Merging Traditional Scientific Computing with Data Science to Develop a New Class of Prediction Engines

Gbocho Masato Terasaki

3rd year Ph.D. Student in Applied Mathematics

UC Merced

December 7^{th}, 2022

A Capstone submitted in partial satisfaction of the requirements for the degree of Master of Science in Applied Mathematics

This is to certify that I have examined a copy of a technical report by

Gbocho Masato Terasaki

and found it satisfactory in all respects, and that any and all revisions required by the examining committee have been made.

Applied Mathematics
Graduate Studies Chair:

_____
Professor Maxime Theillard

Committee Chair / Research Advisor:

_____
Professor Maxime Theillard

Committee Co-Chair / Research Advisor:

_____
Professor Erica Rutter

_____
Date

Contents

Mathematical Modeling of Brain Tumor Growth

1 Introduction

1.1 Background

Glioblastoma Multiforme (GBM) is one of the fastest-growing and is the most common form of brain tumors [13]. It has very aggressive and invasive properties which lead to a median survival of two/three months, if no treatment is performed [4]. Currently, the standard of care for treating GBM is resection followed by radiation and chemotherapy [5]. However, depending on the localization of the tumor, performing surgery might be impossible. Furthermore, surgery does not guarantee complete removal of the tumor cells. With their very diffusive behavior, the remaining tumor cells regrow rapidly [4]. In addition, traditional oral drugs are ineffective as they target tumor cells using bloodstream and GBM occur in the Central nervous system. Moreover, the blood brain barrier represents a physical obstacle through which medication is unable to effectively diffuse through. Consequently, even with available treatments, the mean survival time is less than 15 months [10].

Nowadays, we have sophisticated imaging techniques such as Magnetic Resonance Imaging (MRI) that allow doctors to accurately identify the location and characteristics of brain tumors during preoperative examination as well as well post-treatment management. Basic MRI modalities such as T1-weighted (T1w) and contrast-enhanced (T1CE), T2-weighted (T2w), and T2-fluid-attenuated inversion recovery (T2-FLAIR) sequences, are usually performed at initial diagnosis of GBMs. For instance, T1 MRI usually shows a necrotic region surrounded by an enhancing rim that correlates with high blood vessel density; T2 MRI shows an outer rim corresponding to peripheral/hyper-vascularized region. Hence, from an MRI analysis, it shows that the GBM mass can be divided into three subregions: the inner core (necrotic core), the intermediate layer (enhancing rim), and the peripheral/hyper-vascularized region (edema).

Although we are able to capture the size of the tumor mass at a precise time with standard imaging methods, it is impossible to monitor the entire growth process of GBM cells in a human brain. Additionally, MRI has a limited ability to accurately characterize tumor masses anatomically. Hence, despite being able to exactly pinpoint where the tumor is located, we cannot accurately determine how high the tumor density is [6] . In that regard, different biological models such as *in vivo* models have been explored to better understand the interaction of GBM cells and brain environments. For instance, Rutter *et al.* investigated GBM cell evolution in a Murine environment [12]. Although this approach gives us great insights on the proliferative/diffusive characteristics of gliomas, it does not necessarily translate to what would happen in a human brain since most mice used in such experiments are immunocompromised. On top of that, these experiments come at the detriment of the animal's life.

Thus, despite all the substantial advantages of imaging methods or biological experiments, they are incapable of reliably predict tumor growth in a human brain. Therefore, leaning toward mathematical models might provide better insights on the evolution of GBM cells in a human brain in hopes of discovering effective therapy.
The Kolmogorov–Petrovsky–Piskunov equation, commonly known as the Fisher's equation, is a type of reaction-diffusion equations often used to describe the dynamics of population, in particular to study invasive behavior of growing tumor cells:

$$\frac{\partial C}{\partial t}(\mathbf{x},t) = \underbrace{\nabla \cdot (D\nabla C(\mathbf{x},t))}_{\text{diffusion}} + \underbrace{\rho C(\mathbf{x},t)(1-C(\mathbf{x},t))}_{\text{growth}} \qquad \forall \mathbf{x} \in \mathcal{B} \tag{1}$$

where C indicates the normalized tumor cell density ($0 \leq C(x,t) \leq 1$). We denote $\mathcal{B}$ the brain geometry. In this model, we capture the diffusive characteristic of the GBM by the diffusion parameter D ($mm^2 day^{-1}$), and the proliferative behavior by the growth parameter ρ (day^{-1}).

