CIMAT

Centro de Investigación en Matemáticas, A.C.

A GLOBAL SENSITIVITY INFORMED PARAMETER SELECTION METHOD APPLIED TO A MODEL OF PROSTATE CANCER UNDER INTERMITTENT ANDROGEN SUPPRESSION

T E S I S

Que para obtener el grado de

Maestra en Ciencias

con Orientación en

Matemáticas Aplicadas

Presenta

Laura Figueroa Resa

Director de Tesis:

Dr. Marcos Aurelio Capistrán Ocampo

Autorización de la versión final

Guanajuato, Gto., 5 de noviembre de 2021

Dedication

To my parents and my sister.

Acknowledgements

I would like to thank my advisor, Prof. Marcos Aurelio Capistrán Ocampo, for his guidance and support throughout the development of this thesis. Also, I want to thank Dr. Felipe Medina Aguayo for handling me his notes about "Importance Sampling". Finally, I want to thank CONACyT (Consejo Nacional de Ciencia y Tecnología) for giving me financial support during the entire duration of the Master of Science program offered by CIMAT (Centro de Investigación en Matemáticas).

Abstract

It is known that mathematical models of biological systems often have many parameters, and choosing values for those parameters may be a limitation in practical applications. Indeed, experimental measurement of the model parameters is difficult, and naive parameter estimation leads to parameter identifiability issues. In this thesis we introduce a method for parameter selection informed by global sensitivity analysis and data. We apply these techniques to a model of prostate cancer under intermittent androgen suppression therapy. First, we rank the model parameters according to their contribution to the likelihood variance. We rank one parameter at a time to deal with the likelihood multiple scaling. Second, we use Bayesian model selection to compare the proposed model parametrization with a reference from the literature. We use synthetic data to present our findings and offer a discussion on further applications of our method.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

When putting together mathematical models, we would like to explain and predict a system's behavior. A model has a set of parameters, and often we do not have the tools to know all of them, so we have to estimate some parameters from observations of the system. Mathematical models of biological systems frequently have multiple scales as mentioned in Gutenkunst *et al.* [2], "*...The collected models encompass a diverse range of biological systems, including circadian rhythm, metabolism, and signaling. All the models are formulated as systems of ordinary differential equations, and they range from having about ten to more than 200 parameters...*" and "*...every model we examine exhibits a sloppy parameter sensitivity spectrum...*", so, in general, it is difficult to estimate all the parameters within a system; thus, we have the task of choosing a certain subset of parameters to infer and other subset of parameters to postulate.

The goal of this work is to study the parametrization of mathematical models describing the dynamics of Prostate Specific Antigen (PSA) under Intermittent Androgen Suppression (IAS). We regard model T-5 in [3], henceforth referred to as $\mathcal{M}2$, as a reference model. $\mathcal{M}2$ is a simplification of the state-of-the-art model by Baez and Kuang [4]. Our contribution consists in proposing an alternate parametrization of the $\mathcal{M}2$ model based on global sensitivity analysis, henceforth referred to as model $\mathcal{M}1$, and comparing these models using formal inference and statistical model selection.

We apply the Sobol sensitivity analysis on the model T (mentioned in [3]) parameters, and then compute Markov Chain Monte Carlo (MCMC) for both $\mathcal{M}1$ and $\mathcal{M}2$ models, making assumptions for prior distributions and likelihood in order to get samples to be used to estimate the Bayes Factor between these two models. Notice that, our strategy could be used in other models to do data driven model development.

The theory behind Sobol's sensitivity analysis can be found in [5]. The results of the Sobol's sensitivity analysis can be found in Section 3.2. The theory behind MCMC and the Bayes Factor can be found in [6]. The theory behind assuming a Gaussian for the posterior distribution can be found in the notes [7]. In 3.6, some graphics are presented to visualize some projections of the parameters distributions, and PSA predictions using the corresponding posterior distribution for each model.

1.1 Motivation

Worldwide, the second most frequent type of Cancer among men is Prostate Cancer (PCa).

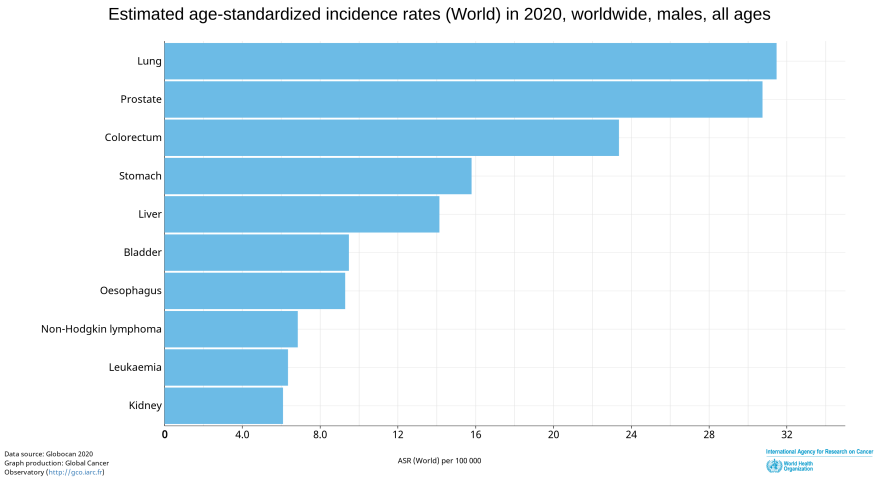

Figure 1.1: Worldwide Cancer Incidence Rates [1]

Since Swanson *et al.* (2001) [8], multiple mathematical models arose

describing different types of PCa.

In 1941, Huggins and Hodges [9] showed that PCa is influenced by androgen –which are male hormones. In addition, they showed that metastatic PCa can be inhibited as androgen is eliminated –either by a surgical method like a bilateral orchiectomy or by estrogen injections. On the other hand, androgen injections contribute to PCa development.

At this point one may think we can control the proliferation of prostate cancer cells by suppressing androgen, but these cells might become androgen-resistant after some time, that is, in spite of the androgen suppression, these cells can proliferate again.

Androgen suppression has been implemented mainly as a treatment for localized or locally advanced PCa. Generally, the androgen is continuously suppressed through time via hormone injections until the malignant cells become hormone-resistant. Some drawbacks of the Continuous Androgen Suppression (CAS) therapy are its side effects like osteoporosis and erectile dysfunction. In the aim of coping with those side effects, there is the Intermittent Androgen Suppression (IAS) therapy. To explain what IAS therapy consist of, let us introduce what is the Prostate Specific Antigen (PSA).

PSA is a biomarker for PCa which is measured in serum. It is worth noting that, formally speaking, the PSA is not an antigen but a protein synthesized by prostate cells, i.e., a proteolytic enzyme. High PSA levels indicate a high probability of PCa whereas low PSA levels indicate a low probability of PCa. IAS therapy goal is to suppress the patient androgen until the PSA reaches a certain inferior threshold where the treatment is suspend, so the patient can take a break from the side-effects; then the PSA starts rising again and when the PSA reaches a superior threshold, the treatment begins again and so on (that is, IAS therapy has a cyclic behaviour) until cancer cells become androgen resistant. As PSA is almost the only observable variable (except maybe the androgen), we use PSA time series for parameter estimation.

1.2 Related work

For PCa, there are several mathematical models that describe the volume or shape of the prostate or the tumor, but we only present models that explicitly describe PSA dynamics.

- In 2001, Swanson *et al.* [8] described the serum PSA dynamics as:

$$\frac{dP}{dt} = \beta_h V_h + \beta_c V_c - \gamma P,$$

  where V_h is the volume of prostate benign cells and V_c is the volume of prostate cancer cells.

- In 2002, Vollmer *et al.* [10] used a log-linear regression to fit PSA data. The equation is:

$$\log(y) = a + bt,$$

  where "*...a is the intercept and reflects the overall amplitude of PSA...*" and b is the relative velocity of PSA with respect to the PSA level, that is, $b = \frac{dy/dt}{y}$.

- In 2003, Vollmer and Humphrey [11] made the distinction between PSA in the prostate tissue milieu, f, and serum PSA, y:

$$\frac{df}{dy} = \alpha - \beta f,$$
$$\frac{dy}{dt} = \beta f \frac{V_p}{V_s} - ky,$$

  where V_p is the prostate volume and V_s is the serum volume.

