UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

Università degli Studi di Modena e Reggio Emilia

Dipartimento Di Scienze Fisiche, Informatiche E Matematiche

Corso di Laurea Magistrale in Matematica

UNDERSTANDING THE OUTCOMES OF EXPERIMENTAL CANCER THERAPIES THROUGH MATHEMATICAL MODELLING

Relatore:
Prof. Marco Prato

Candidato:
Francesco Luppi Pittigliani

Correlatore:
Prof. Federico Frascoli

Anno Accademico 2024–2025

"All models are wrong, but some are useful"

- George E. P. Box

Abstract

Oncolytic virotherapy and immunotherapy are examples of reasonably new techniques that are often combined in the fight against cancer. For some types of cancers, leukemias and melanomas for example, results have been promising. Some other cancers are instead very hard to treat, like brain cancers or pancreatic carcinomas, and progress has been less encouraging with those cases. Overall, there is currently a considerable interest in the scientific community on the use of viruses combined with the stimulation of the immune system to mount ad-hoc responses that can infect and destroy tumour cells, while preserving healthy cells and the surrounding micro-environment. For instance, five oncolytic viruses are now approved for patients' use in tumours as different as glioblastoma, melanoma, head and neck carcinoma, combined with immunotherapies like checkpoint blockade inhibitors, CAR-T, interferon therapy and immune boosting to advance our fight against malignancies.

In this project, the goal is to improve existing mathematical and computational models so that they can better approximate and explain existing clinical and laboratory data. To achieve this, we perform a multi-stage experimental and data analysis focused on the growth dynamics of untreated and treated tumour populations. We employ Maximum Likelihood Estimation (MLE) to fit typical solutions of Ordinary Differential Equation (ODE) for cancer growth - specifically Logistic, Gompertz, and Richards - to experimental datasets. We analyse several metrics, assuming a normal distribution for the experimental noise, and later compare it with an alternative log-normal one. A rigorous persistence analysis is also conducted by iteratively subsampling data points to evaluate model robustness and identify the threshold where structural complexity outweighs parsimony. Furthermore, we assess the forecasting accuracy of these prototypical models, showing that Sigmoidal models possess a good balance between qualitative fitting capacity and complexity in high-density data, whilst still performing well for more sparse clinical observations. Finally, using this framework, we investigate when solutions of ODEs typically used in cancer therapy can predict long-term cyclical behaviour of tumours, with an accuracy that can be useful for clinicians and experimentalists.

This research project has been partly developed during a visiting period at **Swinburne University of Technology** in Melbourne, Australia, aiming to bridge the gap between experimental cancer therapies and mathematical modelling.

INDEX

Chapter 1

Oncolytic Virotherapy

Cancer is the second leading cause of disease related mortality worldwide [1]. Furthermore, concern is widespread due to the severity of symptoms and outcomes. Screening programs for early detection play an important role for patients at various degrees of advancements in the disease and seriousness of symptoms. Improvements in high precision medicine and new type of therapies promise to achieve better survival rates, while plenty of technologies have been developed and implemented, such as immunotherapies. Nevertheless, complete tumour eradication is hard to accomplish in many cases and research for new treatments is still considered high-priority.

It is worth reminding that cancer is not a single disease entity, but rather a vast group of diverse pathological states, characterised by not negligible inter-patient and intra-tumour heterogeneity that defies a one-size-fits-all approach. Recently, new insights into molecular and immunology-based research with the implementation of techniques such us *Virotherapy* have led to better managing some type of malignant diseases. Here, we now focus on this promising, relatively new treatment.

1.1 Viruses: from Enemy to Possible Cure

Intuitively, dealing with a virus infection does not sound desirable: indeed, during entire human history, infectious diseases have often played a critical antagonistic role in our evolution and survival. How did the idea of regulating viral mechanism in favour of medicine even come about?

Oncolytic Virotherapy (OVT) is a treatment that uses biotechnology to convert viruses into therapeutic agents by genetically engineering their behaviour to treat diseases. Apparently, it does not descend from some ad hoc innovative theory formulated for the fight against cancer, as it, instead, originated from unwilling observations. In the early 20th century there have been clinical reports of patients experiencing complete cancer regression after a viral infection. One well-known example, for instance, was that of leukemias cases for which the contraction of influenza was documented to produce beneficial effects [2]. Far from obtaining a successful recovery, it was a first anticipation for a treatment based on the causal use of viral load against malignant diseases. Over one hundred years after first reports, major advances in research made it possible to design and produce viruses, engineered in a more specific and efficient way. Thus, the field gained momentum and it is now considered a promising strategy in the fight against many different tumours.

Let us give a biological overview about the main characters of the topic.

Tumour structure

A **Tumour** is far from being a rigid, isolated and "monolithic" mass; it is a complex, evolving ecosystem and a constitutive part of the **Tumour Micro-environment** (TME). Its composition varies between tumour types, but key common features include immune cells, stromal cells, blood vessels, and **Extracellular Matrix** (ECM), encapsulated into a reciprocal network where the host environment is not a "silent bystander", but rather an active driver of progression [3]. This dynamic interplay supports survival and metastasis, transforming the TME into a continuously evolving entity. In particular, stroma is a supportive tissue, responsible for providing nutrients and with a structural or connective function, whereas the ECM is a large network of proteins and other molecules that surround and support a cell. While both are accountable for tissues growth, the latter also helps cells attach to, and communicate with, nearby cells, and plays an important role in cell movement. The ECM is also involved in repairing damaged tissue, therefore helping in the development - or suppression - of certain diseases, such as cancer.

As a result, TME is considered a fundamental target for therapeutic plans. Until the 20th century, therapies consisted only of excision surgeries; treatments such as radiotherapy and chemotherapy have later been introduced, presenting a quite aggressive and untargeted approach. Since side effects have tremendous impact on the body and in particular on the immune system, more localise and effective therapies, able to leave healthy cells undamaged, are needed.

The complex body of high-quality scientific knowledge available revealed cancer to be a disease involving dynamic changes in the genome. That was possible thanks to the discovery of gene mutations: originating from a programmed genotype, some modified traits are being phenotypically manifested. Tumorigenesis - the formation of new tumours - is a multi-step process where active oncogenes and recessive tumour suppressor genes gain clonal advantage.

To better clarify which paths tumours undertake in order to grow and, eventually, spread out, it is compulsory to make reference to some established concepts in research outcome in medical literature. Authors suggest that "*most if not all cancers have acquired the same set of functional capabilities during their development*" [4], of which figure 1.1 presents a conceptual map. As a result, our carcinogenesis understanding shifts to a type of behaviour that actively reprograms cells.

Virus behaviour

A **Virus** is a unique biological entity, often described as being at the edge of life because it possesses genetic material and evolve, yet lack the cellular structure, independent metabolism, and reproductive abilities of living organisms. They are obligate intracellular parasites, meaning they cannot replicate or carry out metabolic processes outside of a host cell. Structurally, a virus is remarkably simple, typically consisting of a genetic carriage (either DNA or RNA) protected by a protein shell called a capsid, and sometimes a lipid envelope.

For a virus to initiate a successful infection, several critical thresholds must be met. While a single virion is theoretically sufficient, a sufficiently large viral load must enter the host to overcome immediate biological hurdles. Beyond quantity, the success of an infection depends also on fundamental cellular properties.

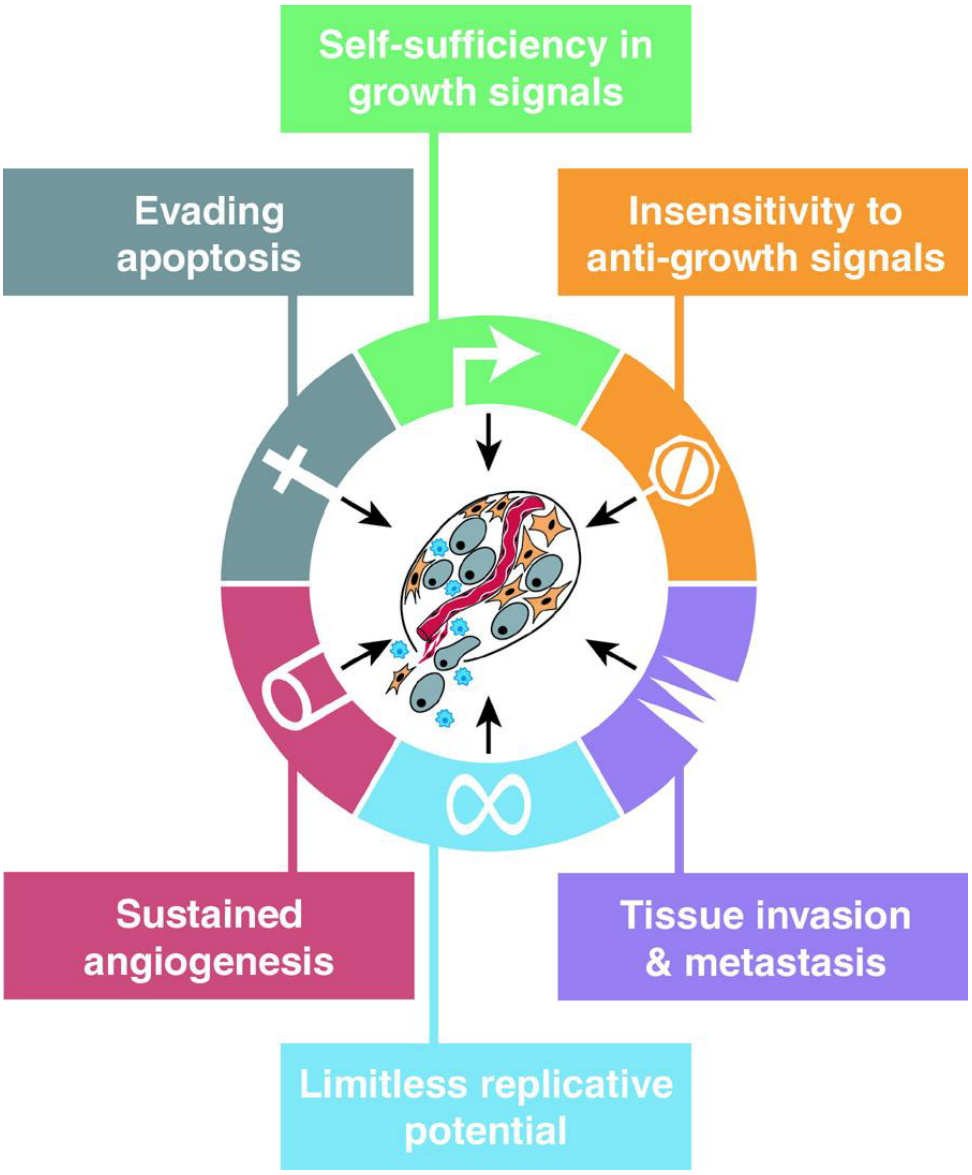


Figure 1.1: "The Hallmarks of Cancer" act as follows: a tumour becomes *self-sufficient* in the generation of its own signals to proliferate, additionally getting *insensitive to anti-growth* signals that normally maintain tissues. Furthermore, damaged cells survive by acquiring a mechanism termed *apoptosis* for preventing programmed cell death. Malignant cells also develop a *limitless replicative potential*, overcoming the natural cellular senescence. Lastly, the *sustained angiogenesis* - the growth of blood vessels - supplies the tumour with oxygen and nutrients, followed by *tissue invasion and metastasis* to distant organs.

First, host cells must be accessible to the virus. Once reached, viral genome has to be internalised into target cells. This process generally follows three pathways: *receptor-mediated binding*, where viral protein spikes act as "keys" for specific receptor "locks", allowing virus to penetrate cellular membrane; *endocytosis*, where the virus gets invaginated within vesicles, exploiting the inherent fluidity of the cellular membrane and *membrane fusion*, where the viral envelope merges with the host bi-layer. Viral binding is possible only if cells are susceptible, meaning that they express the specific surface receptors required for. This cellular affinity for particular tissues is known as tropism, which is the focus of oncolytic virotherapy in the neoplastic mass. Lastly, even if a virus enters, the inoculated cell must also be "permissive", providing the internal molecular machinery and proteins necessary for viral replication. Throughout this process, the virus faces mechanical, chemical, and microbiological barriers, all while an induced activation of the immune system targets its eradication. Once infected, a host cell is often forced to rapidly produce thousands of copies of the original virus, only for being driven to dissolve via lysis and release the genetic material. When not inside an infected cell or in the process of infecting a cell, viruses exist in the form of independent viral particles, or virions.

The Immune System

The **Immune System** is a complex network of cells, tissues, organs, and the substances they produce to help fighting infections and other diseases. Its main function is to recognise and kill antagonistic cells that display specific molecules (antigens) on their surfaces. Immune cells, such as **White Blood Cells** (WBC), are classified in two types: innate immune cells including *neutrophils, macrophages, monocytes* and **natural killer** (NK) cells, and adaptive immune cells mainly represented by **T-cells**. The innate immune system provides the first line of defence against infections and abnormal cells. This defence does not require the recognition of antigens. However, once an innate immune response has been initiated, an adaptive immune response is stimulated, and both work together to eliminate infections or other threats to the body.

Other fundamental therapies against cancer, such as immunotherapies are now part of the clinical options available to patients. For instance, immune checkpoint inhibitors and **Chimeric Antigen Receptor** (CAR) T-cells, revolve around enhancing or exploit immune capabilities for better dealing with illness. Cancer researchers have studied the innate, or nonspecific, immune system, seeking to enhance its potential for cancer immunotherapy. Often, immunotherapies are used in conjunction with OVTs, and their successful partnership depend on viral infections acting as a strategic primer for the immune system. This is capable of transforming a so-called cold immunosuppressive tumours into hot targets for immunotherapy.

The science of *Virology* is relatively young: the first evidence of the existence of viruses came from experiments dating back to the late 19th century. Their understanding accelerated rapidly in the 1950s and 1960s, in large measure because of the advent of cell and tissue culture systems that allowed ex vivo virus propagation. A better understanding of the molecular biology of viral infections and the advancement of **Oncolytic Viruses** (OVs) with genetic engineering has led to the development of OVT. OVs are viruses that can disseminate in cancer cells, and to understand how mathematics can help In order to continue on this topic, it is crucial to present OVs' behaviour.

OVs mechanism of action

For more than a century, scientist experimented with viruses as a form of anti-cancer agents. Interest on the topic has not been stable throughout its evolution: the first approved OVs based treatment represents a milestone that lighted a new spark of fascination in scientists. We shall now focus on the fundamental behaviour of OVT.

The underlying dynamics between an oncolytic virus and a tumour revolves around a few crucial and common steps, for which the **main mechanism** relies on the active action of the pathogen against the malignancy. In the first place, the infection of target cells takes place and the virus efficacy influences heavily any chances of success for the therapy. As previously discussed, virions have different ways of entering a cell, but even admitting we can take advantage of a "well-built" virus able to pass the cellular membrane efficiently, it is still unclear what maximises a positive outcome. The consequent tumour and body response dictates what follows. Due to the complexity of the TME, a plethora of actors are taking part, leading to a complex and often "entangled" scenario. Cell susceptibility to lysis can alter the efficacy, whereas the immune response might act as an inhibitor for the virus spread. The interaction between these phenomena is a crucial aspect for a possible success of the studied OVs. As mentioned,, they are able to transform *cold* tumours into *hot* tumours, more responsive to immunotherapies and therefore less aggressive. Another advantage is that OVs can inhibit or disrupt angiogenesis, leaving cancer cells without oxygen or nourishment. Finally, the virus actively operates to destroy malignant cells by limiting its protein synthesis. Once infected, viruses replicate their genomes inside the host cell, which then undergo cell death through possible various pathways such as apoptosis and necrosis; as a result, the cell cycle ends into lysis and releases viral nucleic acids.