Tracqui *et al.* were the one of the first mathematicians to model GBM growth [15]. Based on the analysis of serial CT scans taken from a single GBM cancer patient (treated with chemotherapy) for one year, they attempted to estimate rates of net proliferation ρ and net diffusion D[15]. Woodward *et al.* used these values as averages to model the effect of resection on GBM evolution from different hypothetical patients [16] and compared with real data obtained from Kreth *et al.* [7, 14].
On the other hand, Swanson and her collaborators have sought to develop a patient specific mathematical model and to estimate patient specific parameter values of D and ρ. Their data consisted of successive Magnetic Resonance (MR) images taken from 70 adult patients with previously untreated glioblastoma and received X-irradiation as well as chemotherapy [14]. Therefore, the ability to estimate these (ideally patient-specific) model parameters as accurately as possible can give a better understanding of the GBM evolution and, thereby, benefit the optimization of therapy.

Yet, most data used in those works included treatment, and truly understanding the growth of GBM in our brain is unfeasible with such data. Conversely, obtaining successive MR (at least 3 time points) images of patients with untreated glioblastoma is quasi-impossible because, due to to the invasive properties of GBM, immediate care will be advised by experts to (try to) extend survival [5]. Most data available are at most two snapshots, since experts usually take at most two MR scans pre-treatment: one at the initial diagnosis, one right before treatment). Given the limited accessibility to MRI scans at multiple time steps, we are unable to learn these parameters for untreated patients using standard learning techniques such as optimization routines (best least-squared fit).

Our goal is to develop a computational pipeline that will allow us to first estimate patient specific parameters from a single MRI snapshot, and then theorize a prognostic using our constructed forward simulator. We aspire that this framework could be applied to help improve real prognostic, and inform clinicians to select optimal treatments.

Our equation model (Equation 1) assumes homogeneity between all cells and leads to a dense tumor mass with an advancing front [3]. It does not explicitly take in account the interaction of the necrotic core, and the surrounding live cells that can contribute to the proliferation and migration

of the tumor cells. Additionally, some have found that the significance of accumulation of fluid resulting from disruption of the blood-brain barrier (edema) might be a prognostic indicator of patient survival [11].

1.2 Computational Pipeline

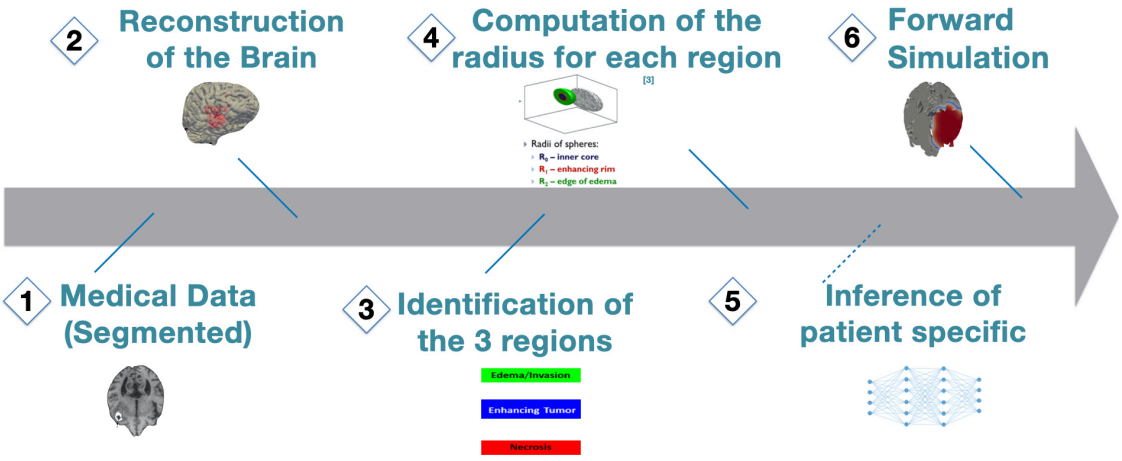

Figure 1: Computational Pipeline

We thereby constructed the following computational pipeline (Figure 1).
In step (1), we obtained MR images for a patient acquired at a single time point with the brain geometry and the tumors segmented [8, 1, 2]. Then, in step (2), using a level set function, we reconstruct the brain geometry. We proceed in this step to reconstruct the tumor geometry as well as its location in the given brain. This reconstruction allows us to identify the three main regions of the GBM: necrotic core, enhancing rim, and edema in step (3). Then (4), using the assumption that all GBM mass are spherical, we calculate the radius of each of the regions. We proceed in step (5) to construct a model that learn patient specific parameters necessary in Equation (1) using the computed radii. Finally (6), we feed the learned parameters into our constructed forward simulator and monitor the GBM growth over $T = 60$ days, which is the mean survival time for GBM cancer without any treatment received [5].