- In 2008, Ideta *et al.* [12] proposed a deterministic model where they considered the dynamics of PSA taking into account that prostate cancer cells can be androgen-dependent (AD), x_1, or androgen-independent (AI), x_2:

$$P(t) = c_1 x_1(t) + c_2 x_2(t).$$

- In 2010, Tanaka *et al.* [13] considered the model by Ideta *et al.* [12] but modified it as a stochastic model:

$$P(t) = x_1(t) + x_2(t) + \xi_P,$$

where ξ_P is Gaussian white noise.

- In 2010, Hirata *et al.* [14] proposed a piecewise linear model considering three types of prostate cancer cells: androgen-dependent (x_1), androgen-independent, that can become androgen-dependent at some point (x_2), and androgen-independent, which can not be androgen-dependent ever again (x_3). Thus the PSA equation is:

$$P(t) = c_1 x_1(t) + c_2 x_2(t) + c_3 x_3(t).$$

- In 2011, Jain *et al.* [15] made a distinction between PSA in tissue and PSA in serum which are represented by a system of coupled ordinary differential equations (ODEs).

  PSA production in tissue:

$$\frac{dP}{dt} = -\alpha_P P E_{\text{frac}} - \gamma_P P \frac{(N+M)^2}{K_P + N + M} - \lambda_P P + \text{prod}_P,$$

  where E stands for the number of prostatic epithelial cells, N is the number of androgen-dependent prostate cancer cells, and M the androgen-independent prostate cancer cells.

  The rate of change of serum PSA concentration is given by:

$$\frac{dP_S}{dt} = \alpha_P P E_{\text{frac}} + \gamma_P P \frac{(N+M)^2}{K_P + N + M} - \lambda_{P_s} P_S.$$

- In 2012, Portz *et al.* [16] took up Ideta *et al.* [12] model but added a cell quota for androgen –for AD cells and AI cells.

$$\frac{dP}{dt} = \sigma_0 (X_1 + X_2) + \sigma_1 X_1 \frac{Q_1^m}{Q_1^m + \rho_1^m} + \sigma_2 X_2 \frac{Q_2^m}{Q_2^m + \rho_2^m} - \delta P.$$

- In 2016, Baez and Kuang [4] compared two models: one with the assumption that there exist only AD cells and another one assuming there are both AD and AI cells.

  A single type of cell population model:

$$\frac{dP}{dt} = bQ + \sigma xQ - \epsilon P.$$

  A two type of cells population model:

$$\frac{dP}{dt} = bQ + \sigma(Qx_1 + Qx_2) - \epsilon P.$$

- In 2019, Draghi *et al.* [17] developed a coupled system of equations which makes the distinction between PSA produced by AD cells, t_{d}, and PSA produced by AI cells, t_{i}.

$$\begin{cases} \dot{p} = \rho_p p(1-p) - \alpha_{pa}(1-a)p - \delta_{pB} pB(t) \\ \dot{a} = \rho_a a(1-a) - \delta_{\mathrm{aL}}\, a\, \mathrm{L}(t) \\ \dot{t}_{\mathrm{d}} = \rho_{\mathrm{d}} t_{\mathrm{d}} - \left[\delta_{\mathrm{d}a}(a) + \delta_{\mathrm{dB}}\mathrm{B}(t)\right] t_{\mathrm{d}}^2 - \alpha_{\mathrm{dp}} t_{\mathrm{d}} p - \epsilon(1-a)t_{\mathrm{d}} \\ t_{\mathrm{i}} = \rho_{\mathrm{i}} t_{\mathrm{i}} - \alpha_{\mathrm{i}p} t_{\mathrm{i}} p + \epsilon(1-a)t_{\mathrm{d}} \end{cases}$$

- In 2019, Phan *et al.* [18] considered the two population model from Baez and Kuang [4] but added an equation for serum androgen.

- In 2019, Phan *et al.* [19] considered the three population model proposed by Hirata *et al.* [14].

$$\frac{dP}{dt} = bQ + \sigma(Qx_1 + Qx_2 + Qx_3) - \epsilon P.$$

- In 2019, Nakanishi and Hirata [20] considered the model by Hirata *et al.* [14].

- Finally, in 2019, Wu *et al.* [3] considered the two population model by

Baez and Kuang [4]. **This thesis is based on this model**.

$$\frac{dP}{dt} = bQ + \sigma(Qx_1 + Qx_2) - \epsilon P.$$

1.3 Contributions

Given a model defined by an initial value problem for a system of ordinary differential equations, and a set of observations depending on the state variables, our method ranks the model parameters, one at a time, according to its contribution to the likelihood variance. Next we parametrize the corresponding statistical model using a prescribed number of leading parameters. We use the model T from Wu *et al.* [3] as a reference to show that our strategy renders a better statistical model. Actually, we argue that there is a relation between global sensitivity and practical identifiability depending on the data at hand.

1.4 Limitations and future work

- In this thesis we use synthetic data. We plan to use clinical data in future work.
- The models we compare have the same number of parameters. In the future, we may apply our findings to decide what is the optimal number of parameters to be inferred given a model and a set of observations.
- Also, we would like to postulate a new system of ODEs or PDEs that models PSA dynamics taking into account not only the total PSA, but making the distinction between free PSA and intact PSA, and other quantities of interest like the Gleason score, age, TNM classification, and prostate volume.

Chapter 2

Theoretical Framework

In this chapter we include a summary of theoretical results used for making our work more self-contained. First, we present the mechanistic model used to carry out our numerical experiments. Next we include a brief summary of global first order Sobol sensitivities. We pose the inverse problem as an inference problem and use the Bayesian paradigm to model the conditional probability of the kinetic parameters with given records of PSA under intermittent androgen suppression therapy. Finally we make a brief summary of Bayesian model selection.

2.1 Model

We are taking into consideration a refined version of the model by Baez and Kuang [4], presented in [3] as model T.

This is a mechanistic model based on Droop's nutrient-limiting theory [21] in the sense that cancer cells need androgen to be able to proliferate. As we mentioned in Section 1.2, this model considers two different types of cancer cells: Those androgen-dependent and those less sensitive to androgen, which are called androgen-independent. However, for both types of cells there is a minimal level of intracellular androgen for them to proliferate, which is called the cell quota (q_i). The equations governing the densities for cancer cells (x_i),

intracellular androgen (Q), serum androgen (A), and PSA (P) are:

$$\begin{aligned}
\frac{dx_1}{dt} &= \mu_1 \left(1 - \frac{q_1}{Q}\right) x_1 - (D_1(Q) + \delta_1 x_1)x_1 - \lambda(Q)x_1 \\
\frac{dx_2}{dt} &= \mu_2 \left(1 - \frac{q_2}{Q}\right) x_2 - (D_2(Q) + \delta_2 x_2)x_2 + \lambda(Q)x_1 \\
\frac{dQ}{dt} &= m(A - Q) - \frac{\mu_1(Q - q_1)x_1 + \mu_2(Q - q_2)x_2}{x_1 + x_2} \\
\frac{dA}{dt} &= \gamma_2 + \gamma_1(A_0 - A) - A_0\gamma_1 u(t) \\
\frac{dP}{dt} &= bQ + \sigma(Qx_1 + Qx_2) - \epsilon P.
\end{aligned} \tag{2.1}$$

Units of x_i, Q, A, and P are, respectively, L, nmol/L, nmol/L, and µg/L.

The androgen-dependent death and mutation rates are given by $D_j(Q) = d_j R_j/(Q + R_j)$, $j = 1, 2$ and $\lambda(Q) = cK/(Q + K)$, respectively, and $u(t)$ is a unit step function, where $u = 1$ indicates that the treatment is on, that is, the androgen is being suppressed (maybe by hormone injections), and $u = 0$ indicates that the treatment is off.

In the article [3], the following values for parameters $c = 0.00015$, $K = 1$, $\delta_1 = 5$, $\delta_2 = 5$, and $\gamma_2 = 0.005$ are fixed, because of the little effect on the model output observed in [18]. For simplicity, the authors also fixed $\mu_1 = \mu_2 = \mu_m$. Considering the sensitivity analysis carried out on [18], a 5 parameter model (μ_m, q_2, d_1, γ_1, and A_0) is considered called T-5 –henceforth referred as model $\mathcal{M}2$.

For details on biological realistic ranges see [18]. However, for simulations we used ranges plus-minus five percent from some reference values –which are the mean parameters estimated in [3] on a set of 28 patients from the Vancouver Prostate Center [22] and some fixed values in [3].