A growing body of evidence suggests the existence of a **secondary effect** from triggering an immune response in the body. After the virions release in cell bursts, the produced materials, such as tumour antigens, allows tumour to be better recognised by the immune system. This can lead to a local response acting on nearby cells, and/or a systemic response to further away locations. That is why some authors refer to OVT also as a type of Immunotherapy [5].

The interaction between OVs and immune system can also lead to negative feedbacks, since emphasising its multiform actions. A heavy immune response may end up targeting virus action, thus limiting its effectiveness [5]. Hence, the need for controlling and improving this dynamic through advanced engineering techniques is crucial.

Strategic bio-engineering

Since solid tumours are complex and heterogeneous systems, research focused on finding ways to strategically intervene on TME and trigger a stronger anti-cancer immune response. Several strategies have been studied and tested to enhance OVs. Through **Genetic Engineering** (GE) many studies focused on crucial aspects: boosting virus dissemination, better degrading tumour cells, stronger control of angiogenesis and improved anti-tumour activity. In order to enhance OVT, many modification strategies aim to increase binding capacity, target accuracy and viral replication have been tried.

A main challenge faced by OVT is the tumour inhibiting T-cells activity supporting its progression and limiting therapeutic benefits. Additionally, a well un-

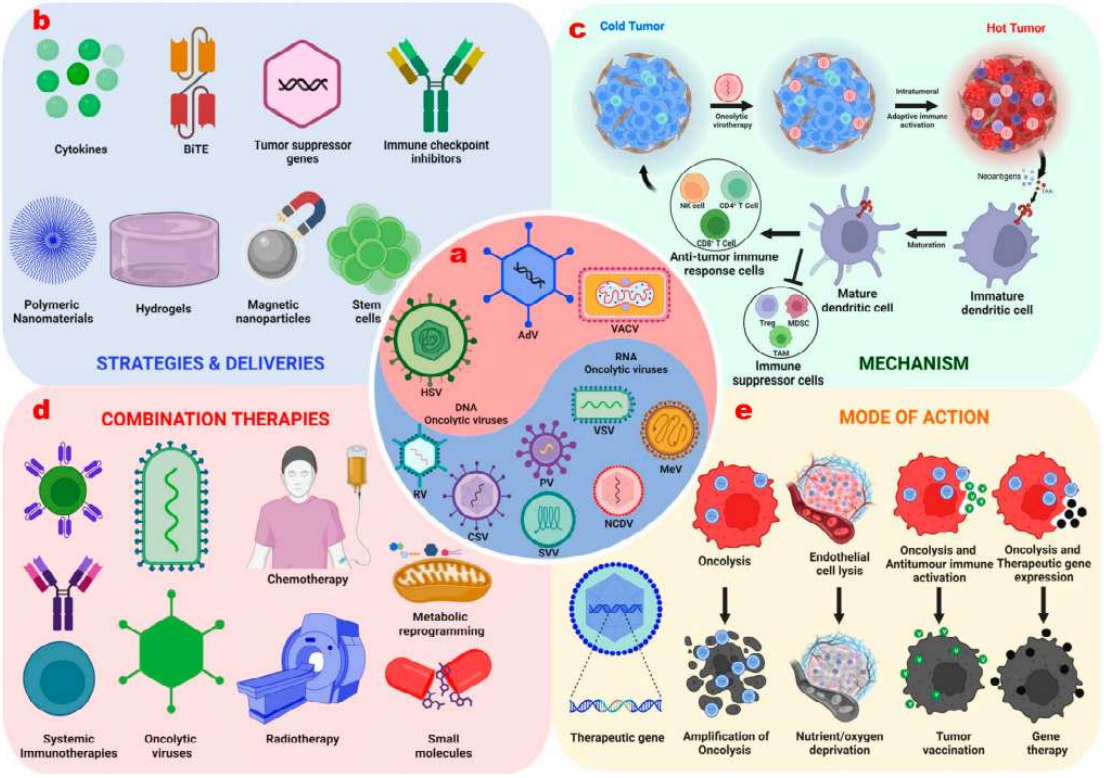


Figure 1.2: Strategies for enhanced OVT. Ideas include Genetic Engineering to combine OVs with other treatments and regulating the delivery process [5].

derstood risk is the potentially compromised individual immune response potential. Some limitations are also related to lack of specificity leading to a virus harming healthy cells, or the possible rise for antiviral immunity, in addition to the difficulties in passing the cell barrier. To overcome these limitations, resources have been invested to optimise the efficacy of therapies. With improvements made in GE and a fundamental understanding of virus genetic biology, interaction with cells, successive apoptosis and lysis, it has been possible to achieve better and more elaborate OVs applications [5]. Results induced improved anti-tumour response and a reduced toxicity to healthy cells.

Existing OVs

A wide range of OVs, including both DNA and RNA viruses, have been developed for cancer immunotherapy. Classified by their genome, a first group contains primarily double-stranded DNA viruses such as **Adenovirus** (AdV), **Vaccinia Virus** (VACV), and **Herpes Simplex Virus** (HSV). The second, RNA group, includes **Coxsackievirus** (CSV), **Poliovirus** (PV), **Measles Virus** (MeV), **Vesicular Stomatitis Virus** (VSV), and **Newcastle Disease Virus** (NDV) [5]. Each of these options shows pros and cons. Some have a safer profile in terms of side-effects, others have a stronger infectivity and have shown more ability to reproduce inside tumour cells. Generally speaking, the first three viria of the list above are the most frequently used and the ones that, also historically, have been associated with better clinical and experimental results.

Stronger approaches in molecular biology of viral infections and advances in genetic engineering have led to the first **Food and Drugs Administration** (FDA)

approved deliveries. It has been the case for a genetically modified form of a Herpes virus for treating Melanoma - well known for being life threatening in case of late diagnosis. A table presenting current FDA approved OVT is presented in table 1.1.

Table 1.1: Overview of commercially approved OVs worldwide.

Name	Year	Cancer	Engineering	Country
Rigvir (ECHO-7)	2004	Melanoma	Unmodified	Latvia
Oncorine (H101)	2005	Head-neck cancer	Modified	China
T-VEC (Imlygic)	2015	Melanoma	Modified	US, EU
DELYTACT (G47Δ)	2021	Malignant glioma	Modified	Japan

Delivery and Combination with other therapies

An effective dose delivery is a key factor for a positive outcome: it is necessary to target tumour without harming healthy cells. Common delivery methods include site-specific, intramural and intravenous injection. For better OVT handling, one must design a high efficacy delivery system; this process is still subject to intense research, but some studies suggest that the use of biological carriers such as stem cells, NK cells and CAR T cells could enhance infection rate [5].

OVs have many advantages over conventional immunotherapies, including precise targeting, effective killing rates, and minimal adverse reaction. They are generally well-tolerated, with most adverse effects being mild-to-moderate, temporary, and flu-like. The most frequent reactions, include fever, fatigue, chills, nausea, vomiting, and injection site pain. Serious adverse events are rare but can include immune-related inflammation or, in immunocompromised individuals, severe infection. Based on clinical data, there have been only few patient deaths directly attributed to the oncolytic therapy itself in over two decades of clinical trials involving thousands of patients, and that occurred for severely immunocompromised individuals.

OVT combination with other therapies has been tested, considering for example the positive influence produced by the synergistic anti-tumour effects of OVs paired with radiotherapy. Future studies may extend this approach with other well-known treatment. Cellular immunotherapy is also considered a candidate for improving OVs action, showing positive results. We will address this topic in depth in a successive section, when describing that we focused on this combination.

To better understand OVT behaviours, we must temporarily trade our microscopes for whiteboards, stripping away the biological surface to reveal the elegant mathematical structure that holds and governs each process.

1.2 Here Comes the Maths

Mathematics is often described as an *exact* science. Within a biological context, however, its exactness lies precisely in its acknowledged *inexactness*. That is to say that any "conclusion" we might draw in this field is nothing but an exemplification and approximation of the real phenomenon. Mathematical modelling does not aim

at predicting biological quantitative outcomes with zero tolerance, but rather at explaining the qualitative mechanisms governing the system. First, it is however appropriate to show why and how we can benefit from it.

The missing link

One may argue: *how do we capture meaningful underlying causal mechanism in such a dis-homogeneous and uncertain environment*? For cancer research it is not always possible to rely on epidemiology for answering most questions with absolute confidence. Studying a large population over a sufficient time frame is not always a possible path for many types of cancer. Even before scaling things up to human experimentation level, at microscopic size is already tremendously difficult to detect and explain patterns exceeding present biological understanding without the help of analytical tools. By learning from key bio-markers, for example, it is possible to study them through the lens of mathematical logic and rigour.

OVT has long been studied before any update in healthcare protocols for human patients. It faced *in vitro* and *ex vivo* trials before any further progress until reaching clinical significance and approval. The scientific hope for finding valid and efficient therapies does not automatically translate into a new result. Indeed, it is usually the case for data to be incomplete or too sparse to eliminate uncertainty. Due to the difficulties derived from this complex multi-factorial biological setting, being able to get a deeper *a priori* understanding of medium to long-term trends becomes very useful. In the last few centuries mathematicians increased their interest in natural phenomena. Many factors contributed: a greater amount of information available from other fields, the needs to make impactful real life decisions and the recent rise of computational power, all gave chance to many applied cases of mathematics.

We now introduce the fundamental research entity, attributing meaning to the initial *inexactness* statement.

Modelling

We will refer to the abstract entity of a **model** as the mathematical language used to describe or simulate the true setting, using one of which is our only rigorous way to infer information from a set of **data**. Being this fundamental understanding the main purpose of research, the way of achieving it may vary. One might get creative with more and more complicated descriptions and involved equations. What mainly inspired this work is try to showcase how minimalist, but smart, models are able to capture key elements on the topic.

The general mathematical-physics approach to derive a model could be stated shortly and recursively in the following few steps: pick a similar, already studied setting, understand its existing model complexity and variables. Hereby, adapting it allows for innovative descriptions of the considered scenario. Obviously there is more room to navigate than these few rules, and that is possible by re-thinking the initial guess adopted. But it is important to understand that in the complex settings, using "first principles" approaches or appealing to "hard reductionism" is unfortunately doomed to fail.

We now present a well known model for "any" growing population - it may refer to the number of individuals as well as shrunk down to cellular reproduction. Our current viewpoint on it is akin to a black box introduction to modelling as a concept, rather than a detailed explanation on this specific dynamic. We use it to point

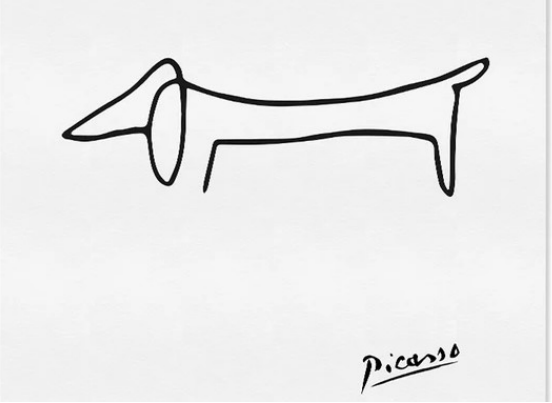


Figure 1.3: A good model for a good boy

out that, in certain cases, "reinventing the wheel" is not necessary if transitioning between models. Specifically, a Malthusian growth law can be generalised to a Logistic framework in a straightforward manner. Trivially, a population concentration at time t is $C(t)$ and can be modelled by a first-order linear differential equation:

$$\frac{dC}{dt} = rC,$$

where r is the intrinsic growth rate This model assumes and idealises an environment with unlimited resources, leading to unbounded exponential growth. The modelling logic evolves by introducing a self-limiting term:

$$\frac{dC}{dt} = rC\left(1 - \frac{C}{K}\right).$$

This second version preserves the original growth trend but changes it through the inclusion of a penalty term that mimics competition for metabolic resources. An alternative, more detailed way for deducing the latter will be presented later on. This simple example captures the logic of universal yet parsimonious mathematical modelling.

A meaningful approach

While traditional models successfully describe trends and patterns in numerous physical contexts, yet they may fail to account for further analysis on more complex dynamics. Such tools could serve as lenses to study mathematical biology, but they are usually treated like mathematical exercises - if not games. For example, current literature often remain silent on the importance of interpretability of key parameters and, even though the approach may keep up for stand alone researching purpose, it cannot be considered complete in an a non trivial field like OVT. A significant challenge we will face is how to deal with elements with low reliability. Indeed, a great disconnect exists between model presentation elegance and the fragmented nature of clinical data. It is important to remark that a meticulous method is needed in every rigorous scientific scenario.

In order to measure how well a system behaves according to data, or better said, how close can we "bend" our model's shape to the real values, we must consider the right key elements. Once we choose a model, we can only tune it by arranging - or optimising - its **parameters**. Bear in mind that attributing values may vary

the output substantially, as is the case for many well-known dynamical systems, for example those susceptibles to chaos and stochasticity. Moreover, **metrics** are essentials for any sort of estimation: they are able to show different sides of the same process. What any researcher fears the most is getting a result that cannot be reproduced because the answers it gave are not significant. In particular, we shall aim at working with models for which a small change of the input determines a reasonably predictable change in the output. We define this characteristic as *stability* and observe that it may get looser with higher complexity.

From Descriptive to Predictive

We propose a data-driven framework designed to better understand underlying dynamics. Our focus consists of the optimisation of existing models on experimental data with a specific purpose: measure OVs in action in given initial conditions and various GE implementation. The novelty of this approach lies in its versatility: capturing trends is not is not always possible, and, as we will show, the analytical and numerical techniques we will explore are capable to show some light on interesting aspects of the topic. The main idea, however, remains simple: data can hide patterns and correlations that can be hopefully investigated through the right mathematical lens.

By then interrogating the stability of the information, this thesis aims to shift from purely describing the behaviour with *a posteriori* analysis to an *a priori* understanding of the phenomenon. In order to do so, appropriate metrics and simulations will be introduced. This fundamental shift can provide clues on how heavily a system is impacted from its initial conditions rather than its evolution.

In order to visualise the magnitude of importance attributed to mathematical modelling in medical literature we cite a very famous example from the work of Murray and Swanson over a really aggressive brain tumour named *Glioma* [6]. By employing a reaction-diffusion equation, they studied tumour growth and invasion. The innovative approach incorporates the effects of heterogeneous brain tissue, in order to measure the spatial diffusion process governed by its proliferation. Thus, they were able to understand the effect of the bulk tumour location, eventually successfully predicting the extent of "invisible" tumour cell infiltration beyond what was detectable via **Magnetic Resonance Imaging** *(MRI)*. That is to say, the model represented a quantitative tool to personalise surgical boundaries and radiation therapy. This breakthrough shifted the perception of mathematical models from abstract theoretical exercises to indispensable diagnostic instruments, capable of visualising the occult spread of disease in the human brain.