In the next section, we illustrate how we reconstruct a realistic brain geometry from the obtained medical data. In Section 3, we explain how retrieve the geometrical properties of the tumor subregions from the reconstruction and describe the model to infer patient-specific parameters. In Section 4, we delineate our constructed forward simulator and use the inferred model parameters to simulate tumor growth. This is followed by the Results and Future Work Section.

2 From the Medical data to the reconstruction of a realistic brain geometry

We obtained MR images of Brain scans of patients diagnosed with gliomas from the BRATS datasets [8, 1, 2]. These images were taken at a single time point. Assuming all brains are similar to an

extent, we subtract a threshold intensity from the raw data to remove any background noise. Then, using a level set function, we reconstruct the whole brain geometry. The level-set approach favors the construction of a non-pixelated brain geometry. The raw data are given in 2 dimension. We can view the 3 dimensional reconstruction as the superposition of all 2 dimensional slices.
In this dataset, we also get segmented data of the brain tumor. The segmentation is done accordingly to heterogeneous histological sub-regions of gliomas, i.e., peritumoral edematous/invaded tissue, necrotic core, active and non-enhancing core. Hence, we identify the different tumor regions from the segmented data and we construct the corresponding level set representation to reconstruct the tumor geometry as well as its location in the given brain. This reconstruction allows us to identify the three main regions of the GBM: necrotic core, enhancing rim, and edema. Below, we can see how our reconstruction matches with the 2 dimensional segmented data (Figure 2, Figure 3), as well as the whole brain and tumor reconstruction (Figure 4).

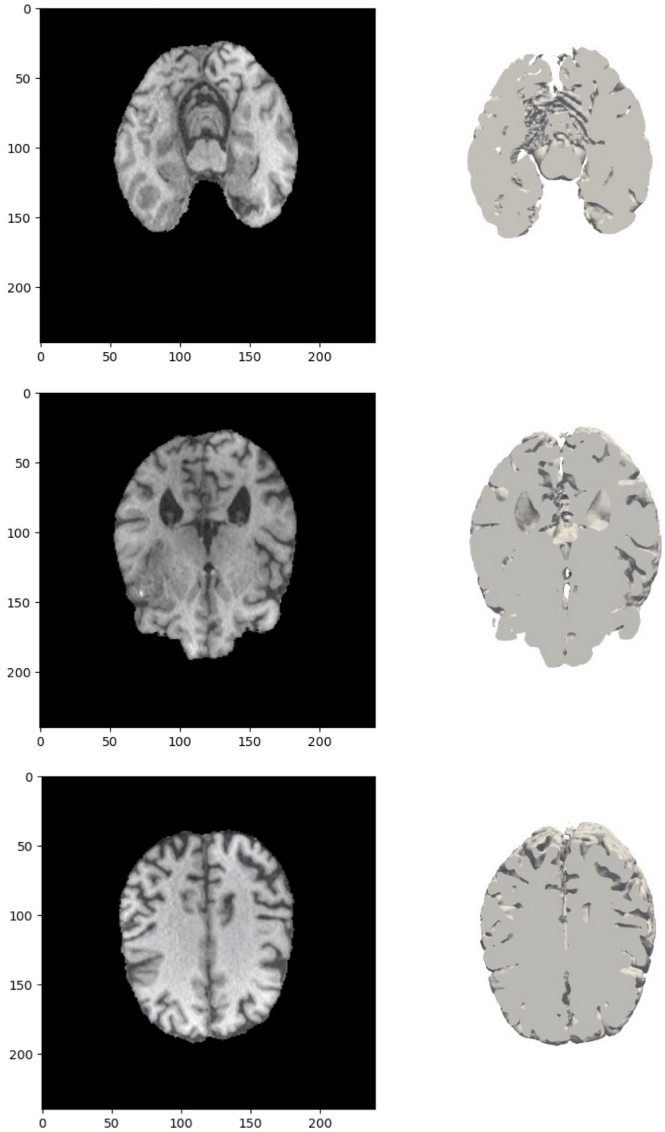

Figure 2: Real Brain MR image data (left) reconstructed Brain in 2D (right)

