the Coordination for the Dean of Student Affairs - PROAES for post-graduate financial assistance.

References

- BARBOSA, O. X.; ASSIS, W. L. S.; GARCIA, V. S.; ALVAREZ, G. B. Computational simulation of gliomas using stochastic methods. *Pesquisa e Ensino em Ciências Exatas e da Natureza*, Cajazeiras, v. 3, p. 199–215, 2019.
- BARNETT, G. H. *High-grade gliomas: diagnosis and treatment*. New Jersey: Humana Press, 2007.
- BELLOMO, N.; CHAPLAIN, M.; ANGELIS, E. *Selected topics in cancer modeling: genesis, evolution, immune competition, and therapy*. Boston: Birkhäuser, 2008.
- BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. *Time series analysis: forecasting and control*. 5th ed. New Jersey: John Wiley & Sons, 2015.

- CANCER INSTITUTE NSW. *What is cancer*. Alexandria: NSW, 2018. Available from: <https://www.cancerinstitute.org.au/understandingcancer/what-is-cancer>. Access in: jun. 2018.
- DOLGIN, E. The mathematician versus the malignancy. *Nature Medicine*, New York, v. 20, p. 460–463, 2014.
- DOMINGOS, S. D. O.; OLIVEIRA, J. F. D.; MATTOS NETO, P. S. G. An intelligent hybridization of arima with machine learning models for time series forecasting. *Knowledge- Based Systems*, Netherlands, v. 175, p. 72–86, 2019.
- FERNÁNDEZ-CARA, E.; PROUVÉE, L. Optimal control of mathematical models for the radiotherapy of gliomas: the scalar case. *Computational and Applied Mathematics*, Petrópolis, v. 37, p. 745–762, 2018.
- HOLLAND, E. C. Glioblastoma multiforme: the terminator. *Proceedings of the National Academy of Sciences*, Washington, v. 97, n. 12, p. 6242–6244, 2000.
- HYNDMAN, R.; KOEHLER, A. B.; ORD, J. K.; SNYDER, R. D. Forecasting with exponential smoothing: the state space approach. Berlin: Springer, 2008.
- HYNDMAN, R. J.; AKRAM, M.; ARCHIBALD, B. C. The admissible parameter space for exponential smoothing models. *Annals of the Institute of Statistical Mathematics*, Tokyo, v. 60, n. 2, p. 407–426, 2008.
- HYNDMAN, R. J.; KOEHLER, A. B.; SNYDER, R. D.; GROSE, S. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, Amsterdam, v. 18, n. 3, p. 439–454, 2002.
- IARC - INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. Latest global cancer data: Cancer burden rises to 18.1 million new cases and 9.6 million cancer deaths in 2018. França: IARC, 2018. Available from: <https://www.iarc.fr/>. Access in: dec. 2018.
- INCA - INSTITUTO NACIONAL DE CÂNCER. Estimativas 2018: Incidência de câncer no Brasil. Rio de Janeiro: INCA, 2018. Available from: <https://www2.inca.gov.br>. Access in: jun. 2018.
- JESUS, J. C. D.; CHRISTO, E. S.; GARCIA, V. S.; ALVAREZ, G. B. Time series analysis for modeling of glioma growth in response to radiotherapy. *IEEE Latin America Transactions*, Canadá, v. 14, n. 3, p. 1532–1537, 2014.
- LEDER, K. *et al.* Mathematical modeling of PDGF-driven glioblastoma reveals optimized radiation dosing schedules. *Cell*, Cambridge, v. 156, n. 3, p. 603–616, 2014.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. C. *Introduction to time series analysis and forecasting*. 2th ed. New Jersey: John Wiley & Sons, 2015.
- MORETTIN, P. A.; TOLOI, C. M. C. *Modelos para previsão de séries temporais*. Rio de Janeiro: IMPA, 1981. v. 1.
- MURRAY, J. D. Vignettes from the field of mathematical biology: the application of mathematics to biology and medicine. *Interface Focus*, London, v. 2, n. 4, p. 397–406, 2012.
- ROCKNE, R.; ALVORD, E. C.; ROCKHILL, J. K.; SWANSON, K. R. A mathematical model for brain tumor response to radiation therapy. *Journal of Mathematical Biology*, New York, v. 58, n. 4/5, p. 561–578, 2008.
- SILVA, J. J. da; ALVAREZ, G. B.; GARCIA, V. S.; LOBÃO, D. C. Modelagem computacional do crescimento de glioma via diferenças finitas em resposta à radioterapia. In: CONGRESSO NACIONAL DE MATEMÁTICA APLICADA E COMPUTACIONAL, 36, 2016, Gramado. *Proceedings [...]*. Gramado: Sbmact Publishing, 2016. p. 1-7.
- SOUZA, E. B.; NEVES, T. A.; ALVAREZ, G. B. Estudo sobre otimização da radioterapia em pacientes com gliomas. In: SIMPOSIO BRASILEIRO DE PESQUISA OPERACIONAL, 47., 2015, Porto de Galinhas. *Proceedings [...]*. Porto de Galinhas: Sbpoc Publishing, 2015. p. 1-12.
- SWANSON, K. R.; ALVORD JUNIOR, E. C.; MURRAY, J. D. A quantitative model for differential motility of gliomas in gray and white matter. Cell proliferation, Cambridge, v. 33, n. 5, p. 317–329, 2000.
- TEIXEIRA, J. G.; IKEHARA, H. C.; BRANDT, S. A. Estrutura de oferta e desenvolvimento agrícola: um modelo de equações recursivas. *Semina: Ciências Exatas e Tecnológicas*, Londrina, v. 2, n. 8, p. 183–188, 1981.
- WALKER, M. D.; STRIKE, T. A.; SHELINE, G. E. An analysis of dose-effect relationship in the radiotherapy of malignant gliomas. *International Journal of Radiation Oncology, Biology, Physics*, New York, v. 5, n. 10, p. 1725–1731, 1979.
- ZHANG, G. P. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, Amsterdam, v. 50, p. 159–175, 2003.

Received: Nov. 24, 2020
 Accepted: Dec. 23, 2020
 Published: Apr. 9, 2021