1.3 Experimental Setting

A justification for the biological context in which we pose our study is hereby presented. We base our approach on the analysis of given datasets from research papers. By detecting common trends, we aim to classify behaviours depending on biological circumstances. The basic idea is intuitive: study the disease evolution and its interaction with OVT over different models. Datasets are extracted from peer reviewed papers by "semi-automatically" running a software; for reference see [7]. This allows us to obtain original data that we can use to construct models, metrics and parameter sensitivities.

Data and experimental settings

Many treatments follow the same type of testing before human experiments and OVTs are often investigated in the lab using animal testing. There are many types of relevant articles addressing that idea; what distinguish one from the other are the techniques involved in preparation, procedures and, obviously, final results. Scientific literature focused on different tumours and various OVs, while all following a defined structure. Indeed, test animals receive doses of tumour inoculation (*xenograft*) and are either split into control (placebo) *(PBS)* group - not undergoing any therapy - or a group receiving an active treatment. After injection, biologists keep track of progresses, measure tumour sizes[1], and finally present values on appropriate graphs. Since groups include several individuals, results are usually showed with the adoption of average and standard deviation parameters of the group responses. Each point typically lies on the coordinates (t, V), which refer to average $V =$ **Volume** of tumour of the considered group at **Time** t. The common duration of these experiments varies between 30-70 days.

Once implanted, tumours tend to grow and reproduce rapidly. A "weak" therapy might not be able to affect such growth and a clear signal of inefficacy is the mass size reaching a certain threshold. If that happens within a certain period of time, the experiment, and the animal, are terminated. Indeed, in vivo experiments are typically concluded when the tumour volume reaches a measure of approximately $1000mm^3$, to comply with ethical guidelines and ensure data integrity. Beyond this value, cancer shifts from manifesting active proliferation to undergoing central necrosis, thus changing its structure and physiology. By becoming a *solid tumour* and outgrowing its vascular supply, the core of the tumour dies due to *hypoxia*, leading to unpredictable growth dynamics. Terminating the study at this volume prevents unnecessary animal distress and ensures that the observed growth curves reflect the biological interaction between the virus and viable cancer cells.

For instance, we take Ref.[8] as a paradigm for this type of data; a snapshot of it can be found in figure 1.4. While induced viral infection may vary between experiments (number of doses, frequencies, delay between shots), we remark how the starting point is usually a fixed initial tumour size of approximately $100mm^3$. At this point, the cancer usually show a small but noticeable decrease until reaching the nadir, which is its lowest point. A trend is then formed depending on the efficacy of the therapy. We aim at graphs with light slopes and as little peaks of acceleration as possible. Inflection points appear to signal important qualitative shifts in the model's trend. This could be linked to changes in the reproductive ability of the tumour cells or indicate a particularly successful tumour infiltration by the immune system and/or the virus. . A quite common, non positive outcome is given by reading a steady growth, with a lack of response to therapy. In that case, studying key parameters may reveal hidden but meaningful explanations. It is also at least interesting to consider the impact of tumour growth at survival level in those cases and compare them with successful responses. A plot showing different outcomes can be found in figure 1.5.

Oncolytic AdV is widely thought as a promising OVT, as it exhibits tumour-selective replication, high rate of viral production and a potent *cytopathic* effect (damage to the host cell). Nevertheless, it is limited by its native tropism. In order to magnify its therapeutic efficacy, several delivery methods mutation have

[1]This step is typically performed via approximation with specific formulas, such as $V = 0.523LW^2$.

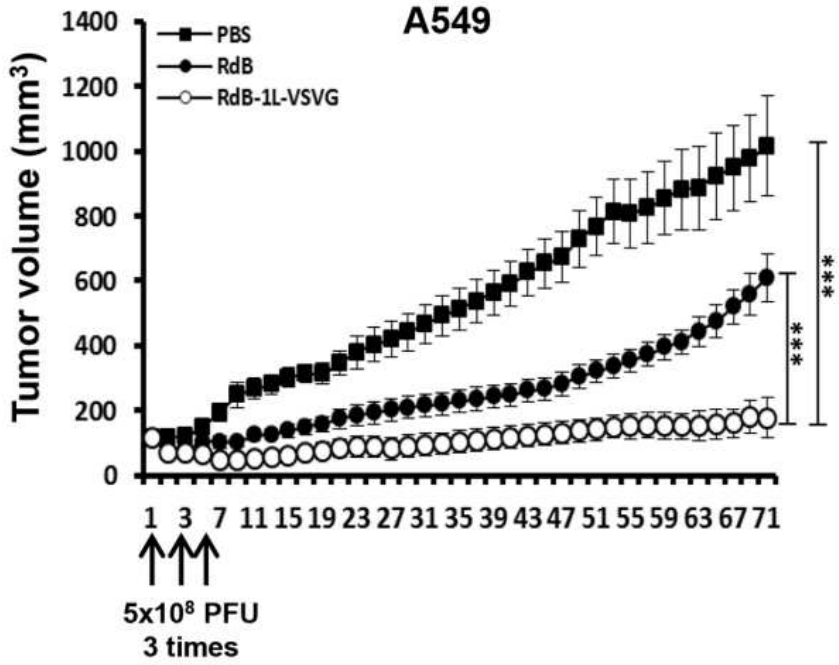


Figure 1.4: Anti-tumour effect of oncolytic Ads in tumour xenograft model, taken from Ref.[8]. This particular experiment focuses on **A549** lung cancer cells with a GE Ad. On the x-axis, days. A **Plaque Forming Unit** (PFU) is a measure of the number of infectious virus particles in a sample, determined by counting the visible plaques (clear areas of cell death). Every group consisted of 7-8 animals. Tumour volume was estimated with the formula: $V = 0.523LW^2$.

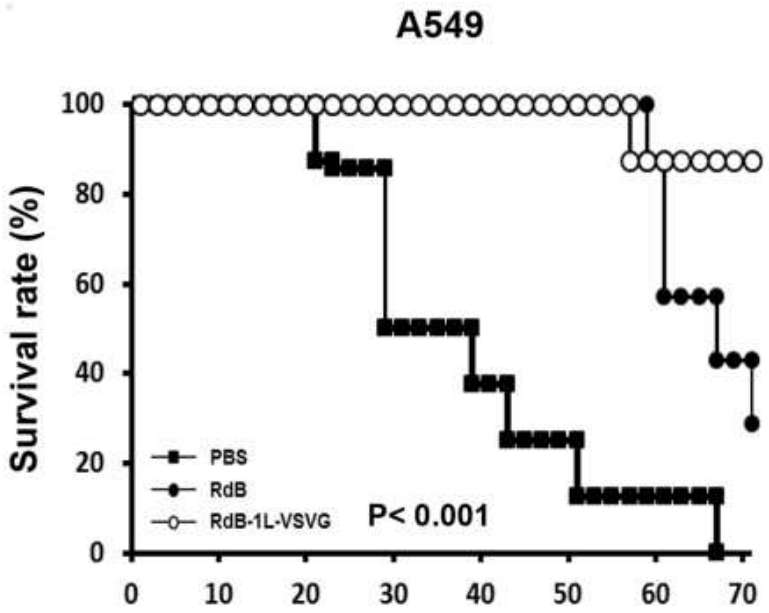


Figure 1.5: Survival curve analysis with Kaplan-Meier plot. This graph is showing each group's individuals survival; each "step" describes the decrease in chance of survival over time. For more details consult [8].

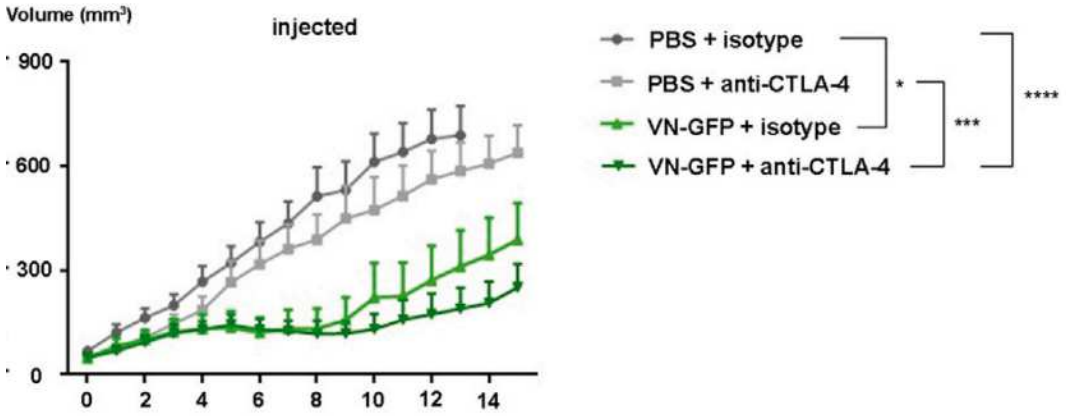


Figure 1.6: Anti-CTLA-4 therapy synergies with recombinant VSV-NDV successfully delaying tumour growth in a melanoma mouse model, taken from Ref. [9].

been studied. For example, GE techniques have been used, such as **Double Mutant Oncolytic** AdV (RdB). A further enhanced virus, encapsulating a **Vesicular Stomatitis Virus Glycoprotein** (VSVG) *epitope*, contains part of an antigen that is recognised by the immune system. Expanded tropism can improve Ad's infection to cancer cells and enhance its therapeutic applications.

On the mathematical side of things, we concentrate at plotted outcomes. In the considered example, the enhanced VSVG Ad is able to overcome the tumour reaction, by successfully preventing it from reproducing over threshold level with a statical significance. The time length for which the therapy was able to keep growth under control is promising and calls for a deeper analysis on the studied mutations.

Types of experiments and datasets

Previous examples are characterised by a *mono-phasic* trend for each group of data. Even though a clearly more successful one exists, a no-decreasing shape is universally seen, but we should admit possibility for more complex behaviours. Taking for example Ref. [9] , advanced GE techniques have been involved and tested, the therapeutic value of an immune-stimulatory OV to sensitise tumours to **Immune Checkpoint Inhibition** (ICI). ICI is a type of cancer immunotherapy that blocks inhibitory signals on T-cells, allowing the immune system to recognise and attack tumour cells. Multiple factors continue to limit their use, but OVs offer the promising possibility to potently synergies because of their multiple mechanisms of action. This novel approach is a case of a new "family" of *chimeric* therapeutic agent. By showing the efficacy of OVT in combination with other therapies, we are able to reach a stronger anti cancer reaction that, even though it may not be able to long-term control tumour growth, is able to reduce growth significantly. This trend is a promising outcome since it proves how malignant proliferation can be delayed significantly. As we can see from the plot in figure 1.6, a *bi-phased* behaviour is present, leading to non trivial evolution, especially in the chimeric OVs group. More examples of OVT experimental data are those of figures 1.7, 1.8 and 1.9.

We can mathematically distinguish research results based on their potency, frequently coinciding with the biological complexity involved in the studied system. We refer to paper with simpler OVs such as the one in figure 1.4 as ***Category I*** and to the more advanced ones where multiple therapies are embodied as the ex-

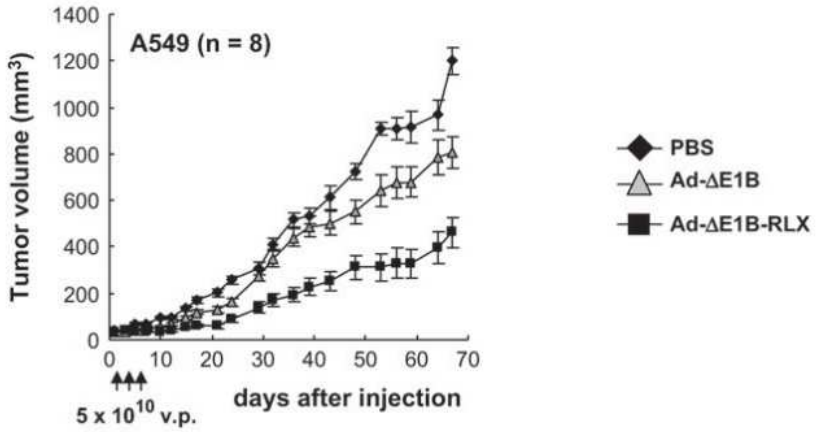


Figure 1.7: GE relaxin-expressing AdV (*VE)* cohort compared with standard AdV (*V*) and control (*PBS*) groups against *A549* cells. Data extracted from Ref. [10]

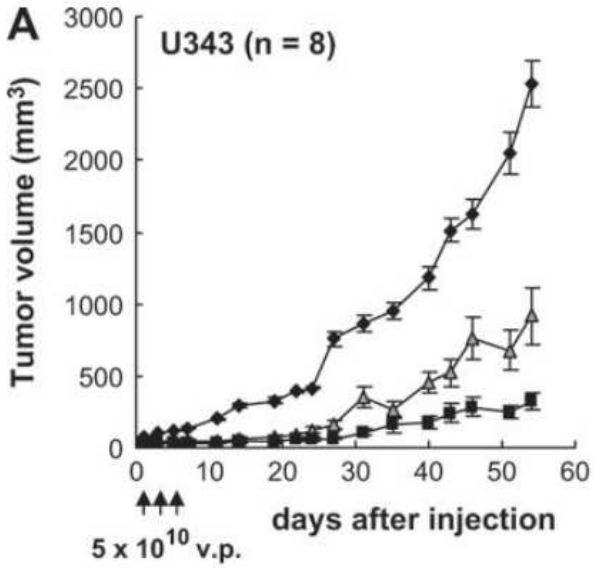


Figure 1.8: *U343* cohorts. Since OVs delivery is performed in the first 5 days of the experiment; data are taken from Ref. [10].

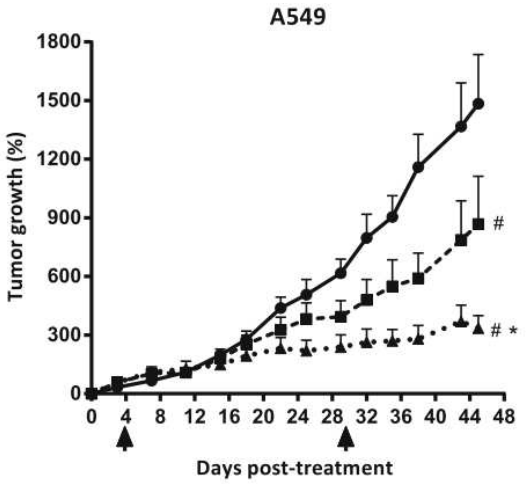


Figure 1.9: Datasets from Ref. [11].

ample just discussed as ***Category II***. We remark that, due to the differences in setting, tumour type, injection modality and frequency, datasets cannot be compared with exact one-to-one measurements, but rather in a *qualitative* fashion. We are interested in growth curves and meaningful interactions, while a promising final tumour size may reveal important insights. We therefore aim to show how different therapeutic approaches could be well explained through appropriate models, and their efficacy measured by using appropriate mathematical analysis.

All the experimental datasets and the MATLAB (R2025b) implementations of the developed functions are provided as an original suite of codes, accessible at: `https://github.com/francescoluppigit/CancerTherapiesModelling`.