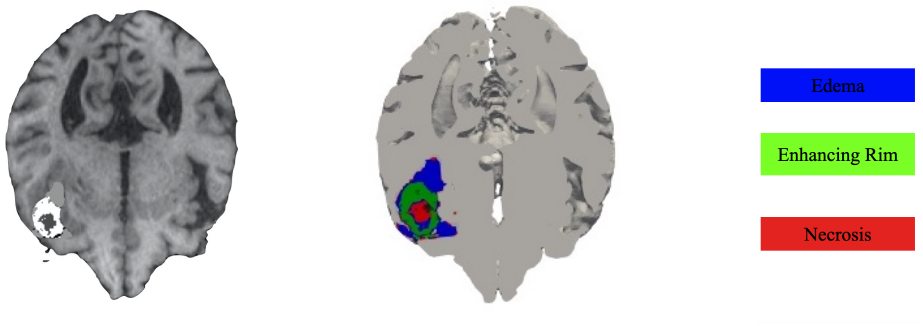

Figure 3: Real Brain/Tumor MR image data (left) reconstructed tumor in 2D (right)

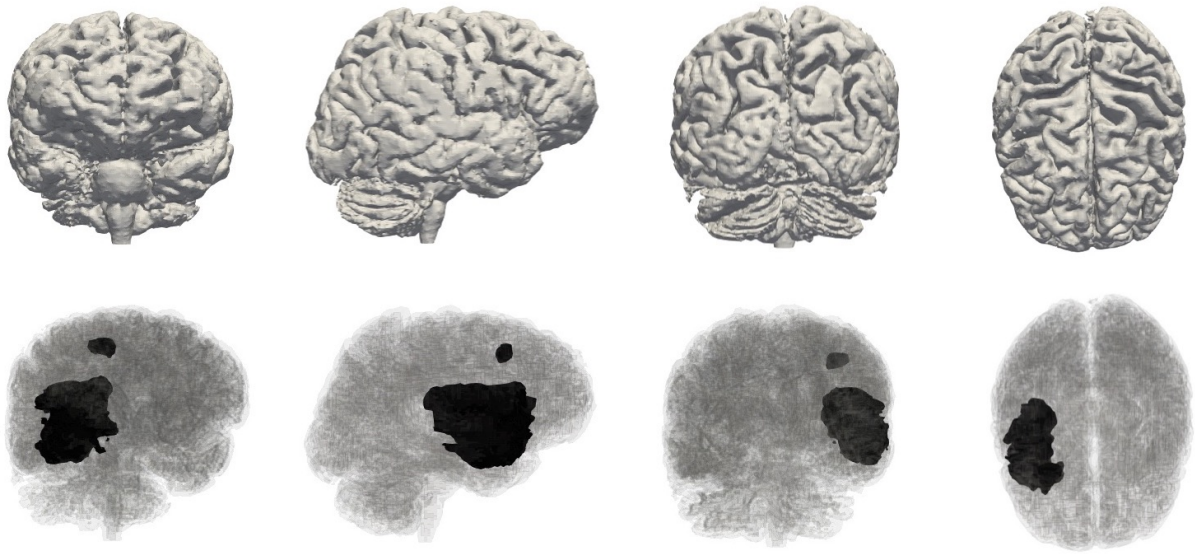

Figure 4: Whole Brain with tumor Reconstruction

3 From the reconstruction of the tumor to the inference of the patient-specific parameters

It has been recognized that GBMs morphologically typically appear (at least at initial diagnosis) as roughly spherical. Consequently, we approximate each of the three regions (necrotic core, enhancing rim, edema) of the tumor as a sphere. We hypothesize that tracking the radii over a given time will yield insights into the relative contributions of cellular proliferation, motility, necrosis, and edema to the observed image features, and thereby have a better understanding of GBM growth. Han *et al.* estimated patient specific parameters from a single snapshot via traveling wave analysis, but a reduced model of Equation (1) onto 1 dimension [3]. Additionally, their analysis was sensitive to parameter values, sometimes providing imaginary numbers, reproducing their same results numerically was quite arduous. Yet, this motivated us to explore possibilities to learn these parameters using these radii obtain from a single snapshot.

In Figure 5 and Figure 6, we can see that each subregion of the GBM can be approximated spherically. Therefore, we calculate numerically the volume of each region. To do so, we assign a density of 1 to all discritized cells and integrate it over the whole brain domain. We find each radius using the volume of sphere formula (*i.e.* $r_i = \sqrt[3]{\frac{3}{4\pi}V_i}$).