For the forward mapping, we took $u = 1$ when $0 \leq t \leq 300$ or $600 \leq t \leq 900$ or $1200 \leq 1500$.

Parameter	Description	Mean	Unit
μ_1	maximum proliferation rate (AD cells)	0.071	$[\text{day}]^{-1}$
μ_2	maximum proliferation rate (AI cells)	0.071	$[\text{day}]^{-1}$
q_1	minimum AD cell quota	0.613	$[\text{nmol}][\text{day}]^{-1}$
q_2	minimum AI quota	0.1971	$[\text{nmol}][\text{day}]^{-1}$
b	baseline PSA production rate	0.0379	$[\mu\text{g}][\text{nmol}]^{-1}[\text{day}]^{-1}$
σ	tumor PSA production rate	0.8667	$[\mu\text{g}][\text{nmol}]^{-1}[\text{L}]^{-1}[\text{day}]^{-1}$
ϵ	PSA clearance rate	0.0565	$[\text{day}]^{-1}$
d_1	maximum AD cell death rate	0.0687	$[\text{day}]^{-1}$
d_2	maximum AI cell death rate	0.0633	$[\text{day}]^{-1}$
δ_1	density death rate for AD cells	5	$[\text{L}]^{-1}[\text{day}]^{-1}$
δ_2	density death rate for AI cells	5	$[\text{L}]^{-1}[\text{day}]^{-1}$
R_1	AD death rate half-saturation	1.2499	$[\text{nmol}][\text{L}]^{-1}$
R_2	AI death rate half-saturation	2.7351	$[\text{nmol}][\text{L}]^{-1}$
c	maximum mutation rate	0.00015	$[\text{day}]^{-1}$
K	mutation rate half-saturation level	1	$[\text{nmol}][\text{day}]^{-1}$
γ_1	primary androgen production rate	0.3742	$[\text{day}]^{-1}$
γ_2	secondary androgen production rate	0.005	$[\text{day}]^{-1}$
m	diffusion rate from A to Q	0.7188	$[\text{day}]^{-1}$
A_0	maximum serum androgen level	11.63	$[\text{nmol}][\text{day}]^{-1}$
$x_1(0)$	initial population of AD cells	0.01	[L]
$x_2(0)$	initial population of AI cells	0.0001	[L]

Table 2.1: Parameters description

2.2 Sensitivity Analysis

We applied Sobol's sensitivity analysis using the `saltelli` and `sobol` modules from the Python library `SALib` (Sensitivity Analysis Library) [23]. For theoretical details see [5].

According to Saltelli *et al.* [5], first order Sobol sensitivities are defined as follows

"*...Given a model of the form* $Y = f(X_1, X_2, \ldots X_k)$*, with* Y *a scalar, a variance based first order effect for a generic factor* X_i *can be written as:*

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right),$$

where X_i is the i -th factor and $\mathbf{X}_{\sim i}$ denotes the matrix of all factors but X_i. The meaning of the inner expectation operator is that the mean of Y is taken over all possible values of $\mathbf{X}_{\sim i}$ while keeping X_i fixed. The outer variance is taken over all possible values of X_i. The associated sensitivity measure (first order sensitivity coefficient) is written as:

$$S_i = \frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)}{V(Y)}.$$

Then, $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right) + E_{X_i}\left(V_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right) = V(Y)$.

S_i is a normalized index, as $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)$ varies between zero and $V(Y)$.$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)$ measures the first order (e.g. additive) effect of X_i on the model output, while $E_{X_i}\left(V_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)$ is customarily called the residual.

Another popular variance based measure is the total effect index:

$$S_{Ti} = \frac{E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)}{V(Y)} = 1 - \frac{V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)}{V(Y)}.$$

S_{Ti} measures the total effect, i.e. first and higher order effects (interactions) of factor X_i. One way to visualize this is by considering that $V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)$ is the first order effect of $\mathbf{X}_{\sim i}$, so that $V(Y)$ minus $V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)$ must give the contribution of all terms in the variance decomposition which do include X_i..."

2.3 Inverse problems theory

We follow the statistical inversion approach where all variables are considered random variables that can be modeled with a probability distribution, so the posterior probability distribution gives us a parameter estimation [6].

2.3.1 Bayes' Theorem

The Bayes' theorem of inverse problems (as presented in [6]) relates two different probabilities densities for observable and non-observable random variables.

Theorem 1 (Bayes' theorem of inverse problems) *Assume that the random variable $X \in \mathbb{R}^n$ has a known prior probability density $\pi_{\mathrm{pr}}(x)$ and the data consist of the observed value $y_{observed}$ of an observable random variable $Y \in \mathbb{R}^k$ such that $\pi\left(y_{observed}\right) > 0$. Then the posterior probability distribution of X, given the data $y_{observed}$ is*

$$\pi_{post}(x) = \pi\left(x \mid y_{observed}\right) = \frac{\pi_{pr}(x)\pi\left(y_{observed} \mid x\right)}{\pi\left(y_{observed}\right)}.$$

2.3.2 Bayes Factor

The Bayes factor is a likelihood ratio of the marginal likelihood for two competing hypotheses, see [24].

From Theorem 1 we have that the likelihood of a model, $\mathcal{M}$, is:

$$\pi(y \mid \mathcal{M}) = \frac{\pi(\mathcal{M} \mid y)\pi(y)}{\pi(\mathcal{M})}$$

Thus, the Bayes factor for $\mathcal{M}1$ and $\mathcal{M}2$ is:

$$B := \frac{\pi(y \mid \mathcal{M}1)}{\pi(y \mid \mathcal{M}2)} = \frac{\frac{\pi(\mathcal{M}1|y)\pi(y)}{\pi(\mathcal{M}1)}}{\frac{\pi(\mathcal{M}2|y)\pi(y)}{\pi(\mathcal{M}2)}} = \frac{\pi(\mathcal{M}1 \mid y)\pi(y)\pi(\mathcal{M}2)}{\pi(\mathcal{M}2 \mid y)\pi(y)\pi(\mathcal{M}1)} = \frac{\pi(\mathcal{M}1 \mid y)\pi(\mathcal{M}2)}{\pi(\mathcal{M}2 \mid y)\pi(\mathcal{M}1)}$$

Assuming that $\pi(\mathcal{M}1) = \pi(\mathcal{M}2)$, then

$$B = \frac{\pi(\mathcal{M}1 \mid y)}{\pi(\mathcal{M}2 \mid y)} = \frac{\int \pi(y \mid \phi_1)\pi(\phi_1)\, d\phi_1}{\int \pi(y \mid \phi_2)\pi(\phi_2), d\phi_2} = \frac{Z_1}{Z_2},$$

where Z_1 and Z_2 are the normalizing constants for models $\mathcal{M}1$ and $\mathcal{M}2$, respectively.

A value of $B > 1$ means that $\mathcal{M}1$ is more strongly supported by the data than $\mathcal{M}2$ [25].

2.3.3 Statistical Models

As stated previously, we will not use real clinical data, but synthetic ones. The generations of these are as follows:

$$y_{\text{observed}} = y + \sigma\epsilon,$$

where y is the solution of the model with the reference values, $\epsilon \sim \mathcal{N}(0,1)$ and $\sigma = \frac{\max y}{100}$.

We have synthetic PSA data y^i_{observed} at times $t_1, t_2 2, \ldots, t_{1500}$.

Also assume that the data has a normal distribution and x_0 known

$$y^k_{\text{observed}} \sim \mathcal{N}(P(t_i, \theta), \sigma),$$

where $\theta = (\theta_1, \ldots, \theta_5)$ is the parameter vector of the model and $y_{\text{observed}} = (y^1_{\text{observed}}, \ldots, y^{1500}_{\text{observed}})$.

Assuming that the observations are independent from each other, the joint distribution of the observed PSA is a good approximation to the conditional probability (likelihood) $\pi(y_{\text{observed}}|\theta)$, which would simply be defined by the product of the individual probability density functions of the observations:

$$\pi(y_{\text{observed}}|\theta) = \prod_{k=1}^{1500} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_{\text{observed}} - P(t_k,\theta)}{\sigma}\right)^2},$$

where σ is the one calculated for the creation of the synthetic data.

Now we need an a priori probability density function $\pi_{prior}(\theta)$.