Theory put to the test

It would be naïve to think the body-tumour interaction is only related to exogenous therapy, since many other factors are involved. However, being able to understand main trends sheds some light on new experimental directions and help understanding existing hurdles. In order to do so, as a first step we approximate therapeutic effects taking place between OVs as a single entity and the cancer itself. We shift our focus on the total tumour size, temporarily ignoring its complexity and the diffusion of infected cells that act with different purpose. By adopting this approach we are able to describe a first family of model and undergo analysis through various dataset.

It is important to stress repeatability aspect of things: since every experiment follows a general structure but implements it through various steps, we, again, stress the significance of calculations. A model might not be able to estimate the exact expected growth, but should aim at capturing main trends. While bearing in mind our goals, we shall also put data to the test to see if hidden correlations occur. Studying the distribution of errors, also gives us the possibility to reach a critical and objective interpretation of the phenomenon.

The following chapter describes a family of models consisting of a single **Ordinary Differential Equation** (ODE), as a first step for a better understanding of OVT dynamics, and how their solutions help in characterising experimental trends.

Chapter 2

One-dimensional ODE Growth Models

In an effort to balance biological realism with mathematical parsimony, we address general populations growth models that can be also applied to cellular dynamics. We can leverage their analytical solutions to draw conclusions about parameter sets and consequent behaviours.

2.1 Shape and Constraints

As we already mentioned, cancer cells tends to grow at an exponential rate if let free in a Petri dish cultivated tissue. A consequent trivial modelling approach would be thinking of the dynamic as a classic Malthusian law:

$$\frac{dC}{dt} = rC, \tag{1}$$

that can be solved by separating the variables and integrating both sides. We calculate:

$$\int \frac{1}{C}\, dC = \int r\, dt \implies \log(C) = rt + k.$$

Applying the exponential function to both sides and considering the initial condition $C(0) = C_0$, the general solution is:

$$C_E(t) = C_0 e^{rt}$$

where r represents the *intrinsic growth rate*. An example of exponential shaped data is presented in figure 2.1.

Exponential limitations

This first idea soon reveals a plethora of limitations related to the absence of any growth constraint. One shall admit the chance for an interaction between malignant cells and an exterior entity acting against it, whereas an exponential curve responds to no such limits. Indeed, by taking into account experimental data from clinical trials and experiments, they typically show a better overall "handling" of the disease. Even though early-stage tumour growth might employ an exponential rate, real

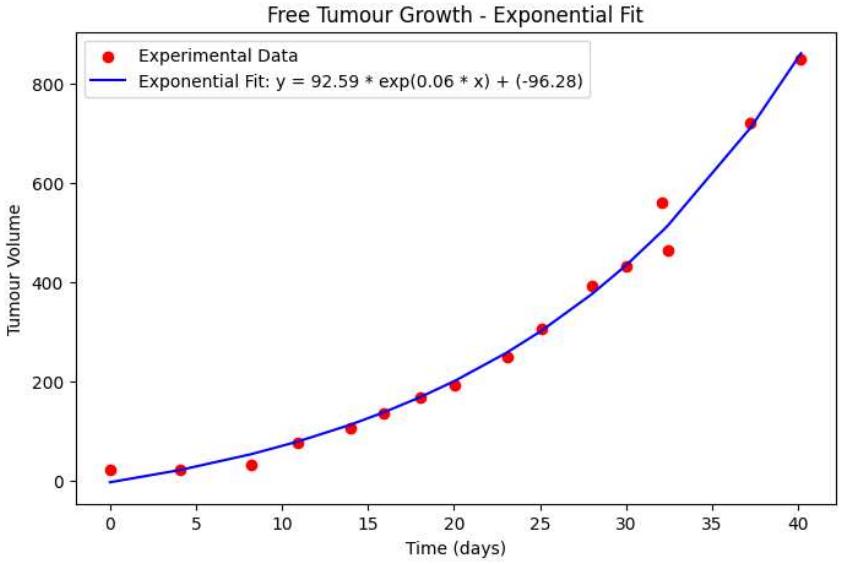


Figure 2.1: Experimental data tested with an exponential fit, taken from Ref [12]. Parameters optimisation has been conducted with the *curve_fit()* Python function.

biological systems tend to face constraints. The way in which the TME is acting heavily depends on available resources: location has to be favourable, nutrients are determined by vascularisation and a type of **intra-species competition effect** is present, among various elements of the environment. Since, as we presented in the previous chapter, virions do not replicate until infecting the targeted cell, neither they act against it, we consider, as a form of mathematical exemplification, those inoculated cells to be the main tumour antagonist. This exemplification, temporarily ignoring other factors, such as the Immune System, serves us as a lens to read the main framework. A fight for survival begins as soon as the viral load successfully starts interfering with the (no longer) "solitary" tumour dynamic. This behaviour forces us to include terms in our equations able to model a more realistic cancer growth.

Capable of deducing mono-phasic growing trends, the exponential model is not suitable for multi-actors mediated constrained situations. That is to say that even by tuning the parameters is not possible to achieve a fundamental change in its behaviour: it can only delay its rise slightly or flect to be less steep. In every case, it will eventually "explode" without any *upper bound*. While easier to describe, it is clearly and demonstrably not desirable for positive outcomes in patients: propagation is positively and directly related to the worsening of patient conditions. Since a *good* experimental result shows more complex and mixed behaviour, this first given approximation does not permit to deduce sufficient information.

An important remark has to be made: an isolated therapeutic agent shows different response compared to the full system combined action. Indeed, due to intrinsically present mechanisms, biologists expect an induced effect, similar to the *invasive species* dynamics of well known systems. This species effect translates into a reduction in malignant proliferation due to competition on resources, until reaching a *plateau*. That is why historically other shapes have been considered for describing growing population dynamics and we aim at modelling two (or more) factors in the frameworks, using ODE sigmoidal models.

2.2 First Approach: Sigmoidal Models

Setting ourselves in a continuum modelling framework, we define $C(t) \geq 0$ as the *population concentration* at time $t \geq 0$. A regulating parameter is $r > 0$: its intrinsic *growth rate*. Let us then introduce a class of mathematical models defined as:

$$\frac{dC}{dt} \coloneqq rCf(C). \tag{2}$$

Their name is obviously referring to the trajectories they are able to describe: an horizontal 'S', with an inflection point whose position determines, according to the related parameters, how the curve approaches the characteristic plateau.

The discriminant element for the family is $f(C)$, the so-called **crowding function**. It allows to include and regulate the competition effects, limiting the absolute rate of growth as C increases. By taking into consideration $K > 0$ as the **carrying capacity** - the maximum population size of a species that a specific environment can sustainably support over the long term without damaging the habitat - it is considered common practice to choose $f(C)$ as decreasing, in the sense of

$$\frac{df}{dC} < 0$$

while, intuitively, $f(K) = 0$ (i.e., there is zero growth at carrying capacity).

Fundamental questions about model selection and parameter estimation are critical if the intention is to deduce qualitative inference. More details about parameters significancy and ranges will be presented after introducing biologically significant equations.

Logistic

The logistic function was first introduced in the 19th century by Pierre Francois Verhulst as an adjustment to the exponential curve. Its 'S' shaped behaviour can be considered as a paradigm for our analysis. Initially growing exponential, then, as it approaches saturation, it slows down to linearity, and once "mature", reaches its growth limit.

The standard logistic crowding function is:

$$f_L(C) \coloneqq 1 - \left(\frac{C}{K}\right).$$

The term $\dfrac{C}{K}$ describes the proportion of the carrying capacity that has been reached. As the fraction approaches values close to 0, the term itself is considered negligible, thus taking it back to an almost exponential behaviour. The **Inflection Point** (IP) - turnover for slowing down growth - is reached at $\dfrac{C}{K} = 0.5$. If $C \to K$, the dynamic is approaching equilibrium, while values above 1 are considered as overshoot, meaning unsustainability for the system.

By substituting the crowding function into the family equation, we obtain:

$$\frac{dC}{dt} = rC\left(1 - \left(\frac{C}{K}\right)\right) \tag{3}$$

Interestingly, this is a separable ODE:

$$\frac{1}{C(1 - \frac{C}{K})}\, dC = r\, dt$$

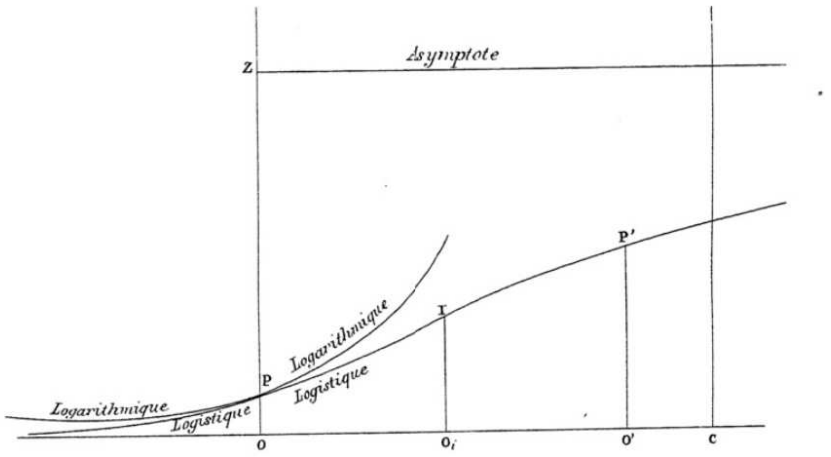


Figure 2.2: "Memoire sur la population par M. P. Verhulst" comparing the newborn "logistique" curve to the classic exponential (here referred as "logarithmic"), taken from Ref. [13].

Using partial fractions:

$$\frac{1}{C(1-\frac{C}{K})} = \frac{1}{C} + \frac{1}{K-C}\cdot\frac{1}{K},$$

so we integrate both sides:

$$\int \left(\frac{1}{C} + \frac{1}{K-C}\cdot\frac{1}{K}\right) dC = \int r\, dt$$

$$\log|C| - \log|K-C| = rt + A$$

$$\log\left(\frac{C}{K-C}\right) = rt + A.$$

Applying exponential to both sides:

$$\frac{C}{K-C} = Be^{rt}, \quad B = e^{A}$$

and solving for $C(t)$

$$C(t) = \frac{KBe^{rt}}{1+Be^{rt}}.$$

Let C_0 be the initial condition, then:

$$\frac{C_0}{K-C_0} = B \quad \Rightarrow \quad B = \frac{C_0}{K-C_0},$$

and we finally arrive at:

$$\boxed{C_L(t) = \frac{KC(0)}{C(0) + [K-C(0)]\exp(-rt)}} \tag{4}$$

We remark that having access to an analytical solution is a clear advantage for any possible fit with available data, as we will show shortly.

The next model incorporates a logarithmic element to the crowding function.

Gompertz

The function was originally designed by Benjamin Gompertz to describe human mortality in actuarial studies, but has later been modified to detail populations. This curve describes a type of growth that is slower at the start and end than any other middle point. The upper asymptote is approached much more gradually by the function than the left-side asymptote. This contrasts with the logistic function in which both are approached symmetrically. Its crowding function is

$$f_G \coloneqq \log\left(\frac{K}{C}\right).$$

In order to solve it, we first make a substitution step: let $u = \log(C)$. Then, $\frac{du}{dt} = \frac{1}{C}\frac{dC}{dt}$. Substituting this into the equation gives:

$$\frac{du}{dt} = r(\log K - u).$$

To solve this separable ODE, we rearrange it to:

$$\frac{du}{\log K - u} = r\ dt.$$

Integrating both sides:

$$-\log(\log K - u) = rt + A$$
$$\log K - u = e^{-(rt+A)} = Be^{-rt} \quad \text{where } B = e^{-A}.$$

Finally we back-substitute by replacing u with $\log C$:

$$\log C = \log K - Be^{-rt}$$

and exponentiate:

$$C(t) = Ke^{(-Be^{-rt})}$$

Using the initial condition C_0, we find $B = \log(\frac{K}{C_0})$. Thus:

$$\boxed{C_G(t) = K \exp\left[\log\left(\frac{C_0}{K}\right)\exp(-rt)\right]} \tag{5}$$

The main difference between Gompertz and Logistic models lies in the former possessing an unbounded proliferation rate as $C \to 0$, in fact

$$\lim_{C\to 0^+} f_G(C) = \lim_{C\to 0^+} r\log\left(\frac{K}{C}\right) = +\infty,$$

whereas in the latter

$$\lim_{C\to 0^+} f_L(C) = \lim_{C\to 0^+} r\left(1 - \frac{C}{K}\right) = r.$$

This result anticipates how the Gompertz law might be non ideal for initial tumour absence conditions modelling and, more generally, to account for extinguished cellular proliferation. Nonetheless, Gompertz growth is often used in fitting cancer data as firstly observed in Ref. [14]

Richards

Named after F. J. Richards, a plant physiologist of last century, this model extends the two previous curves by adding a *shape parameter*; it is also referred to as the *generalised logistic*. The name comes from the introduction of the power β to the fraction in the crowding function:

$$f_R \coloneqq \left(1 - \left(\frac{C}{K}\right)^{\beta}\right)$$

The new parameter controls the symmetry of the sigmoidal curve and determines the exact point of maximum growth, called *IP*, enabling the curve to be more accurate at describing various biological growth patterns. Since every experimental environment owns its unique characteristics, we expect Richards curve to capture diverse dynamics in better ways than previous sigmoids.

In the same fashion as in the others sigmoidal functions, we derive an analytical solution. In the first place, we rearrange the ODE to group the variables C and t:

$$\frac{dC}{C\left[1 - \left(\frac{C}{K}\right)^{\beta}\right]} = rdt$$

Integration on the left-hand side is performed using the substitution $u = (C/K)^{\beta}$ and leads to:

$$\int \frac{dC}{C\left[1 - \left(\frac{C}{K}\right)^{\beta}\right]} = \frac{1}{\beta}\log\left|\frac{(N/K)^{\beta}}{1-(C/K)^{\beta}}\right|$$

Setting both sides equal:

$$\frac{1}{\beta}\log\left|\frac{(C/K)^{\beta}}{1-(C/K)^{\beta}}\right| = rt + A_1$$

Applying the initial condition C_0 to determine the constant A, and solving the equation for $C(t)$, yields the general analytical solution:

$$\boxed{C_R(t) = \frac{KC(0)}{[C(0)^{\beta} + (K^{\beta} - C(0)^{\beta})\exp(-\beta rt)]^{1/\beta}}} \tag{6}$$

If $\beta = 1$, the model reduces to the standard Logistic equation, whereas for $\beta \to 0$ it approaches the Gompertz formula with a $r_G = \beta \cdot r_R$; see figure 2.3 for a qualitative comparison between the three curves.

Models dynamic and parameters

Growth models are often classified by their number of structural, biological parameters p; Logistic and Gompertz's kinetics rely on the value of $p = 3$ (r, C_0, K) and Richards model runs over a set of 4 (r, C_0, K, β).