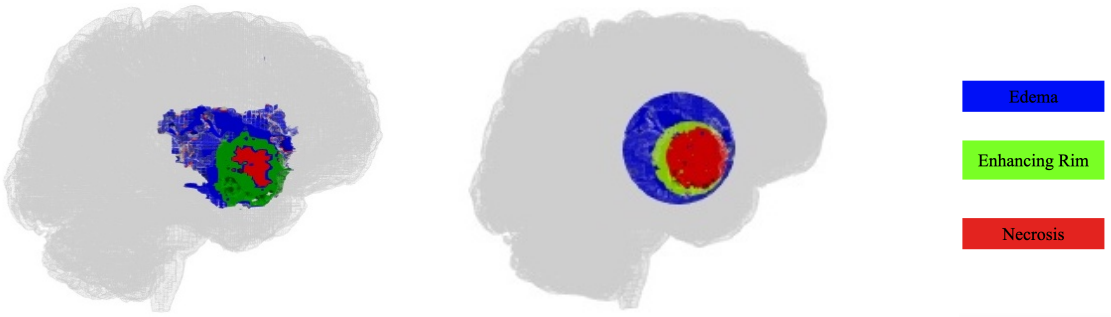

Figure 5: Real Tumor (left) Spherical tumor (right)

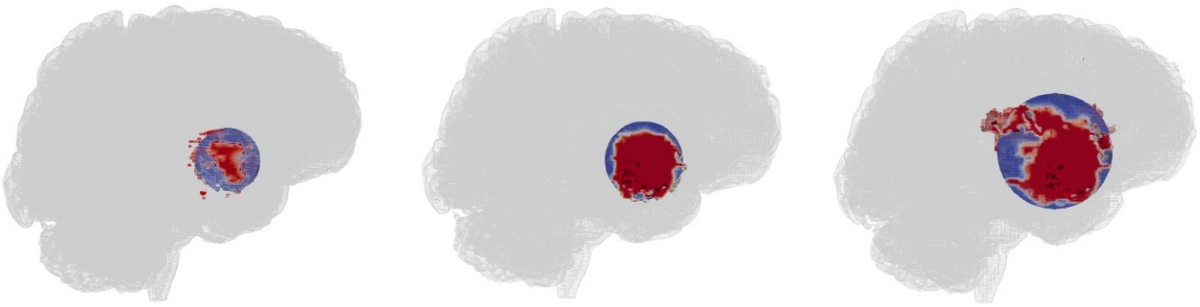

Figure 6: Necrotic Core (left), Enhancing Rim (middle), Edema (right)

We find $R_0 = 12.90mm$**,** $R_1 = 21.15mm$**,** $R_2 = 28.35mm$.

Table 1 represents the patient-specific parameters D and ρ obtained by Han *et al.* given a patient's tumor information ($R0, R1, R2$ in mm) [3]. We opted to first investigate a native model, a linear regression model based on the results provided by Han *et al.* (see Table 1). Table 1 also shows the predicted parameters $\hat{D}$ and $\hat{\rho}$ from our constructed linear regression model.

Patient	$R0$	$R1$	$R2$	D	ρ	$\hat{D}$	$\hat{\rho}$
1	14.87	20.73	27.77	0.2852	0.2102	0.4332	0.2068
2	20.48	26.34	38.24	1.2791	0.2624	1.1835	0.2685
3	6.61	10.91	15.24	0.0825	0.1736	0.0319	0.1769
4	22.87	26.96	37.03	0.9825	0.2590	0.9295	0.2558
5	8.17	14.20	25.10	0.9769	0.2520	1.0170	0.2491
6	8.29	15.83	20.35	0.0687	0.1652	0.0177	0.1651

Table 1: Patient-specific parameters inferred from both Han's and our model given $R0$, $R1$, and $R2$ (in mm).

We use a 90%/10% train-test ratio to construct a model for the growth parameter as well as for the Diffusion parameter D. We obtained the following linear models:

$$\rho = 5.142 \cdot 10^{-3} R_0 - 1.671 \cdot 10^{-2} R_1 + 1.209 \cdot 10^{-2} R_2 + 1.409 \cdot 10^{-1}$$
$$D = 1.473 \cdot 10^{-2} R_0 - 1.672 \cdot 10^{-1} R_1 + 1.534 \cdot 10^{-1} R_2 - 5.784 \cdot 10^{-1}$$

Therefore, using the previous computed radii, we obtained the following parameters specific for Patient A.

$$\rho = 0.1966$$
$$D = 0.4229$$

It is important to note that this investigation is a proof of concept and we need to either improve the model or find a way to increase the data size.

4 Forward simulation

Now that we have obtained the parameters, we can feed it to the forward simulator and observe tumor growth over our given $T = 60$ days. To construct the forward simulator, we discretize Equation (1) using the Finite Volume method.