Since all parameters have positive ranges, we can normalize them to $[0,1]$, so we can model them as Beta distributions. Suppose that $\theta_i \sim \text{Beta}(\alpha_i, \beta_i), i = 1, \ldots, 5$, are independent, i.e., $\pi_{\text{prior}}(\theta) = \pi_{\text{prior}}(\theta_1) \times \cdots \times \pi_{\text{prior}}(\theta_5)$. Bayes

Theorem implies the equality:

$$\pi(\theta|y_{\text{observed}}) \propto \pi(y_{\text{observed}}|\theta) \times \pi_{\text{prior}}(\theta)$$
$$= \prod_{k=1}^{1500} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_{\text{observed}} - P(t_k,\theta)}{\sigma}\right)^2} \times \text{Beta}(\alpha_1, \beta_1) \times \cdots \times \text{Beta}(\alpha_5, \beta_5),$$

The prior distribution for each parameter, θ_i is a Beta distribution:

$$\pi_{\text{prior}}(\theta_i) = \frac{\Gamma(\alpha+\beta)\theta_i^{\alpha-1}(1-\theta_i)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}, \tag{2.2}$$

where α and β are the shape parameters for the Beta distribution, $\theta_i \in [0, 1]$, for $i = 1, \ldots, 5$, and Γ is the Gamma function ($\Gamma(n) = (n-1)!$ for n in $\mathbf{Z}^+$).

We choose α and β to be equal to 1.1 for every parameter.

Our last step is to carry out Markov Chain Monte Carlo (MCMC) to construct a statistically meaningful sample of the posterior conditional distribution $\pi(\theta|y_{\text{observed}})$.

2.3.4 Markov Chain Monte Carlo

There are no analytical methods to explore the posterior distribution (2.2) provided the non-linearity of the inference problem. We resort to Markov Chain Monte Carlo (MCMC) to explore the posterior distribution. We use the twalk method of Christen and Fox [26].

MCMC is a stochastic collocation method used for approximating the support of a target distribution like (2.2), that consists of *letting the posterior distribution determine a set of points, a sample, that supports well the distribution. These sample points can then be used for approximate integration* [6].

In the next Subsection 2.3.4, we will show why it is plausible to propose a Gaussian (Normal) approximation for the posterior distribution.

Posterior Approximation by a Normal distribution

For any probability density function (pdf) that is smooth and well peaked around its point of maxima, Laplace proposed to approximate it by a normal pdf. It is a simple 2-term Taylor expansion trick on the log pdf. If $\hat{\theta}$ denotes the point of maxima of a pdf $h(\theta)$, then it is also the point of maxima of the log-pdf $q(\theta) = \log h(\theta)$ and we can write:

$$\begin{aligned} q(\theta) &\approx q(\hat{\theta}) + (\theta - \hat{\theta})\dot{q}(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^2 \ddot{q}(\hat{\theta}) \\ &= q(\hat{\theta}) + 0 + \frac{1}{2}(\theta - \hat{\theta})^2 \ddot{q}(\hat{\theta}) \quad [\text{ because } \dot{q}(\hat{\theta}) = 0] \\ &= \text{ const } - \frac{1}{2}(\theta - \hat{\theta})^2 \ddot{q}(\hat{\theta}) \\ &= \text{ const } - \frac{(\theta - \bar{a})^2}{2b^2} \end{aligned}$$

with $\bar{a} = \hat{\theta}$ and $b^2 = \{-\ddot{q}(\hat{\theta})\}^{-1}$ (notice $\ddot{q}(\hat{\theta}) < 0$ because $\hat{\theta}$ is a maxima). But the right hand side of the last display matches the log-pdf of Normal $\left(\tilde{a}, \bar{b}^2\right)$. Hence the pdf $h(\theta)$ is approximately the Normal $\left(\bar{a}, \tilde{b}^2\right)$ pdf with $\bar{a} = \hat{\theta}$ and $\tilde{b}^2 = \{-\ddot{q}(\hat{\theta})\}^{-1}$.

All one needs is that the log-pdf is smooth at the maximum and peaks well at it so that the quadratic approximation is good. We only need to know the point of maximum $\hat{\theta}$ and the curvature $-\ddot{q}(\theta)$ at this point.

The same technique could be applied to a posterior pdf $\xi(\theta \mid x) = \text{const} \times L_x(\theta)\xi(\theta)$. The log-pdf in this case is $q(\theta) = \text{const} + \ell_x(\theta) + \log \xi(\theta)$. Typically we will not know the value of the constant term at the front. But it does not affect when computing the point of maximum $\hat{\theta}$ and the curvature $-\ddot{q}(\theta)$ at $\theta = \hat{\theta}$.

Now we need the following:

Theorem 2 (Bernstein-von Mises) *Consider the model $Y_1, \cdots Y_n \overset{IID}{\sim} g(y_i|x)$, $x \in X$. Under some regularity conditions on the pdfs/pmfs $g(\cdot|x)$, including that all of them have the same support, and that for each $y_i, x \mapsto \log g(y_i|x)$ is twice continuously differentiable, we have that for any prior $\pi_{pr}(x)$*

which is positive, bounded and twice differentiable over X,

$$\sup_z \left| P(x \leq z | Y = y) - \Phi\left(\{-\ddot{q}(\hat{x})\}^{1/2} (z - \hat{x}) \right) \right| \approx 0$$

for all large n.

Bernstein-von Mises applies to most of the standard models for which a conjugate prior family exists. Therefore for large n, the conjugate posterior should look like a bell curve too.

In Section 3.4, the required hypotheses on the posterior probability, for both model $\mathcal{M}1$ and model $\mathcal{M}2$, are verified in order to use Theorem 2 (Bernstein-von Mises) and thus be able to approximate it by a Gaussian.

Chapter 3

Results

In this Chapter we present our findings regarding the model sensibility analysis, formal inference and model selection. The sensibility analysis is aimed at choosing a set of parameters to make the inference, and to compare the corresponding inference results with those of a reference model parametrization. The rationale of our approach is based on the following empirical argument: under mild regularity assumptions in Bayesian linear regression, the posterior distribution variance should shrink with respect to the prior distributions along the direction of those parameters that are informed by the data. Likewise, if a parameter is not informed by the data, then its marginal posterior distribution should resemble the prior model. Numerical evidence in the prostate cancer model studied in this thesis indicates that the posterior distribution is unimodal. Consequently, given a fixed set of synthetic data, we use the Sobol sensitivity analysis to rank the model parameters in terms of how much does each parameter variance contribute to the variance of the likelihood on a given support, and we keep the five top ranked parameters as our model parametrization. The results are shown in Section 3.2. Due to the multiple scales present in the model according to Gutenkunst *et al.* [2], we rank one parameter at a time by sequentially stripping the model from the leading parameter and fixing it to some reference value.

Afterwards, we use Bayes factors to carry out formal model selection and we

show that the proposed method gives the better model. To accomplish this task, we use Markov Chain Monte Carlo to sample the reference posterior distribution and the one proposed in this thesis. Later, we make a Gaussian approximation of each posterior distribution based on the MCMC samples. We compute the Bayes factors using the normalizing constant of the Gaussian approximations.

One very important feature of the dynamical system being investigated in this thesis is numerical stiffness. To address this problem we use a suite of numerical integration methods for ordinary differential equations given by `odeint` from Python's library `Scipy`. The numerical integration method `odeint` receives the Jacobian of the dynamical system right hand side as an argument and approximates the Jacobian eigenvalues to decide automatically whether to use an Adams-Bashforth method (non-stiff) or a backward differentiation formula (stiff).

We believe that our findings on model parametrization may be generalized to other cancer dynamics models, as well as to models of other more general biological systems. Furthermore, expert insight, i.e. from medical doctors, on the scale of model parameters would be very useful to carry out more deep analysis.

3.1 Model Stiffness

When numerically solving model $\mathcal{M}2$, we see that the ODE system is stiff. See Figure 3.1. Given the numerical nature of our research, we require some form of unsupervised robustness in the numerical integration of the model.