Parameter	**Symbol**	**Biological Interpretation**
Volume at Time t	$C(t)$	Tumour concentration at time t
Initial Volume	C_0	Tumour initial concentration
Intrinsic Growth Rate	r	Growth rate in the exponential phase
Carrying Capacity	K	Maximum sustainable tumour volume
Shape Parameter	β	Controls the the IP

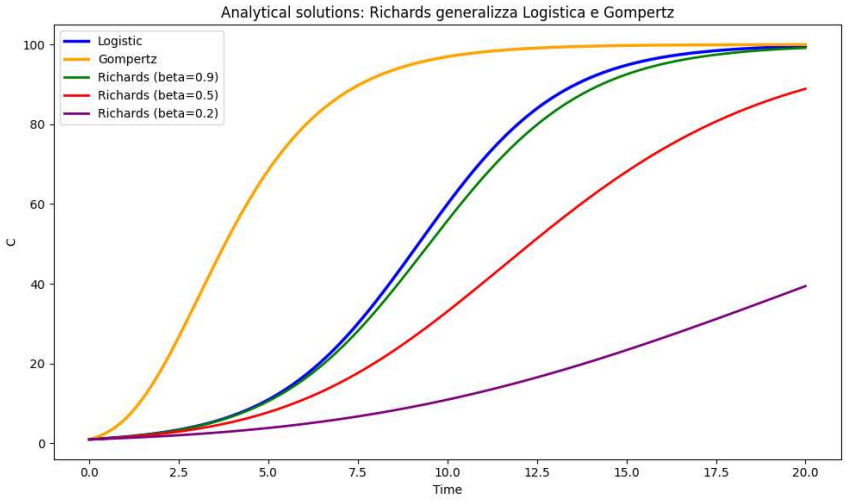


Figure 2.3: Sigmoidal shape according to considered models, plotted through their analytical solution. Generalised Logistic's β is close to 1 (green), thus the curve is almost overlapping with the standard Logistic.

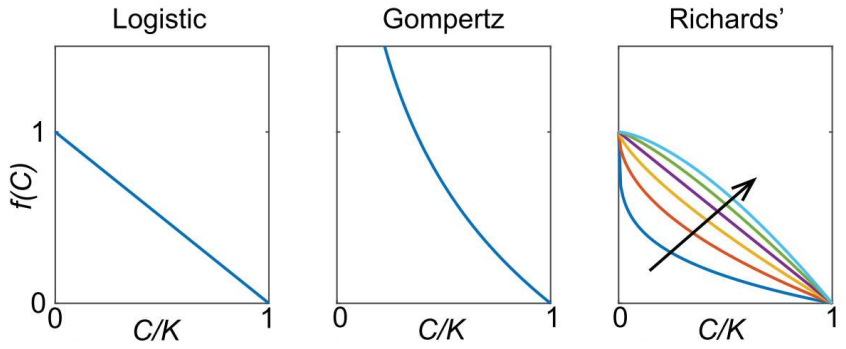


Figure 2.4: Plots of $f(C)$ as a function of $\dfrac{C}{K}$ showing saturation across the three models; Richards arrow pointing toward increasing β values. See Ref. [15] for more detailed parameters analysis.

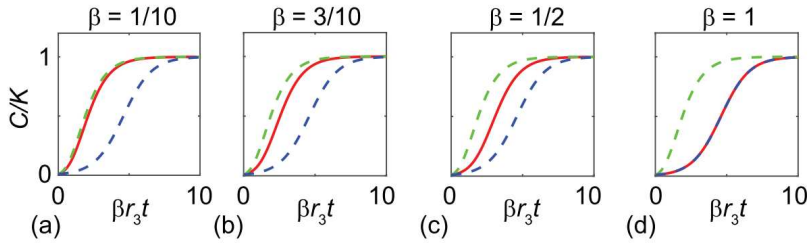


Figure 2.5: Richards model solution for different parameters election. ($r_3 = r$) refers to Richards and Logistic growth rate, while we consider $\beta \cdot r$ as the rate for Gompertz. The x-axis thus consists of the negative argument for the inner exponential term in the solution formulation.

In oncological modelling, the IP represents the critical juncture where the tumour growth rate reaches its maximum value before environmental or internal constraints begin to slow it down. This point is analytically determined by setting the second derivative of the growth function to zero, $\frac{d^2C}{dt^2} = 0$.

For the symmetric Logistic model, the maximum growth occurs at a volume $C_L = \frac{K}{2}$ and a specific time:

$$t_{Ip} = \frac{1}{r} \log\left(\frac{K - C_0}{C_0}\right).$$

In contrast, the Gompertz model exhibits an earlier, asymmetric inflection at $C_G = \frac{K}{e}$ reached at:

$$t_{Ip} = \frac{1}{r} \log\left(\log\left(\frac{K}{C_0}\right)\right).$$

The Richards model further generalises these dynamics through the shape parameter β, yielding $C_R = K(1+\beta)^{\frac{-1}{\beta}}$ for

$$t_{Ip} = \frac{1}{\beta \cdot r} \log\left(\frac{(K/C_0)^\beta - 1}{\beta}\right).$$

Identifying these coordinates serves the essential purpose of quantifying the peak of tumour aggressiveness and the timing of the transition to a resource-limited phase. From a clinical perspective, a delay or shift in t_{Ip} between treated and untreated datasets provides a robust metric for evaluating the efficacy of therapies like viral oncolysis in slowing down the peak proliferation rate.

Crowding functions' trajectory for each model are presented in figure 2.4, whereas figure 2.5 displays Richards' behaviour across several values of the shape parameter.

We now make reference to Ref. [15] on sigmoidal growth models. In that work, Simpson and his collaborators discussed the impact of a *good* choice for parameters on the re-growth of hard coral. The focus is on **parameter identifiability**: whether collected data are sufficient to draw unique and sufficiently precise parameter estimates. The aim of this thesis is to show how, while being often overlooked, addressing identifiability reveals hidden clues on the phenomenon under study. The whole purpose of this chapter is to finally adapt sigmoids according to data and plot a *good* fit, while being able to explain it, rather than just guessing the data shape. The way we "choose" those values undeniably change the final solution and its adherence to the time series. One must also be aware and consider biological constraints. Once we get an insight through optimisation techniques for optimal values, we must be sure about their mathematical significance. Methods that neglect investigating if parameters are identifiable have the potential risk of leading to misleading mechanistic interpretations. Moreover, a final insight is provided by taking into account fit residuals as they help understand how the model performs over time. A frequent tendency to underestimate data suggests, for example, that the curve is not suitable to properly describe that tumour evolution.

As Simpson and collaborators show, a minor change in rigorously studied values leads to an order of magnitude difference in time scale. We will unpack techniques involved in those estimates in the following sections; we just remark that for specific data, simple sigmoidal shaped time series can be accurately predicted by several candidate mathematical models. Only if parameters are identifiable, some models are better than others in terms of not only overall prediction, but also in data consistency.

2.3 Maximum Likelihood Estimation

We leverage an optimisation technique called **Maximum Likelihood Estimation** (MLE), which is a statistical method for estimating the parameters vector θ of a model given a set of observational data $\mathcal{D} \coloneqq (t_i, \mathcal{D}_i)_{i=1}^n$. The core principle is to find the parameter values that maximise the probability of obtaining the observed data.

Likelihood function

When studying a model f, the likelihood function for each of the n independent observations is the probability density function evaluated at data point D_i:

$$p(\mathcal{D}_i \mid t_i, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mathcal{D}_i - f(t_i, \theta))^2}{2\sigma^2}\right), \quad i = 1, \ldots, n;$$

$f(t_i, \theta)$ is the model estimate at time t_i for one specific parameters choice, such that $\theta \in \Theta$ parameters space. The *likelihood function* of the entire dataset is the product of these individual probabilities:

$$\mathcal{L}(\theta \mid \mathcal{D}) = \prod_{i=1}^{n} p(\mathcal{D}_i \mid \theta, \sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mathcal{D}_i - f(t_i; \theta))^2}{2\sigma^2}\right). \tag{7}$$

We remark that in equation 7 the underlying statistical assumption for data is the Normal distribution, which is a conventional choice in the literature.

In particular we look for $\hat{\theta}$, the parameter vector such that:

$$\hat{\theta} = \arg\max_{\theta} \mathcal{L}(\theta \mid \mathcal{D})$$

This $\hat{\theta}$ represents the "best-fit" parameter set and its corresponding likelihood value

$$\hat{\mathcal{L}} \coloneqq \mathcal{L}(\hat{\theta})$$

serves as the global maximum benchmark. However, to simplify computation, we take the natural logarithm of this function, which transforms the product into a sum. This step is valid because the logarithm is a strictly increasing function, meaning the value $\hat{\theta}$ that maximises the log-likelihood (ℓ)

$$\ell(\theta) \coloneqq \log \mathcal{L}(\theta \mid \mathcal{D}) = \sum_{i=1}^{n} \left(-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(\mathcal{D}_i - f(t_i, \theta))^2}{2\sigma^2}\right),$$

also accomplishes the same task for the original likelihood, so that

$$\hat{\theta} = \theta_{mle}.$$

Next, we study the **Negative Log-Likelihood** ($-\ell$) as our objective function to minimise. Having assumed that the set $\mathcal{D}$ is finite, we are lead to find the so-called *MLE estimates* θ_{mle} so that it attains:

$$\boxed{\hat{\ell} \coloneqq \ell(\hat{\theta}) = \min_{\theta} \left[-\ell(\theta)\right]} \tag{8}$$

We recall the **Sum of Squares of Residuals** (SSR): a well known metric for modelling errors, defined as

$$\text{SSR} \coloneqq \sum_{i=1}^{n} \left(\mathcal{D}_i - f(t_i; \theta)\right)^2,$$

so that we can restate the objective function as:

$$\ell(\theta) = -\left[n\frac{1}{2}\log 2\pi + n\log\sigma + \frac{\text{SSR}}{2\sigma^2}\right]. \tag{9}$$

Since the first term of the log-likelihood function is constant with respect to θ, we deduce that

$$\ell(\theta) \propto \text{SSR},$$

so that maximising ℓ is equivalent to studying the quadratic error, while simultaneously achieving two goals. First, it better fits on a curve *vs* data level, due to the minimisation error; second it finds the MLE for the noise parameter σ that best explains the *residual*, e.g. the geometric distance between the data and the curve, dispersion. Since we are studying models reliability, the main focus, even though data can have different interpretations depending on the metric we use, remains residuals, as their dispersion and distribution contribute substantially to global and local meaning of the time series.

Point-wise MLE meaning

The likelihood contribution of a single observation to the total likelihood function can be simply stated as the "probability of having a result from a certain data with a certain parameter". At the level of a single optimised point, the algorithm computes the probability density of the measured tumour volume $\mathcal{D}_i$ by centering a Gaussian distribution on the model's prediction $f(t_i;\theta)$. This operation effectively transforms the residual into a probabilistic weight. When the optimisation method adjusts the parameters to reach the stationary point where the gradient of the log-likelihood vanishes [2], it is mathematically interesting that this process looks for the value of σ that best explains the residual dispersion across the entire timeline. By minimising the individual log-likelihood contributions, the method ensures that the resulting estimator $\hat{\theta}$ represents the configuration where the observed sample variance is most consistent with the assumed underlying noise distribution. In the current context of growth modelling, we assume that the error ϵ_i is independently and identically distributed, and follows a Gaussian distribution with mean zero and standard deviation σ:

$$\mathcal{D}_i = f(t_i;\theta) + \epsilon_i; \qquad \epsilon_i \sim \mathcal{N}(0,\sigma^2).$$

Consequently, the global fit is achieved not merely by minimising Euclidean distances, but by aligning the statistical "peak" of the likelihood surface with the densest concentration of experimental evidence.

2.4 Curves Fitting

With the purpose of showing how MLE is working in practice, let us refer to datasets depicted in figure 1.4 . As we mentioned, mice groups were selected depending on the

[2] The choice of strategy is critical because the likelihood function landscape is often non-convex, containing multiple local minima that might mislead the optimiser. However, we partially addressed this issue through adaptive choice for initial parameters settings, depending on information such as values of $\mathcal{D}_0$, $\max\mathcal{D}_i$ and $\min\mathcal{D}_i$.

type of OVT received, which, for the sake of clarity, have been renamed respectively as "*PBS*", "Virus" (V) and "Virus enhanced" (VE); this last one expresses the GE mutations implemented in the studied Ad. At first glance, with just the help of some mathematical intuition, we can get an idea of which kind of behaviour is taking place: the **nadir** is reached later for the VE group, coherently with the expected immediate proliferation happening in *PBS* mice. The placebo group is the most diverse one: in fact the standard deviation is spreading significantly to roughly 30% in the final third of the time interval. Surprisingly, V data is showing an increased rate of growth, with a clear convex and exponential trend toward the end of the experiment, possibly due to some sort of "momentum" gained by the tumour over viral load.

We conducted an optimisation of parameters analysis for each time series, and plotted resulting curves on the same graph, to illustrate how we evaluate experimental data.

Fit accuracy

We undergo a full MLE MATLAB implementation consisting of a main function reading datasets, and setting initial estimates for then, by specifying the considered model's corresponding likelihood function and bounds. Error is minimised through *fmincon* and we are able to compute best fit parameters, finally plotting corresponding curves.

To give a quantitative idea of the overall model precision for each case we use a well-known metric: the **Akaike Information Criterion** (AIC). It is considered a measure of information loss, since it penalises models for two aspects: poor fit (high residuals) and excessive complexity (too many parameters). Therefore, the model with the lowest AIC is considered the most parsimonious, as it retains the maximum amount of information with the minimum number of variables. In particular, for a given model running over k parameters, the criterion is defined as

$$\mathrm{AIC} \coloneqq 2k + 2\hat{\ell},$$

which describes the trade-off between goodness of fit and simplicity. In other words, AIC account for either the risk of over-fitting and the risk of under-fitting.

It goes without saying that in such a complex biological setting, a single number is not enough to explain the entire phenomenon. It is rather to be seen as a tool and a limitation lies in the fact that it can only be used to compare models between the same dataset, otherwise any analysis that reduces the amount of data would result in a drop in its absolute value. Moreover, available data are of small sample size and we have to make sure that a model is not overemphasising the effects of randomness instead of capturing growth trends. We introduce an enhanced version of this metric in which there is a penalty term accounting for the number of points considered,

$$\mathrm{AIC}_c \coloneqq \mathrm{AIC} + \frac{2k(k+1)}{n-k-1}.$$

The added ratio ensures a consistent measurement in case of interval restriction or simply a "small" number of points (Ref. [16]), with thresholds such as $k > \frac{n}{40}$. Since most of our dataset consists of at most 35 points, it will be consider the main judge for fit accuracy.

It is important to note that the total number of estimated parameters k must include the residual standard deviation σ derived from the MLE. Consequently,

we defined $k = p + 1$, assigning $k = 4$ to the Logistic and Gompertz models, and $k = 5$ to the Richards model. This statistically conservative choice impacts the AIC_c penalty term: we can think of the AIC measure in terms of first order penalisation for k, while its modified AIC_c includes a second order penalisation in the numerator. Therefore, instances where the Richards model still emerges as the optimal choice provide strong evidence that its additional shape parameter β is biologically necessary and not a mere mathematical artefact.