4.1 Finite Volume Method

Reconsider Equation (1):

$$\frac{\partial C}{\partial t} = \underbrace{\nabla \cdot (D \nabla C)}_{\text{diffusion}} + \underbrace{\rho C(1 - C)}_{\text{growth}}$$

We discretize the diffusion term implicitly, and we linearize the reaction term:

$$\frac{C^{n+1} - C^n}{\Delta t} = D(\nabla \cdot \nabla C^{n+1}) + \rho C^{n+1} - \rho (C^n)^2.$$

Then,

$$(1 - \Delta t \rho) C^{n+1} - \Delta t D(\nabla \cdot \nabla C^{n+1}) = C^n - \Delta t \rho (C^n)^2.$$

We sub-divide the spatial domain into finite cells. Then, for a particular cell, i, we take the integral over that cell Ω_i, and get:

$$\int_{\Omega_i} (1 - \Delta t \rho) C^{n+1} - \int_{\Omega_i} \Delta t D(\nabla \cdot \nabla C^{n+1}) = \int_{\Omega_i} C^n - \Delta t \rho (C^n)^2$$

And, applying the divergence theorem to the second term yields:

$$\int_{\Omega_i} \alpha C^{n+1} - \int_{\partial \Omega_i} \lambda \nabla C^{n+1} \cdot \mathbf{n} = \int_{\Omega_i} F(C^n)$$

where[1] $\alpha = 1 - \Delta t \rho$, $\lambda = \Delta t \cdot D$ and $F(C^n) = C^n - \Delta t \rho (C^n)^2$.
A zero flux boundary condition (*i.e.* $\nabla u \cdot \mathbf{n} = 0$) is applied in this model.

We can simplify the explanation of the Finite Volume method in 2 dimension (Figure 7): we compute the average tumor density over the consider cell Ω_i, then we compute the flux through all boundaries of the cell, and finally multiply the right hand side using values obtained at the previous time-step with the cell volume.

[1] We choose our Δt to satisfy the condition that $\alpha > 0$ for the method to be stable.

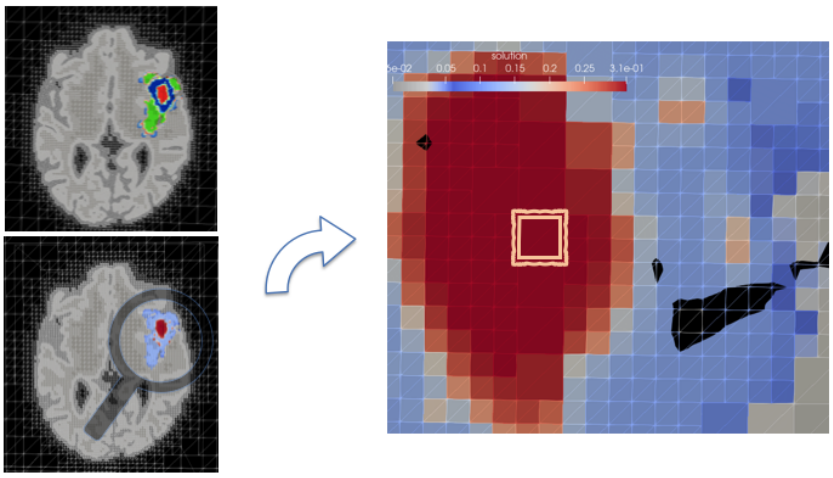

Figure 7: Finite Volume Method in 2D.

4.2 Results

Using the computed parameters (D, ρ) for the patient we considered, we simulated tumor growth over $T = 60$ days. Hence, we build a numerical method to solve the equation above with $\rho = 0.20$, $D = 0.43$, $\Delta t = T/N$, $(n = 300)$, and the following initial conditions:

$$\begin{aligned} u_0 &= 0.5 \quad \text{(necrotic core)} \\ u_0 &= 0.3 \quad \text{(enhancing rim)} \\ u_0 &= 0.1 \quad \text{(edema)} \end{aligned}$$

We obtain the following results (see Figure 8). We can see the tumors growing and proliferating in the brain as time passes. We slice the brain geometry across the three coordinate planes (xy-plane, yz-plane, and xz-plane accordingly) and we show snapshots of the tumor evolution at $T = 0, 20, 40$, and 60 days. At the final time $T = 60$ days, we measure the spherical equivalent necrotic tumour volume and find a radius of $R_0 =$. If we consider that death occurs when the spherical equivalent tumour volume is typically of radius is 30 mm [9]. Hence, this simulation seems to agree with the expected survival time of a untreated GBM.