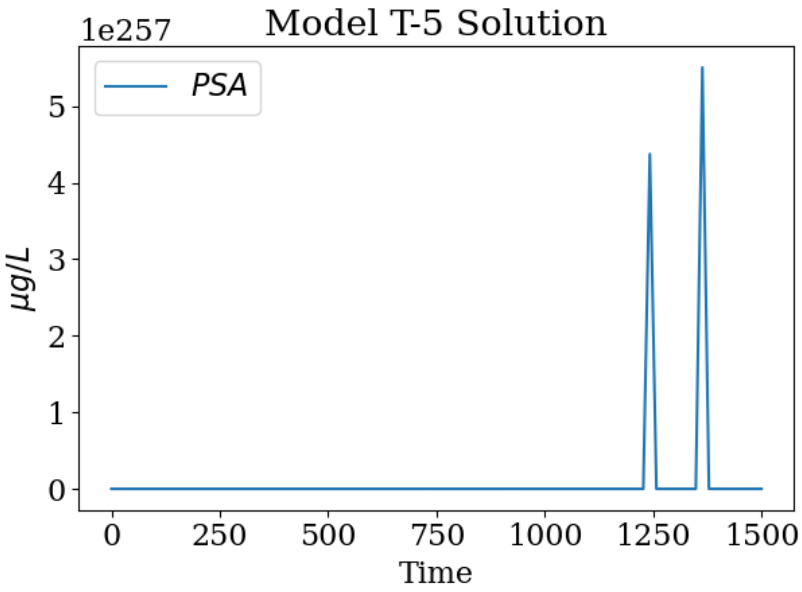

Figure 3.1: $\mathcal{M}2$ solution for PSA without passing the model jacobian as an argument to `odeint`.

Thus, we use the `odeint` method from Python's library `scipy` with the Jacobian of the dynamical system right hand side as an argument [27]. Of note, `odeint` approximates the eigenvalues of the dynamical system Jacobian as it marches on time, and automatically decides whether to use an Adams-Bashforth multistep method in case the ODE is non-stiff, or a backward differentiation formula in case the ODE system is stiff. Using this robust suite of numerical integration circumvents the verification task as suggested by Oden *et al.* [28]. Although the topic of numerical analysis is not the goal of this work, we remark that the verification of the numerical model used in this thesis is of paramount importance.

3.2 Sensitivity Analysis

Below we present the first 5 iterations of parameter ranking based on the Sobol sensitivity analysis and parameter stripping (Tables 3.1, 3.2, 3.3, 3.4, 3.5). Namely, selected parameters are fixed to a reference value after being selected.

Parameter	**S1**	**S1_conf**	**ST**	**ST_conf**
μ_1	0.001765	0.002340	0.001469	0.001967
μ_2	0.008909	0.023479	0.103712	0.153503
q_1	-0.007577	0.016155	0.088806	0.178843
q_2	0.000418	0.000963	0.000427	0.000587
b	0.069536	0.064823	0.626695	0.087236
σ	0.005903	0.030135	0.117871	0.207771
ϵ	0.059577	0.076895	0.592916	0.076401
d_1	-0.002053	0.004958	0.014018	0.025704
d_2	0.000698	0.001000	0.000446	0.000629
δ_1	0.001453	0.002137	0.000532	0.000698
δ_2	0.001386	0.001630	0.001747	0.002363
R_1	0.000328	0.000870	0.000429	0.000710
R_2	0.000451	0.000909	0.000282	0.000556
c	0.000457	0.000909	0.000282	0.000556
K	0.000456	0.000909	0.000282	0.000556
γ_1	0.000552	0.001037	0.000420	0.000581
γ_2	0.000432	0.000912	0.000378	0.000613
m	0.004968	0.006191	0.006480	0.001093
A_0	0.056413	0.060459	0.611050	0.078197
$x_1(0)$	0.000000	0.000000	0.000000	0.000000
$x_2(0)$	0.000000	0.000000	0.000000	0.000000

Table 3.1: Sobol's sensitivity analysis first iteration

Parameter	S1	S1_conf	ST	ST_conf
μ_1	0.001834	0.002779	0.000877	0.001193
μ_2	0.001425	0.009563	0.011471	0.001634
q_1	0.000483	0.001182	0.000373	0.000711
q_2	-0.009320	0.011554	0.086095	0.121678
σ	0.000481	0.001167	0.000383	0.000713
ϵ	0.167161	0.090656	0.840193	0.088059
d_1	-0.000103	0.000206	0.000007	0.000015
d_2	0.000695	0.001260	0.000599	0.000846
δ_1	0.002290	0.005248	0.002100	0.003267
δ_2	0.000980	0.001753	0.003686	0.006965
R_1	0.000615	0.001171	0.000375	0.000711
R_2	0.000792	0.001441	0.000445	0.000739
c	0.000453	0.001197	0.000374	0.000711
K	0.001787	0.002528	0.001588	0.002235
γ_1	0.002532	0.002912	0.000948	0.001038
γ_2	0.000680	0.001171	0.000381	0.000711
m	-0.004912	0.022978	0.341728	0.584830
A_0	0.128800	0.095533	1.573242	1.494242
$x_1(0)$	0.000000	0.000000	0.000000	0.000000
$x_2(0)$	0.000000	0.000000	0.000000	0.000000

Table 3.2: Sobol's sensitivity analysis second iteration

Parameter	S1	S1_conf	ST	ST_conf
μ_1	-0.039606	0.041518	0.688884	0.714722
μ_2	-0.040813	0.043622	0.727035	0.721380
q_1	-0.040344	0.042354	0.715501	0.742359
q_2	-0.040522	0.041981	0.715743	0.741829
σ	-0.042626	0.037561	1.068618	1.610828
d_1	0.051334	0.333238	11.480921	40.277976
d_2	-0.040342	0.042372	0.715498	0.742362
δ_1	-0.040709	0.041591	0.715677	0.741981
δ_2	-0.032981	0.135129	1.174270	1.512754
R_1	-0.053605	0.068618	1.102131	1.956318
R_2	-0.042633	0.038469	0.836839	0.661484
c	-0.040310	0.042444	0.715501	0.742362
K	-0.061227	0.078264	1.301971	2.271343
γ_1	-0.040289	0.042534	0.715620	0.741809
γ_2	-0.039670	0.044176	0.715824	0.741628
m	-0.039765	0.046627	0.725013	0.712885
A_0	0.244901	0.712170	0.960341	0.054252
$x_1(0)$	0.000302	0.001387	0.000134	0.000616
$x_2(0)$	0.000307	0.001409	0.000138	0.000636

Table 3.3: Sobol's sensitivity analysis third iteration

Parameter	S1	S1_conf	ST	ST_conf
μ_1	-0.000560	7.103237	0.278018	3526.185610
μ_2	0.000229 3	.026604	0.038075	686.712488
q_1	-0.000001	0.015700	0.000001	0.008295
q_2	-0.000002	0.031764	0.000003	0.053873
σ	-0.001305	20.450230	1.637895	25661.060112
d_1	-0.000000	0.002778	0.000000	0.000529
d_2	-0.000330	4.015757	0.080535	981.269114
δ_1	-0.000423	7.045382	0.297338	5242.698437
δ_2	-0.000126	2.089244	0.329717	5649.855534
R_1	-0.000331	6.683270	0.085493	1741.953153
R_2	-0.000110	1.559100	0.236599	3343.092075
c	-0.000241	4.558030	0.058654	1112.079940
K	-0.000573	9.871518	0.610213	10513.702803
γ_1	-0.000203	4.432037	0.085847	1337.942492
γ_2	-0.001586 2	2.413685	2.316874	39645.296903
m	0.000281	0.096618	0.000046	0.533588
$x_1(0)$	0.000000	0.000000	0.000000	0.000000
$x_2(0)$	0.000000	0.000000	0.000000	0.000000

Table 3.4: Sobol's sensitivity analysis fourth iteration

Parameter	S1	S1_conf	ST	ST_conf
μ_1	-0.588844	1.047543	9.296985	17.495542
μ_2	-4.877749	10.495671	3776.069264	6825.979154
q_1	-87.130723	146.736350	85340.839095	144079.889107
q_2	-0.062627	0.082180	0.027882	0.034970
σ	-0.020792	0.038218	0.015175	0.024385
d_1	-87.512239	163.997137	190661.015896	357302.568189
d_2	-0.041333	0.043291	0.011357	0.014925
δ_1	-74.368736	148.821202	145298.391234	290887.254693
δ_2	-0.013875	0.025933	0.003721	0.006895
R_1	-41.165961	91.668827	32881.089294	74020.444562
R_2	-0.002366	0.003824	0.000108	0.000069
c	0.000965	0.003397	0.000077	0.000053
K	-34.282095	51.256507	33162.603239	48156.641390
γ_1	-26.693361	56.792984	11906.694737	25312.941983
γ_2	-0.050788	0.059807	0.023560	0.028995
$x_1(0)$	-0.000633	0.001210	0.000007	0.000013
$x_2(0)$	0.000257	0.000492	0.000001	0.000002

Table 3.5: Sobol's sensitivity analysis fifth iteration

In view of the results shown above, we propose a model (henceforth referred to as $\mathcal{M}1$) with 5 parameters: b, ϵ, c, m, and A_0.