Residuals

Another important factor in this first layer analysis is the corresponding *error*

$$e_i = \mathcal{D}_i - f(t_i, \hat{\theta})$$

which we refer to as **residuals**. When we apply the fit, what we are left with is the distance between predicted points and experimental measures, but if a model is not explaining data well, some features of the time series might be ignored. It is the case for *hetero-skedasticity* to occur, for which waves of magnifying amplitude are shown and it is an indication that noise is gaining momentum, while underexplaining the given data. As a result, later points weight "more" since the error gets amplified and lead to an overall worse fit for cases in which the model is compensating the tails while overlooking initial or inner segments. Ideally, normalised residual should consist of *white noise*: randomly distributed with respect to zero, with symmetry and a homo-skedastic constant variance. This condition also prevent models from *overfitting*, i.e. when the curve shows profiles that follow and fit noise rather than real data trends.

We hence consider **first-order Residuals Autocorrelation**, referred as AR(1) through the **Durbin-Watson** (DW) test:

$$\text{DW} = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}.$$

Values typically range from 0 to 4, where a value near 2 indicates that the residuals are uncorrelated. Values significantly below 2 (typically DW< 1.5) suggest positive serial autocorrelation, implying that the model may systematically under- or over-estimate specific phases of the growth curve. This test should not be taken as an absolute result, but rather like an accessory tool to valuate model consistency. Stiff models, such as the ones we are dealing with, are not suitable for complex behaviour showing several shifts in shape, so we must consider these limitation when conducting AR(1) analysis.

Fit: A549 AdV

Our first fitting attempt is on the dataset of figure 1.4. We classified it as Category 1 class of experiment due to its naturally mono-phasic tendency and the fact that it does not involve combination therapies. Figure 2.6 serves us as a first example. The time interval spreads across 70 days and contains a time-volume reading for every two days, which lead to AIC having a "high" baseline value for each model. A major shift in growth is present in the first 10 days, drastically impacting in overall fit precision. To fairly present data, we also conduct a second analysis by treating those points as outliers and ignoring them.

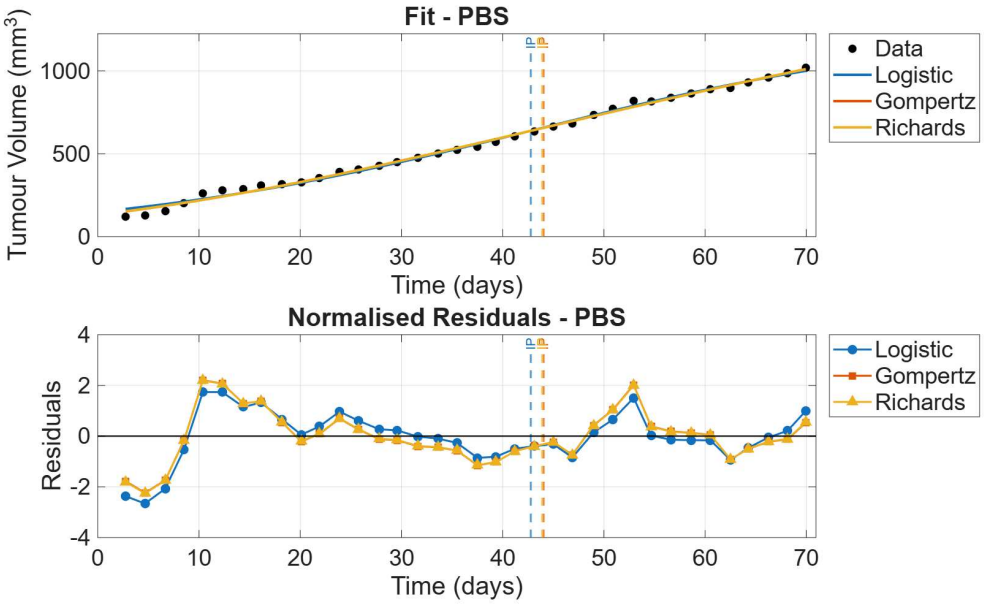
Curves are close to each other, except for the exponential. Vertical dotted lines mark IPs position. We remark that a delayed model prediction of the switch to the second half of the growth slope implies a faster estimated growth rate for the final end of the interval. Interestingly, Richards is squeezed to almost overlap with Gompertz, proving the former adaptability to various context, and indeed we observe a value $\beta = 0.02$.

With the second run of optimisation we are able to achieve a meaningful improvement in relative metric, in addiction to a qualitative forward shift of IPs. Another improvement across every model is a noticeable increase in the DW test. However, a remark has to be made: due to the complexity of data, stiffs model are penalised by the compromise they have to make in the context of fitting. Since *PBS* dataset shape does not reflect an ideal sigmoidal trend, it is hereby expected for some amount of residual correlation to occur, and ideal values larger than 1.5 are not always to be reached. Any progress in the direction of limiting the systemic compounding of error, has to be considered a partial goal.

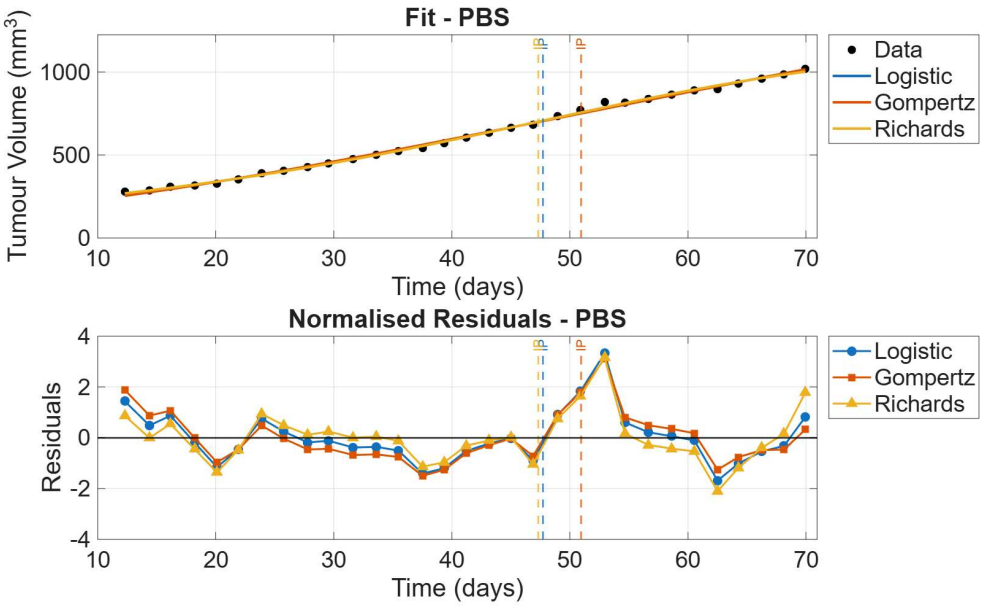
Top dataset in figure 2.7 (a), as mentioned, presents a convex, more aggressive behaviour. What stands out specifically is the AIC_c value across all the models, according to which the exponential curve would be the best guess for this case. We then switched to the corrected metric for "unbiased" results: residuals plotted in the bottom portion of the image confirm that each model is successfully accomplishing a mathematical optimisation task by adapting the fit. Sigmoidal models tend to overlap, with a minor exception for the Gompertz curve revealing an IP at day 62. The reason underneath these similarities lies in the qualitative trend itself: if the optimisation identifies the growth to be sufficiently fast, the carrying capacity gets estimated to be really high thus approximating an exponential growth. In simple terms, we are seeing the first portion of the "S". By tuning the upper-bounds passed to the *fmincon*, we might be able to "hint" models with a more (or less) noticeable rate of increase. For this dataset, setting K as a smaller constraint does not translate in better fits; instead the opposite is true. Raising it from $2 \max \mathbf{D_i}$ to $10 \max \mathbf{D_i}$ yields to noticeable decrease in AIC_c values for both Logistic and Gompertz, while not affecting the Richards model at all.

However, biologically, as we can see from the resulting plot in 2.7 (b), models just shows more of their exponential phase, further delaying the IP. One reason for which the generalised logistic is not acting in the same way probably has to do with the β parameter. That extra degree of freedom implies a broader handling of the fit according to available data, such that the optimisation might trade one value increase for different value decrease. This consideration holds, since β drops significantly from 9.6 down to 7.1 while the carrying capacity reaches 4014 mm^3 - which surprisingly is not as high as both Logistic and Gompertz. The maximum allowed value in those cases is 6148 mm^3. Even though all values carry mathematical meaning, the consequent biological interpretation is erroneous since it would translate to an almost immediate shift toward the saturation phase in the Richards case, or a widely not sustainable tumour proliferation for the Logistic and Gompertz case. The "wave" occurring within residuals leaves no doubt about each model behaviour, alternating under and over estimates with the same, as expected, dynamic as the exponential curve itself. Indeed, DW values are very far away from being optimal. All these signals are consistent with an unreliable data trend.

While previous analysis might appear trivial due to the unviable result we ended up with, it is at least inspiring as an example of balancing between mathematical rigour and biological interpretation. If we were to aseptically look at numbers, even

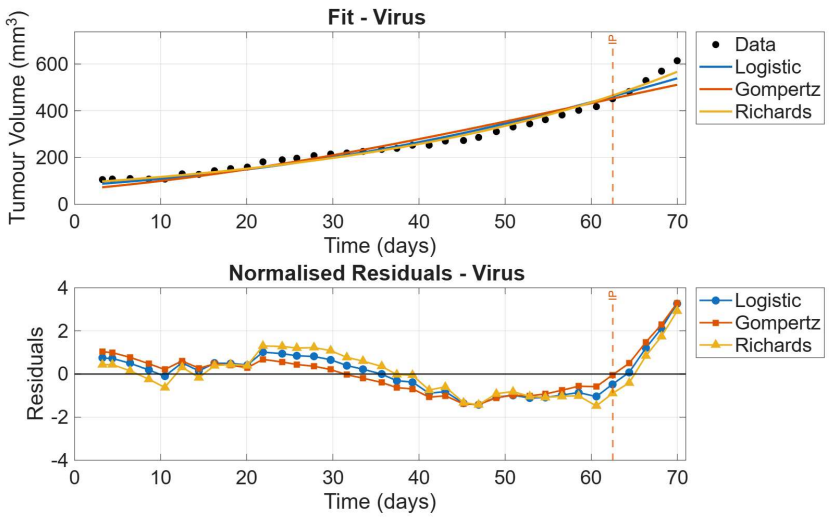


(a) With outliers analysis.

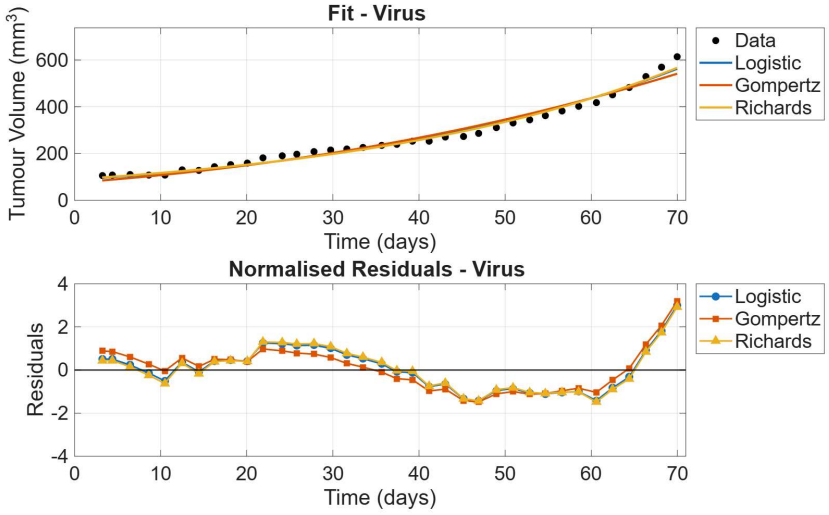


(b) Without outliers analysis.

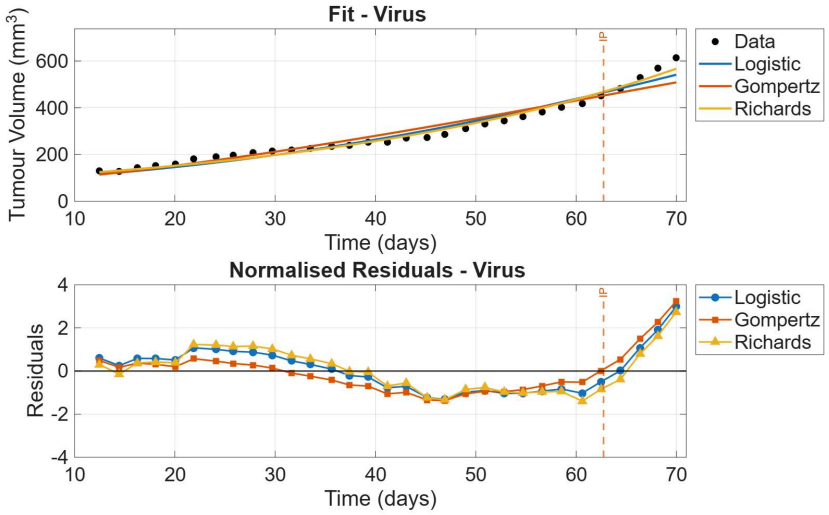
Figure 2.6: Time interval analysis of the *PBS* dataset, taken from Ref. [8].



(a) Pre-nadir analysis on the V group with $K_{ub} = 2 \max \mathbf{D_i}$.



(b) Pre-nadir analysis on the V group with $K_{ub} = 10 \max \mathbf{D_i}$.



(c) Post-nadir analysis on the V group ($K_{ub} = 2 \max \mathbf{D_i}$).

Figure 2.7: Comprehensive analysis of the V dataset, taken from Ref. [8].

allowing further metrics, we would not achieve a significant answer. Indeed, there is no such a thing as a "good answer" for that specific series: experimental data are limited to showing trends and they are not meant or superimposed to follow specific rules. We cannot force a curve to plateau where the data still sees an open horizon, or, as Ronald Coase expressed it, we should be aware that "*if you torture the data long enough, it will confess to anything*". We must not infer information we might wish for, and simply accept the previous exponential shape as not suitable for well specified sigmoidal models.

We point out that in V time-series an initial decrease in tumour size occurs right after the OVs dose; such a trend forces to reaching the nadir at day 11. If we aim at working with sigmoidal mono-phasic behaviours, we shall consider some sort of flexibility: given that the curve will not be able to bend repeatedly, we should compare this first fit to the one of a second, restricted, interval excluding those points. They are not to be seen as outliers, rather as a simple lag required by the OVs to effectively inoculate cells and induce an immune response.