5 Conclusion

We developed a computational pipeline to simulate a GBM evolution that is specific to a given patient. From their single pre-operative MR scan, we aimed to estimate the model parameters to simulate the tumor growth in hopes to predict a prognostic.
It would have been interesting if there were data available at multiple time points for an untreated patient so that we can compare the simulated growth and the actual growth. One way to conduct such comparison might be to consider experiments done on mice as various data are available in this framework. Also, the linear regression constructed in Section 3 is a proof of concept to learn the model parameters from a single MRI snapshot. As discussed, the data considered for this model is very small. Since accessibility to data in this framework is limited, we opted to create a hypothetical dataset using our numerical solver.

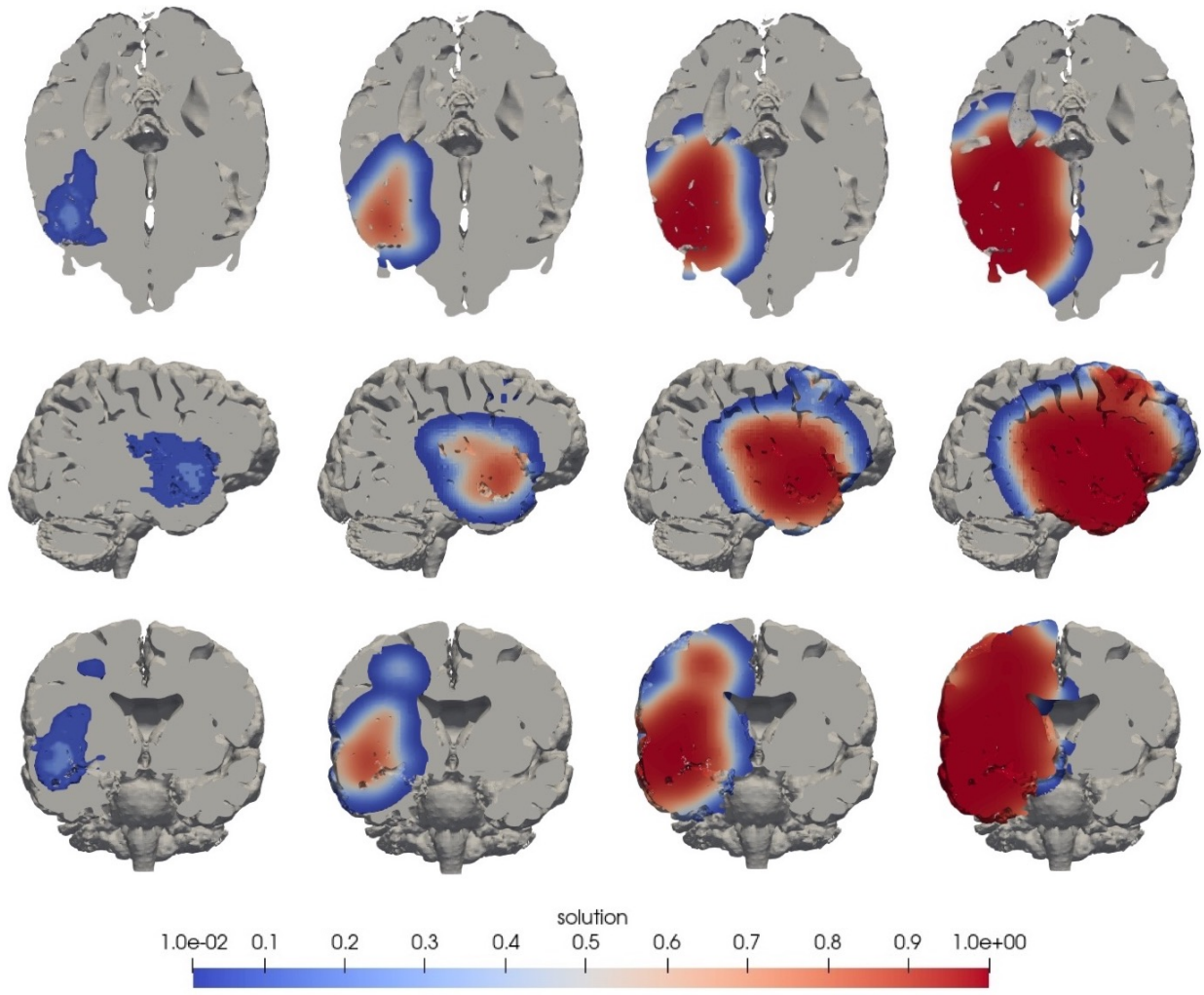

Figure 8: Forward Simulation for $T = 0, 20, 40, 60$ days.