3.3 MCMC convergence analysis

In this section we offer numerical evidence of the Markov Chain Monte Carlo.

The next plots show the convergence of the MCMC samples for each model (Figures 3.2 and 3.3).

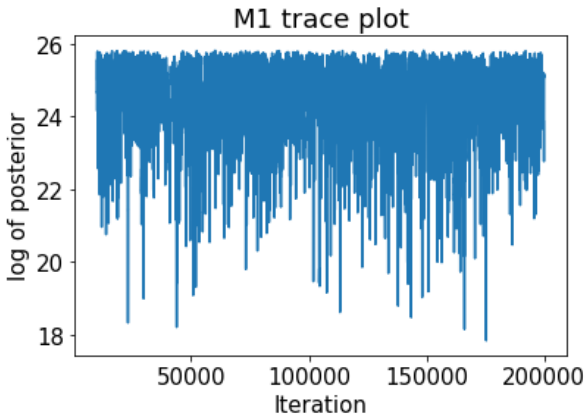

Figure 3.2: $\mathcal{M}1$ trace plot

For model $\mathcal{M}1$, the Integrated Autocorrelation Time (IAT) is 96.4, so the IAT divided by the number of parameters is 19.3.

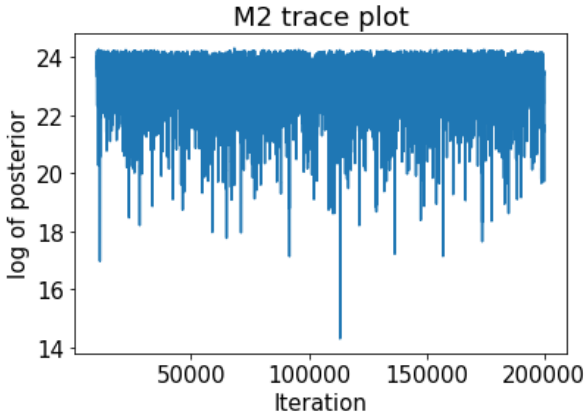

Figure 3.3: $\mathcal{M}2$ trace plot

For model $\mathcal{M}2$, the Integrated Autocorrelation Time (IAT) is 109.1, so the IAT divided by the number of parameters is 21.8.

3.4 Normalizing Constants

As seen on Subsection 2.3.4, we can elicit the posterior probability density function, $\pi(x \mid y_{\text{observed}})$, as a Gaussian if it meets certain conditions. For simplicity, we will just write y instead of y_{observed}.

(i) $\log \pi(x \mid y)$ is twice continuously differentiable.

(ii) $\pi_{\text{pr}}(x)$ is positive, bounded and twice differentiable over the parameter space.

The hypotheses in MCMC are a Gaussian likelihood and a priori Beta distribution for the parameters. But since they all have bounded positive ranges, condition (ii) is partially met. For condition (ii), we do the following:

By Bayes' Theorem 1 we know that:

$$\log(\pi(x|y)) = c_1 + \ell_x(y) + \log(\pi_{\text{pr}}(x)).$$

Assuming $x = (b, \epsilon, c, m, A_0)$ is a vector of parameters, y is the data, and $\ell_x(y)$ is the log-likelihood of the data, y, that depends on the parameters, x, then we have

$$\ell_x(y) = \sum_{i=1}^{n} \left[\log\left(\frac{1}{\sigma_x \sqrt{2\pi}} \right) - \frac{(y_i - \mu_x)^2}{2\sigma_x^2} \right],$$

and

$$\begin{aligned} \log(\pi_{\text{pr}}(x)) = &-5\log(\Gamma^2(1.1)) + 5\log(\Gamma(2.2)) + \log(b^{0.1}) + \log((1-b)^{0.1}) \\ &+ \log(\epsilon^{0.1}) + \log((1-\epsilon)^{0.1}) + \log(c^{0.1}) + \log((1-c)^{0.1}) \\ &+ \log(m^{0.1})\log((1-m)^{0.1}) + \log(A_0^{0.1}) + \log((1-A_0)^{0.1})). \end{aligned}$$

Thus,

$$q(x) = \log(\pi(x|y)) = c_3 + n\log\left(\frac{1}{\sigma_x\sqrt{2\pi}}\right) - \sum_{i=1}^{n}\frac{(y_i-\mu_x)^2}{2\sigma_x^2} + \log(b^{0.1}) + \log((1-b)^{0.1}) + \log(\epsilon^{0.1}) + \log((1-\epsilon)^{0.1}) + \log(c^{0.1}) + \log((1-c)^{0.1}) + \log(m^{0.1})\log((1-m)^{0.1}) + \log(A_0^{0.1}) + \log((1-A_0)^{0.1})).$$

By the Bernstein von-Mises Theorem 2, we can say that the posterior distributions for each model is a Gaussian whose mean and variance can be estimated by the mean and variance from a MCMC sample of the parameters.

In Table 3.6 we present the estimations of the normalizing constants for each model.

Model	Normalizing Constant
$\mathcal{M}1$	$3.3797029802338296 \times 10^{12}$
$\mathcal{M}2$	$9.38764868.8826965 \times 10^{8}$

Table 3.6: Normalizing constants estimations

3.5 Bayes Factor

The core of this work is presented in this Section. As seen in Subsection 2.3.2, $B = \frac{Z_1}{Z_2}$, thus:

$$\hat{B}_{\mathcal{M}1,\mathcal{M}2} = \frac{\hat{Z}_{\mathcal{M}1}}{\hat{Z}_{\mathcal{M}2}} = \frac{3.3797029802338296 \times 10^{12}}{9.38764868.8826965 \times 10^{8}} = 3.6001592009469846 \times 10^{3}$$

which is greater than 1 therefore, $\mathcal{M}1$ **is more supported by the data than** $\mathcal{M}2$.

3.6 Distributions and Predictions

The next plots (Figures 3.4 and 3.5) are generated using a Python library called `corner` [29] which helps to visualize the distributions for each parameter and a two-dimensional projection for the covariance across parameters.