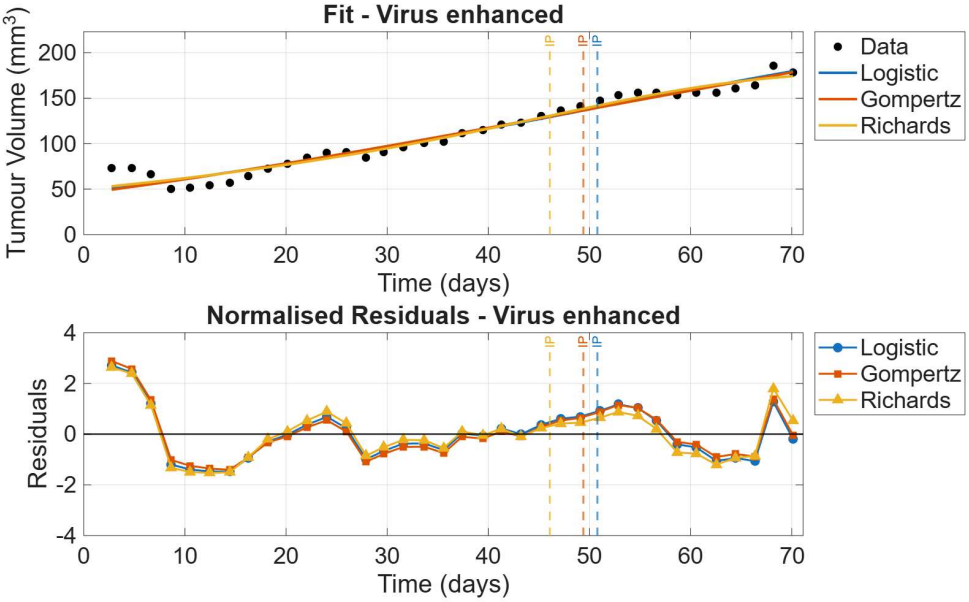
Including a second optimisation attempt with the truncated left-hand tail, as seen in figure 2.7 (c), leads us to some sort of similar behaviour regarding the carrying capacity change. We stick to the $2 \max \mathbf{D_i}$ value, since the approach would not benefit from exponential and overlapping profiles. An immediate difference is the lower relative AIC_c value, meaning that the curves are better able to distinguish noise from information. Furthermore, we can notice a more lay-down profile of the residuals, thus a wider wave with less amplitude, accounting for the observably better fit. This slight change should not surprise us, since dealing with a more coherent time series, as one would expect, leave space for more precise fits. The reason is simple: due to the diminished need to "average" the curve in between points, the distance from point optimisation is able to be more accurate. We remark that on the last days of the experiments, a broader difference between models predictions can be seen.

Finally, we study the flatter, OVT effective, *VE* dataset. We remark that what looks like a major difference in the data profile, has to do to with the stretch in plot $y-$axis caused by the dimensional difference from the previous ones. By conducting the same type of pre/post-nadir analysis, we observe a more marked difference than previously, acting in the same way for each model. Since the left hand side of the (a) fit plot is drastically different in shape compared to the subsequent data, the consequent interval adjustment results in sensibly lower values. This time, what stands out is the qualitative difference in estimates alongside the whole (b) fit. Indeed, the approximately 15 days anticipation of the IP reveals a more pronounced separation of models around day 50. Moreover, models tend to "agree", as a mean of almost overlapping along the entire series, and the DW values, again, show a significant improvement even if not reaching a merely white noise range. Nevertheless, residuals plots are not free from the presence of "waves" in the data.

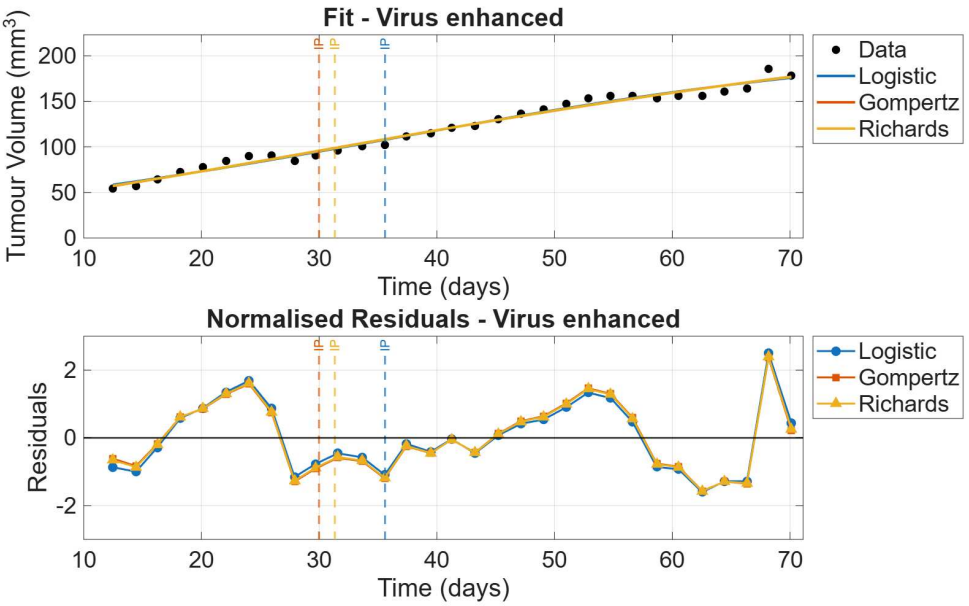
What should be noticed here, is that Richards is achieving the best relative score, in spite the penalisation term, in the overwhelming majority of plots.

2.5 Profile Likelihood

To assess parameter identifiability, we employ the **Profile Likelihood** method as detailed by Simpson and his collaborators in Ref.[15]. This technique is used to determine practical parameter identifiability - that is, whether a finite and noisy



(a) Pre-nadir analysis.



(b) Post-nadir analysis.

Figure 2.8: Comprehensive analysis of the *VE A549* dataset

dataset contains sufficient information to precisely estimate a model's parameters. The name, usually shorted as **profiling**, comes from generated parameters profiles plot.

We work with a *normalised* log-likelihood function, shifted towards a maximum value of zero, so that

$$\ell_0(\theta) \coloneqq \ell(\theta) - \sup_{\theta} \ell(\theta),$$

and study it across various combinations of parameters values. In particular, for each ψ *interest* parameter, we then set the others as η *nuisance* ones, so that $\theta = (\psi, \eta)$. This allows us to study its log-likelihood profile in the following fashion:

$$\hat{\ell}_0(\psi, \eta) = \sup_{\eta} \ell_0(\theta)$$

In practical terms, once calculated the initial MLE fit and parameters, we optimise η for each value of the current parameter of interest. We work by explicitly defining a $\psi^- < \psi < \psi^+$ interval (uniformly discretised by N points) to then pass it to the same *fmincon* optimisation process: this extrapolates a sequence of y-values corresponding to the change in overall ℓ_0, according to the considered combination of values. For example, in the logistic case, studying the growth rate as an interest parameter yields to:

$$\hat{\ell}_0(r) = \sup_{K, C_0, \sigma} \ell_0(r, K, C_0, \sigma).$$

By then symmetrically evaluating the function over a set of sorted parameters values we are able to plot a parameter, log-likelihood graph. Since we normalised the objective function, we expect any of θ to produce a negative ℓ_0 value. In order to practically identify a parameter, we should note worse values of the shifted likelihood, as we go through all the other considered parameters combination; the slope for which this may happen determines the **profile**. Sharper parabolic-like decreases and "thin" curves are crucial for stronger interpretations, whereas flat or wide tails, either in one or both directions, are an indication of uncertainty: an interest parameter may change in size considerably, but the effect on fitting is being hidden by other η values. We call a parameter of the first scenario *interpretable*.

Confidence Intervals

The profile likelihood for ψ is defined as:

$$\mathcal{L}(\psi) = \sup_{\eta} \mathcal{L}(\psi, \eta)$$

To determine the **Confidence Interval** (CI), we employ Wilks' Theorem, which states that under the null hypothesis $H_0 : \psi = \psi_0$, the likelihood ratio statistic

$$\Lambda \coloneqq \frac{\mathcal{L}(\psi_0)}{\mathcal{L}(\hat{\theta})}$$

is such that

$$-2 \log(\Lambda) = 2 \left[\hat{\ell} - \ell(\psi_0) \right] \sim \chi^2_d.$$

In particular, vector ψ_0 has the number of fixed parameters that differs from the one $\hat{\theta}$ by 1. The formula means that twice the normalised log-likelihood converges in distribution to a chi-squared distribution with $d = 1$ degree of freedom, χ^2_1, as the sample size $n \to \infty$.

By considering a $95\% = 100(1-\alpha)\%$ confidence ($\alpha = 0.05$), the critical value is

$$\chi^2_{1,0.95} \approx 3.84$$

Consequently, the confidence interval for ψ is the set of values such that:

$$\hat{\ell} - \ell_0(\psi) \leq \frac{3.84}{2} = 1.92 \tag{10}$$

In our implementation, we plot the normalised profile likelihood $\ell_0(\psi)$. The extrema of the confidence interval are determined by the intersection of this profile with the constant threshold. If a profile is flat, the CI may not be determined by this threshold but rather by the *a priori* bounds of the parameter (e.g., $C_0 > 0$), which is a clear sign of non-identifiability.

We remark that these considerations apply both to the uni-variate (single interest parameter) profiles, just like it would for bi-variate (multi parameters) profiles, with a coherent choice for a mesh of NxN points. We limit our analysis to the former case, with a special uni-variate combination of parameters in the Richards case.

A549 Profiles

We present a complete summary of results with respect to figure 1.4 datasets. We studied parameters profiles for each model, considering a sequence of $N = 100$ points, but denser grids help reducing the cornered look of some graphs. Figure 2.9 shows plots for each $\psi \in \{r, K, C_0\}$ in the Logistic model. Note that the focus is on the post-nadir interval only.

As a result of the exponential trend commented in previous sections, we expect the carrying capacity of the V dataset to be practically non identifiable. By considering the corresponding profile, the $\hat{l}$ value falls right on to the upper bound set for the parameter itself in our implemented *initialise_param* MATLAB function; the profiling is now forced to follow along the only allowed values, and get stretched out widely, specifically pointing to possibly greater upper bound values. By repeating the run on much bigger allowed K values, the shape does not get affected, therefore suggesting the parameter is simply mathematically correct, but biologically not relevant. Regarding every other parameter, we end up with a good identifiability profile and, in addiction, values are kept into significant biological ranges. The profiling is coherent to the fit and confers it a more trustworthy look. A similar qualitative view is offered by the analysis conducted on the Gompertz model, with some minor differences: r values are attained generally for smaller ranges, and a general asymmetric decrease on the second half of the data appears, especially for C_0. It might have to do with the shape difference itself, since the asymmetrical Gompertz model present a IP occurring earlier than the exact half, as it happens in the Logistic. Moreover, the MLE estimates for the initial tumour concentration are greater for the Logistic curve, while interestingly in both cases the corresponding value is inferior than the actual series initial value.

In their work, Simpson and collaborators consider C_0 as a fixed value, giving better overall fits and reduced AR(1) values according to the DW test. However, in our setting, this method did not produce the same improvement, especially giving the worsening effect on the plot due to the range of initial values preferred in the present run of optimisation. Our hypothesis is that a greater amount of points require more flexible data handling.

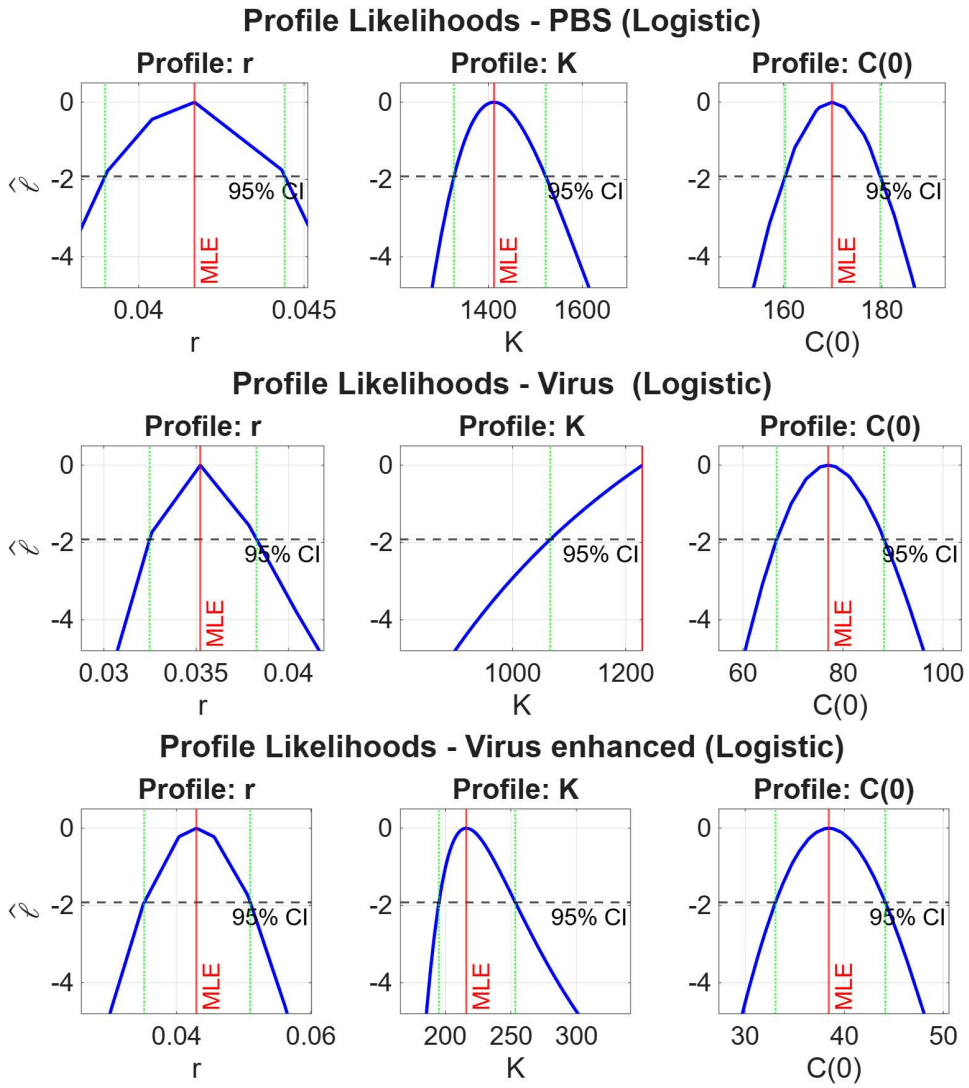


Figure 2.9: Comprehensive Logistic profiles of *A549 AdV* datasets, obtained using the data in Ref. [11].