6 Futher exploration: Inference of patient specific parameters using virtual brains

Let us consider the same domain, initial condition, and $\Delta t = T/N = (60/300)$.
We define

$$DD = \{D = 0.05 + i \times \frac{1.3 - 0.05}{n_d} \quad \text{for} \quad i = 0, 1, ..., 20\}$$
$$RR = \{\rho = 0.1 + i \times \frac{0.6 - 0.1}{n_r} \quad \text{for} \quad i = 0, 1, ..., 10\}$$

For each parameter pair $(D, \rho) \in DD \times RR$, we run the forward simulation and compute the radius of the GBM subregions at each iteration. This gives us a collection of single-snapshot data with associated model parameters.

We, thereby, aim to construct a new linear regression using these hypothetical data. Since we have $(n_d+1) \times (n_r+1)$ possible pairs, if we choose $n_d = 20$ and $n_r = 10$, we obtain 69300 snapshots

in total.
Thus, we construct a new linear regression model using these hypothetical data. We use a 80%/20% train-test ratio for the new (updated) linear regression model for both D and ρ.
Hence, given the radii computed for the brain we considered in Section 3 , we obtain a new estimate for the model parameters:

$$\rho = 0.2661$$
$$D = 0.3530$$

There is still a lot to be investigated and improved in our model. However, we have succeeded in building a bridge between our simulation environment (C++) and our parameter learning environment (python). This will facilitate the exploration of various models of step 5 (Figure 1) in our future work.

Bibliography

[1] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data*, 4(1):1–13, 2017.

[2] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. T. Shinohara, C. Berger, S. M. Ha, M. Rozycki, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint arXiv:1811.02629*, 2018.

[3] L. Han, S. Eikenberry, C. He, L. Johnson, M. C. Preul, E. J. Kostelich, and Y. Kuang. Patient-specific parameter estimates of glioblastoma multiforme growth dynamics from a model with explicit birth and death rates. *Mathematical biosciences and engineering: MBE*, 16(5):5307, 2019.

[4] G. Iacob and E. B. Dinca. Current data and strategy in glioblastoma multiforme. *Journal of medicine and life*, 2(4):386, 2009.

[5] T. P. Jigisha, P. P. Pier, and P. C. Vikram. Glioblastoma multiforme. *American Association of Neurological Surgeons*.

[6] M. J. Kransdorf and M. D. Murphey. Radiologic evaluation of soft-tissue masses: a current perspective. *American Journal of Roentgenology*, 175(3):575–587, 2000.

[7] F. W. Kreth, P. C. Warnke, R. Scheremet, and C. B. Ostertag. Surgical resection and radiation therapy versus biopsy and radiation therapy in the treatment of glioblastoma multiforme. *Journal of neurosurgery*, 78(5):762–766, 1993.

[8] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2014.

[9] J. Murray. Glioblastoma brain tumours: Estimating the time from brain tumour initiation and resolution of a patient survival anomaly after similar treatment protocols. *Journal of biological dynamics*, 6(sup2):118–127, 2012.

[10] A. D. Norden and P. Y. Wen. Glioma therapy in adults. *The neurologist*, 12(6):279–292, 2006.

[11] W. B. Pope, J. Sayre, A. Perlina, J. P. Villablanca, P. S. Mischel, and T. F. Cloughesy. Mr imaging correlates of survival in patients with high-grade gliomas. *American Journal of Neuroradiology*, 26(10):2466–2474, 2005.

[12] E. M. Rutter, T. L. Stepien, B. J. Anderies, J. D. Plasencia, E. C. Woolf, A. C. Scheck, G. H. Turner, Q. Liu, D. Frakes, V. Kodibagkar, et al. Mathematical analysis of glioma growth in a murine model. *Scientific reports*, 7(1):1–16, 2017.

[13] S. Sathornsumetee, J. N. Rich, and D. A. Reardon. Diagnosis and treatment of high-grade astrocytoma. *Neurologic clinics*, 25(4):1111–1139, 2007.

[14] K. Swanson, R. Rostomily, and E. Alvord. A mathematical modelling tool for predicting survival of individual patients following resection of glioblastoma: a proof of principle. *British journal of cancer*, 98(1):113–119, 2008.

[15] P. Tracqui, G. Cruywagen, D. Woodward, G. Bartoo, J. Murray, and E. Alvord Jr. A mathematical model of glioma growth: the effect of chemotherapy on spatio-temporal growth. *Cell proliferation*, 28(1):17–31, 1995.

[16] D. i. w. Woodward, J. Cook, P. Tracqui, G. Cruywagen, J. Murray, and E. Alvord Jr. A mathematical model of glioma growth: the effect of extent of surgical resection. *Cell proliferation*, 29(6):269–288, 1996.