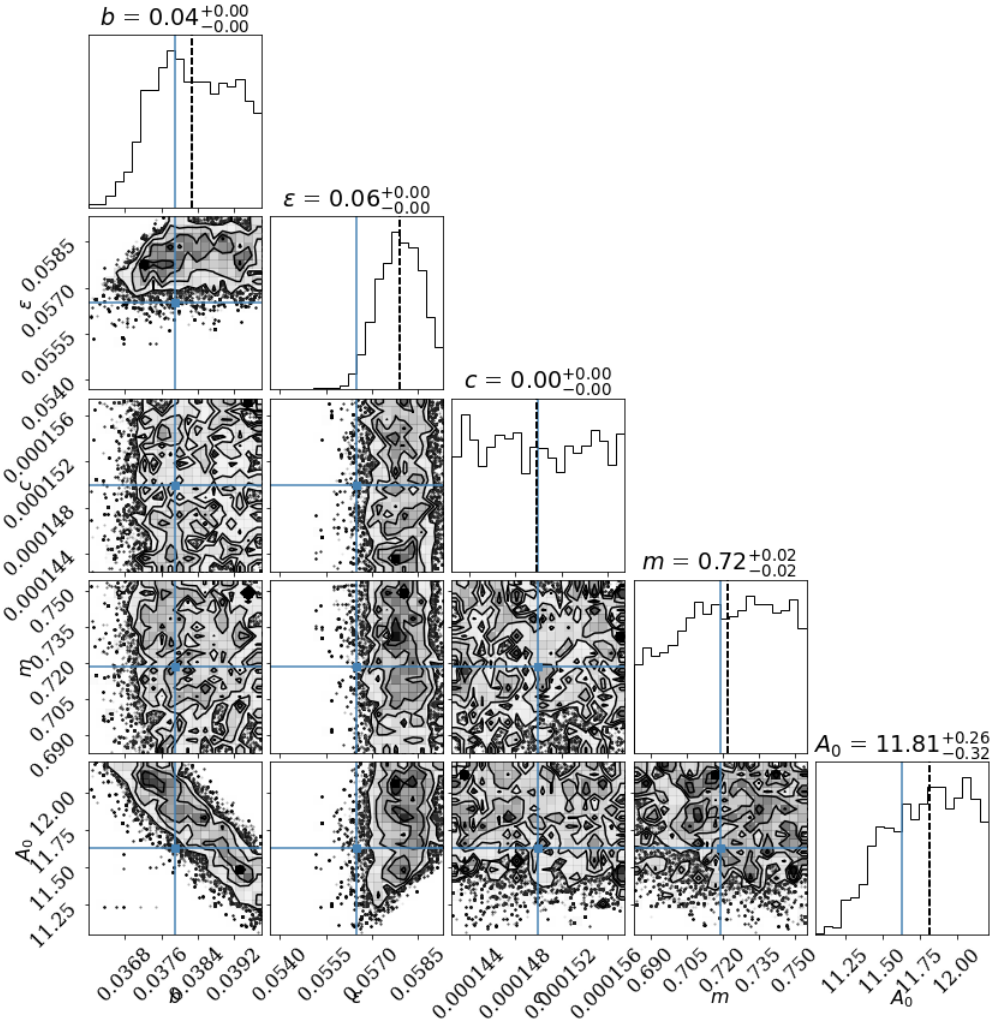

Figure 3.4: Corner $\mathcal{M}1$

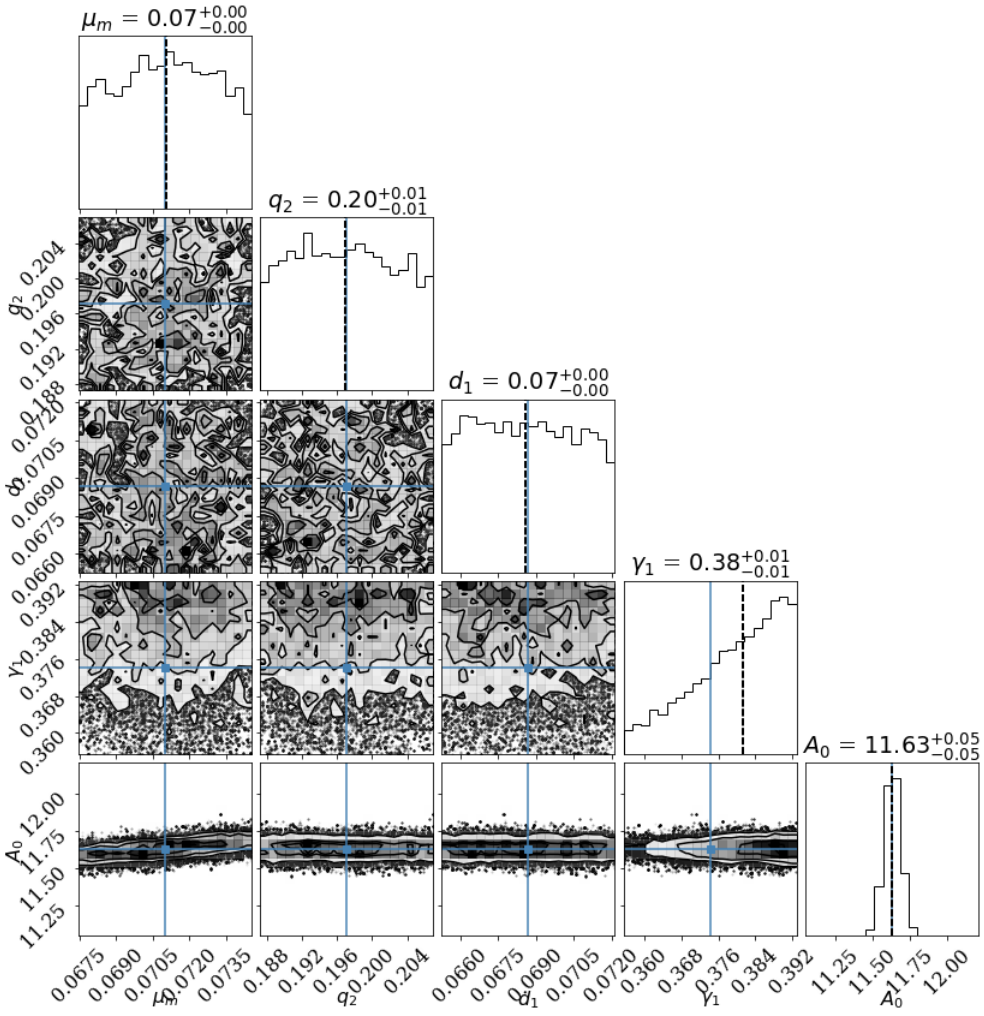

Figure 3.5: Corner $\mathcal{M}2$

Now that we have an approximation for the posterior distribution, we can estimate PSA values with the Maximum A Posteriori (MAP) and the Conditional Mean (CM). See Figures 3.6 and 3.7.

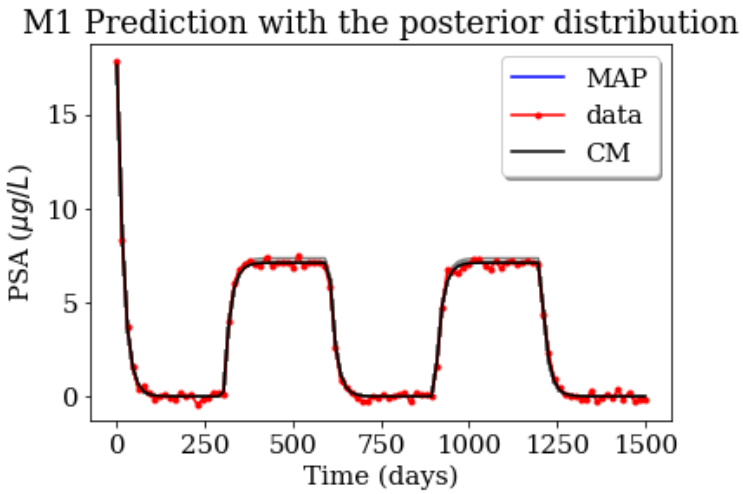

Figure 3.6: Prediction $\mathcal{M}1$

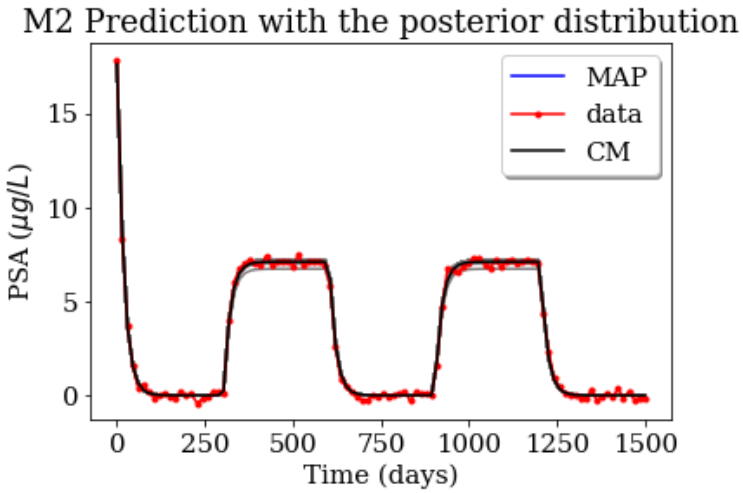

Figure 3.7: Prediction $\mathcal{M}2$

Chapter 4

Discussion

In this thesis, we introduce a general method for sensitivity informed parameter selection in models for biological systems. First, we perform a Sobol sensitivity analysis, taking the likelihood (conditional probability) as a quantity of interest. We address the "sloopy parameters" problem, mentioned in [2], of multiple scales in the parameters, by sequentially stripping the model out from the leading parameter and fixing it to a reference value in order to rank the model parameters in terms of how much does each parameter variance contribute to the variance of the likelihood, and maintain the desired number of parameters which contribute the most. If a parameter does not contribute or contributes a little to the likelihood variance, then that parameter is non identifiable. On the other hand, the parameters contributing the most to the likelihood variance shrink their posterior distribution variance when conditioned to observed data, which makes our inference easier. It is noteworthy that some parameters can be fixed by a specialist in the field of the model we are dealing with.

In this work, we design synthetic data by simulating from the reference model, adding Gaussian noise with a signal to noise ratio of 100. We generate synthetic data fixing a seed with Python's library `random` to `random.seed(10)`.

We make our model selection from a Bayesian viewpoint, that is, we are using the Bayes factor for model comparison. For that purpose, we elicit informative prior distributions for the model parameters, using a likelihood model which is

known by construction, and Markov Chain Monte Carlo (MCMC) to sample the posterior distributions for each model. Afterwards, we fit a Gaussian distribution to the samples obtained through MCMC to exploit the fact that we can compute the normalization constant of each distribution in closed form.

The validity of our findings relies on the fact that the arising posterior distribution for the problem at hand is unimodal. This supports that we can approximate the posterior distribution in a neighborhood of the Maximum A Posteriori (MAP) by a Gaussian distribution invoking the Bernstein-von Misses theorem.

Regarding future work, we believe that more general sensitivity informed parameter selection problems in biological systems may be derived from the method proposed in this thesis. Indeed, we want to explore whether a bimodal posterior distribution may be approximated by a mixture of Gaussian distributions to select the model parameters. The motivation is given by biological system where bistability plays an important role; to name one, glucose metabolism.

Chapter 5

Conclusions

We proposed a general method for global sensitivity informed parameter selection in models of biological systems. We applied our method to a model of prostate cancer under intermittent androgen suppression. The validity of our methods relies on the fact that the posterior distribution for the problem at hand can be approximated by a Gaussian distribution in a neighborhood of the maximum a posteriori.

Model parameters are ranked according to how much does the variance of each parameter contribute to the variance of the likelihood. Parameters are ranked one at a time stripping the model out from the leading parameter and fixing it to a reference value.

Our findings can be extended in several directions. We may explore what is the optimal number of parameters that can be inferred with a specific model and given a set of data. We may generalize our findings to explore models whose posterior distribution exhibits bistability. Notice that, bistability is a very important problem arising in biological systems such as glucose metabolism. We may explore other approaches to carry out model selection such as importance sampling.

Acronyms

AD Androgen-dependent.

AI Androgen-independent.

CAS Continuous Androgen Suppression.

CM Conditional Mean.

IAS Intermittent Androgen Suppression.

IAT Integrated Autocorrelation Time.

MAP Maximum a posteriori.

MCMC Markov Chain Monte Carlo.

ODE Ordinary Differential Equation.

PCa Prostate Cancer.

PDE Partial Differential Equation.

PSA Prostate-Specific Antigen.

TNM Tumor, Nodes, and Metastases.

Bibliography

[1] Global Cancer Observatory. International Agency for Research on Cancer, 2021.

[2] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLoS Computational Biology*, 3(10), October 2007.

[3] Zhimin Wu, Tin Phan, Javier Baez, Yang Kuang, and Eric J. Kostelich. Predictability and identifiability assessment of models for prostate cancer under androgen suppression therapy. *Mathematical biosciences and engineering: MBE*, 16(5):3512–3536, 2019.

[4] Javier Baez and Yang Kuang. Mathematical Models of Androgen Resistance in Prostate Cancer Patients under Intermittent Androgen Suppression Therapy. *Applied Sciences*, 6(11):352, November 2016. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

[5] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270, February 2010.

[6] Jari Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. 160. Springer-Verlag New York, United States of America, 2005.

[7] STA 250: Statistics. Notes 11. Laplace Approximation to the Posterior.

[8] Kristin R. Swanson, Lawrence D. True, Daniel W. Lin, Kent R. Buhler, Robert Vessella, and James D. Murray. A Quantitative Model for the Dynamics of Serum Prostate-Specific Antigen as a Marker for Cancerous Growth: An Explanation for a Medical Anomaly. *The American Journal of Pathology*, 158(6):2195–2199, June 2001.

[9] Charles Huggins, R. E. Stevens, and Clarence V. Hodges. Studies on Prostatic Cancer: II. The Effects of Castration on Advanced Carcinoma of the Prostate Gland. *Archives of Surgery*, 43(2):209–223, August 1941. Publisher: American Medical Association.

[10] Robin T. Vollmer, Shin Egawa, Sadahito Kuwao, and Shiro Baba. The dynamics of prostate specific antigen during watchful waiting of prostate carcinoma. *Cancer*, 94(6):1692–1698, 2002. _eprint: https://acsjournals.onlinelibrary.wiley.com/doi/pdf/10.1002/cncr.10443.

[11] Robin T. Vollmer and Peter A. Humphrey. Tumor Volume in Prostate Cancer and Serum Prostate-Specific Antigen: Analysis From a Kinetic Viewpoint. *American Journal of Clinical Pathology*, 119(1):80–89, January 2003.

[12] Aiko Miyamura Ideta, Gouhei Tanaka, Takumi Takeuchi, and Kazuyuki Aihara. A Mathematical Model of Intermittent Androgen Suppression for Prostate Cancer. *Journal of Nonlinear Science*, 18(6):593, October 2008.

[13] Gouhei Tanaka, Yoshito Hirata, S. Larry Goldenberg, Nicholas Bruchovsky, and Kazuyuki Aihara. Mathematical modelling of prostate cancer growth and its application to hormone therapy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1930):5029–5044, November 2010. Publisher: Royal Society.

[14] Yoshito Hirata, Mario di Bernardo, Nicholas Bruchovsky, and Kazuyuki Aihara. Hybrid optimal scheduling for intermittent androgen suppression of prostate cancer. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(4):045125, December 2010. Publisher: American Institute of Physics.

[15] Harsh Vardhan Jain, Steven K. Clinton, Arvinder Bhinder, and Avner Friedman. Mathematical modeling of prostate cancer progression in response to androgen ablation therapy. *Proceedings of the National Academy of Sciences*, 108(49):19701–19706, December 2011. Publisher: National Academy of Sciences Section: Biological Sciences.

[16] Travis Portz, Yang Kuang, and John D. Nagy. A clinical data validated mathematical model of prostate cancer growth under intermittent androgen suppression therapy. *AIP Advances*, 2(1):011002, March 2012. Publisher: American Institute of Physics.

[17] Clément Draghi, Fabrice Denis, Alain Tolédano, and Christophe Letellier. Parameter identification of a model for prostate cancer treated by intermittent therapy. *Journal of Theoretical Biology*, 461:117–132, January 2019.

[18] Tin Phan, Kyle Nguyen, Preeti Sharma, and Yang Kuang. The Impact of Intermittent Androgen Suppression Therapy in Prostate Cancer Modeling. *Applied Sciences*, 9(1):36, January 2019. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[19] Tin Phan, Changhan He, Alejandro Martinez, Yang Kuang, 1 School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA, and 2 Department of Physics,Department of Physics, Tempe, AZ 85287, USA. Dynamics and implications of models for intermittent androgen suppression therapy. *Mathematical Biosciences and Engineering*, 16(1):187–204, 2019.

[20] Ayako Nakanishi and Yoshito Hirata. Practically scheduling hormone therapy for prostate cancer using a mathematical model. *Journal of Theoretical Biology*, 478:48–57, October 2019.

[21] M. R. Droop. Some Thoughts on Nutrient Limitation in Algae1. *Journal of Phycology*, 9(3):264–272, 1973. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1529-8817.1973.tb04092.x.

[22] Nicholas Bruchovsky, Laurence Klotz, Juanita Crook, Shawn Malone, Charles Ludgate, W. James Morris, Martin E. Gleave, and S. Larry Goldenberg. Final results of the Canadian prospective phase II trial of intermittent androgen suppression for men in biochemical recurrence after radiotherapy for locally advanced prostate cancer. *Cancer*, 107(2):389–395, 2006. _eprint: https://acsjournals.onlinelibrary.wiley.com/doi/pdf/10.1002/cncr.21989.

[23] Jon Herman and Will Usher. Salib: An open-source python library for sensitivity analysis. *Journal of Open Source Software*, 2(9):97, 2017.

[24] Phillip I. Good and James W. Hardin. *Common errors in statistics (and how to avoid them)*. Wiley, Hoboken, N.J, 2nd ed edition, 2006.

[25] Robert E. Kass and Adrian E. Raftery. Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795, June 1995. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572.

[26] J. Andrés Christen and Colin Fox. A general purpose sampling algorithm for continuous distributions (the t-walk). *Bayesian Analysis*, 5(2):263 – 281, 2010.

[27] A.C. Hindmarsh and L.R. Petzold. Lsoda, Ordinary Differential Equation Solver for Stiff or Non-Stiff System, 2005.

[28] Tinsley Oden, Robert Moser, and Omar Ghattas. Computer predictions with quantified uncertainty, part i. *SIAM News*, 43(9):1–3, 2010.

[29] Daniel Foreman-Mackey. corner.py: Scatterplot matrices in python. *The Journal of Open Source Software*, 1(2):24, jun 2016.