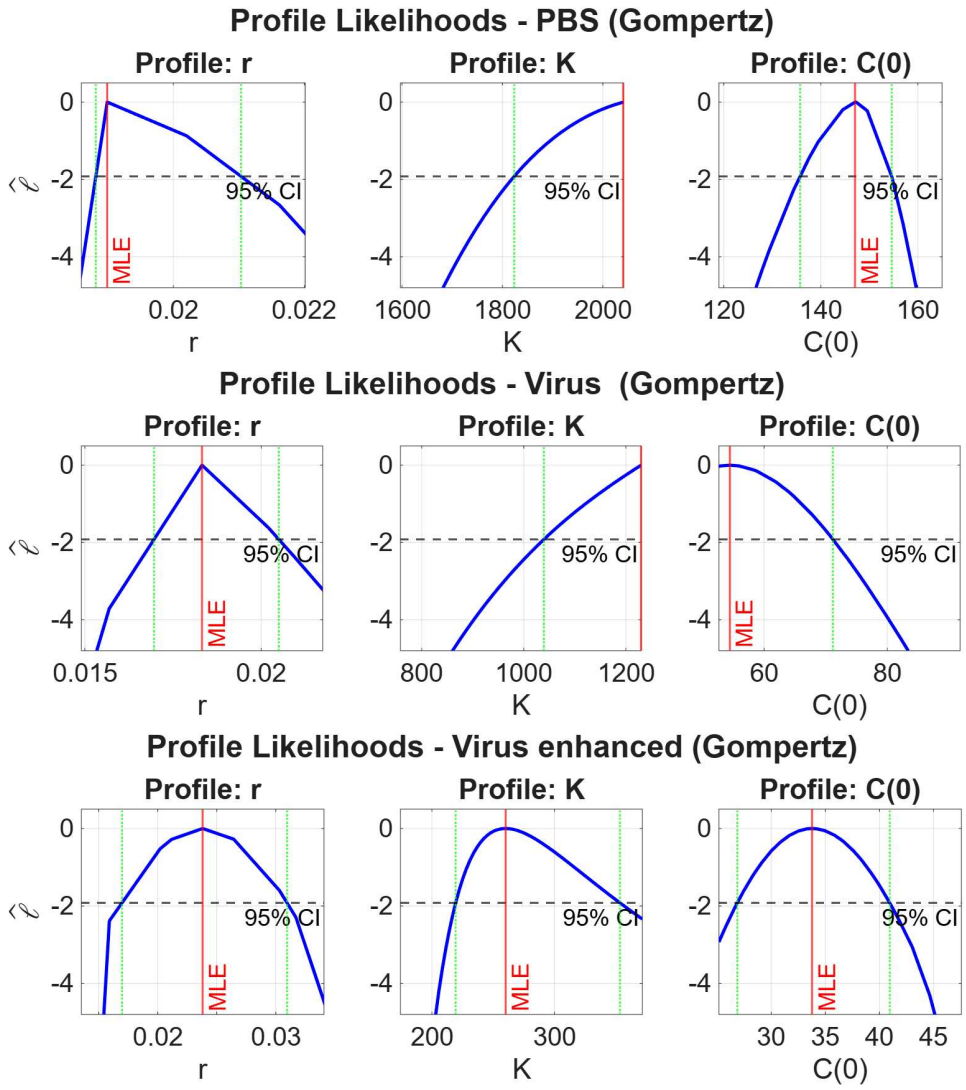


Figure 2.10: Comprehensive Gompertz profiles of *A549 AdV* datasets.

Finally, we undergo the analysis on the more complex Richards model. An important remark is to be made about a key difference: the presence of a shape parameter β. As we mentioned, the intent here is to keep a careful, parsimonious approach and it is compulsory to weight whether the inclusion of one extra degree of freedom allows for better biological interpretation or it is just extending a misspecified method to noisy unreliable data, with the use of heavier computations.

In our analysis, we explicitly include the noise parameter σ within the profile likelihood framework, in order to give a physical measure of the model, rather than dimensionless metrics alone. Although σ is a nuisance parameter that describes the statistical error rather than the biological growth, its identifiability is crucial: a flat profile would indicate that the model is unable to distinguish between biological signal and experimental noise, undermining the reliability of all other structural parameters. The profile likelihoods formulations are the same for every single parameter with the inclusion of β:

$$\hat{\ell}_0(r) = \sup_{K,C_0,\beta,\sigma} \ell_0(r, K, C_0, \beta, \sigma).$$

In figure 2.11 (a), parameters manifest a quite sharp profile, even though not as narrow as the others sigmoidal models. Overall, we are led to a clear and causal view on parameters choice, where any $\theta \neq \hat{\theta}$ choice undeniably leads to a worse ℓ_0 results.

Furthermore, we investigate the uni-variate profile of the combined parameter $\psi = \beta \cdot r$ (the product of the shape parameter and the growth rate) often referred to as the **effective growth velocity**. The name descent from the fact that the product βr measures the maximum slope of the curve in its IP. In the Richards model, β and r often exhibit strong functional correlation, leading to elongated, "open" profiles where the individual parameters cannot be uniquely identified. By profiling their product, we assess whether the model captures a stable effective growth velocity, even when the specific symmetry of the curve (dictated by β) remains uncertain due to the lack of a clear saturation plateau in the data. In figure 2.11 (a) plot is coherent with the profile of each individual parameter, but we attribute more importance to it for (b,c) profiles. As we mentioned, K is not identifiable and, contrary to what happens in Logistic and Gompertz, other Richards parameters get affected, by trying to compensate for it. All the V profiles are clearly not ideal in terms of reliability, but a specific observation has to be made about the last dataset. While growth rate is undeniably flat on one side, and shape parameter is pretty stretched out with a slow decay in ℓ_0, their combination interestingly has a CI width of 0.1, similar to the one from *PBS* data.

Comparisons and conclusions

Some general conclusions that hold for all presented results are possible to be drawn. Richards exhibits broader profile likelihoods compared to the Logistic and Gompertz models, reflecting the inherent functional correlation between the shape parameter β and the growth rate r. This widening of the confidence intervals is a hallmark of increased model flexibility at the expense of individual parameter identifiability. However, the stability observed in the profile of the combined parameter product $r \cdot \beta$ alongside the superior performance in residual analysis and AIC_c scores, suggest that the Richards model provides a more faithful representation of the underlying biological dynamics of the presented datasets. The main drawback for the generalised

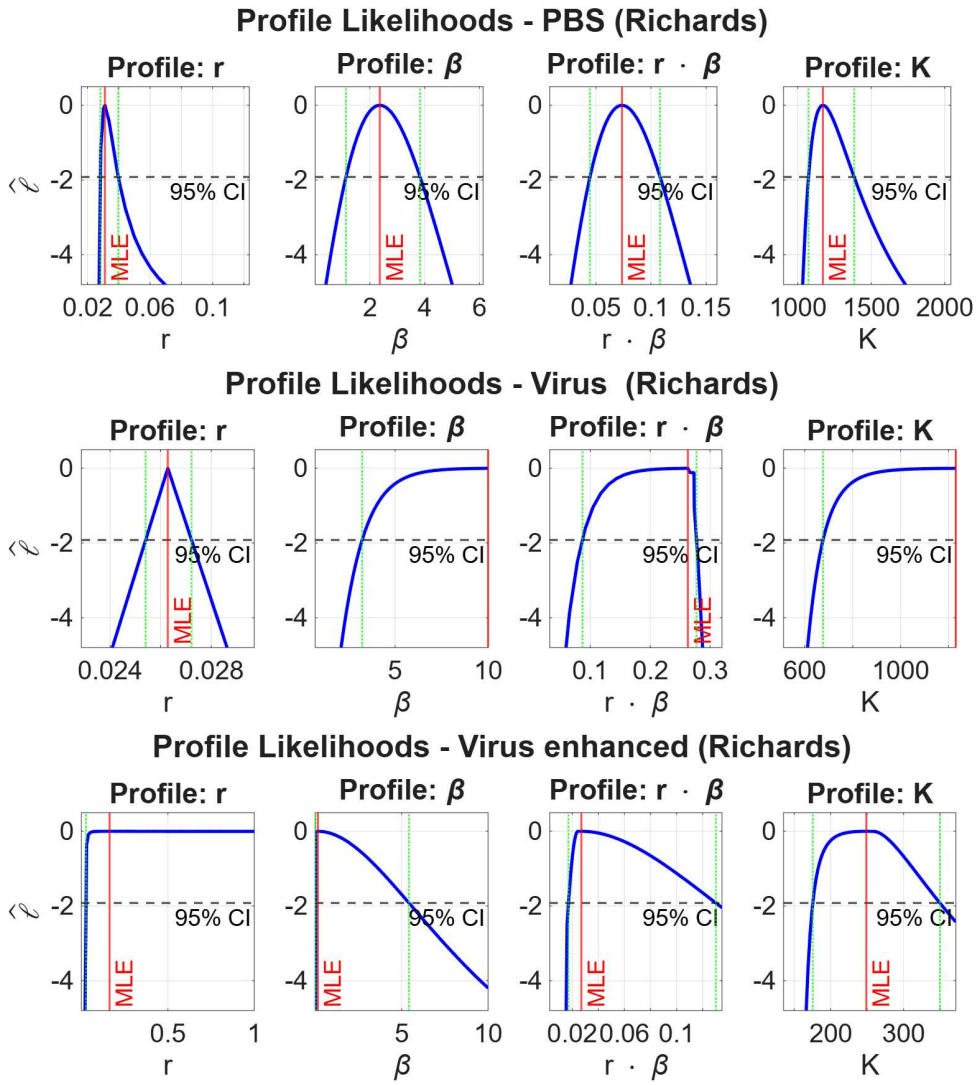


Figure 2.11: Comprehensive Richards profiles of *A549 AdV* datasets, using data from Ref. [8]

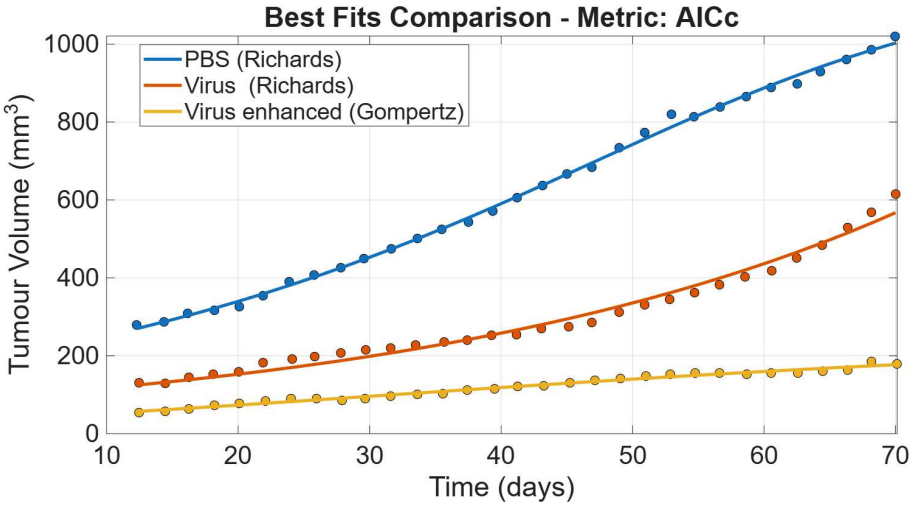


Figure 2.12: A549: post-nadir plot of all datasets with respective "best model" curves according to AIC_c metric

Logistic model is the possible redundancy of β itself in case of easy to capture trends, even though biological significance is usually attained for a set of MLE estimates. While individual structural parameters may carry higher uncertainty, the model remains robust and consistent, provided that its complexity is generally justified by its ability to get potentially asymmetrical growth patterns.

In addiction, we offer an implementation for a MATLAB script that serves the following purposes: for each loaded dataset, it runs MLE optimisation for each of the selected models and, by tuning the considered interval, evaluates the chosen metric. It then reports values and the best model for each dataset, while plotting points and curves in line with research article results. Consider figure 2.12 as an example. This form offers a quick overview of the goodness-of-fit for all data contained in a specific article and shifts its focus from merely mathematical metrics to a biological mean of comparing mice groups response to different OVT settings. Moreover, it lets us switch between different metrics and explore consistently.

We shall now focus on model stability across data sparsity and give answer to another aspect of modelling: whether complexity is an advantage over sporadic data.

2.6 Framework Robustness

On this section we explicitly stress-test our models: we are interested to see if results hold following a new criteria. We define the **Persistence Property** as a measure for the robustness of a model, trying to quantify the degree of resilience when the data interval shrinks. In particular, we will adopt **Monte-Carlo** oriented simulations, along the decrease of available points, and prove statistically how often a model is successful according to a selected metric. We first discuss what is running underneath a single iteration.

Sub-sampling

A first aspect is to decide how to restrict the interval: if we were to cut off chunks of points without any criterion, results would be intrinsically misleading. Removing

a crucial segment of the data would hide major shifts in shape, masking the true behaviour. In order to work with a realistic framework and force models to act reasonably, we have to wisely set the sampling method. Since our approach aims at evaluating models that are consistently successful at multiple fits across a wide range of combination for the reduced dataset, we propose to ignore one point at a time. By repeating the points election every iteration, for a total statistical significant M amount of times, what we hope to see is a stable selection for the best model[3]. The idea is simple: a "wise" model does not need a plethora of redundant information, but rather a minimal look at the graph to be able to get the main growth pattern.

The way we choose our points for every new step of M iterations obviously affects the end result. For the sake of testing, we consider a naive *random* approach, in which no limit is forced on points selection. Our guess is that standard uniform random extraction applied to point removal does not significantly bias the estimation process, as the underlying data noise structure remains preserved. Since the sub-sampling is independent of the growth phase, the residual dispersion and the resulting MLE parameters should converge to the same statistical equilibrium. Ideally, the model's predictive power remains robust, showing that the global fit has a low sensibility to reasonable reductions in sample density. We therefore try to empirically measure to which extent each model follows persistence.

More metrics as lenses for accurate interpretations

Moreover, we have to make sure we provide our analysis with the right tools to comprehensively read data and results. In the context of oncological modelling, the primary challenge lies in balancing the mathematical accuracy of the fit with the biological plausibility of the underlying parameters. To rigorously evaluate the performance of our growth models, we employ a multi-metric approach that transitions from absolute error quantification to information-theoretic selection. The AIC_c value, for example, better suits the current reducing framework for which, just like we previously showed in the pre/post-nadir optimisation runs, the standard AIC metric would bias results when reducing number of points.

The first measure of discrepancy we encountered is the SSR, that provides a direct calculation of the total unexplained variance; its magnitude is intrinsically tied to the number of data points n, making it unsuitable for comparing datasets of different lengths. To resolve this scale-dependency, we utilise the **Root Mean Square Error** (RMSE), defined as

$$\mathrm{RMSE} \coloneqq \sqrt{\frac{\mathrm{SSR}}{n}}.$$

The RMSE translates the total error back into the physical units of the experiment (mm^3), representing the average distance between the observed tumour volumes and the predicted curve. However, neither SSR nor RMSE accounts for the degrees of freedom. This is a critical omission producing a bias when comparing models of varying complexity. To address this, we introduce the **Residual Standard Error** (RSE), calculated as:

$$\mathrm{RSE} \coloneqq \sqrt{\frac{\mathrm{SSR}}{n-k}}.$$

[3]The procedure is not redundant: if we consider a dataset of $n = 25$ observations, at step 15 it yields to a number of possible combinations equivalent to computing the binomial $\binom{25}{15} \simeq 3.27 \times 10^6$.

Unlike RMSE, RSE penalises the loss of degrees of freedom[4]. If a more complex model reduces SSR only marginally, RSE might increase, signalling that the added complexity is not statistically justified. This metric is particularly robust when narrowing time intervals or removing data points, as it monitors whether the precision of the fit is maintained despite a smaller sample size. During our step-by-step subsampling iterations, this metric might help identifying the threshold where further data removal prevents the model from accurately capturing the underlying growth dynamics.

Another possible choice of metric for model selection is the **Coefficient of Determination** R^2 that quantifies the proportion of the total variance in the tumour volume data that is explained by the biological growth model:

$$R^2 := 1 - \frac{\sum_i^n (\mathcal{D}_i - f_i)^2}{\sum_i^n (\mathcal{D}_i - \bar{\mathcal{D}})^2},$$

where $\bar{\mathcal{D}} = \frac{1}{n}\sum_i^n \mathcal{D}_i$. While R^2 provides a measure of goodness-of-fit, it does not account for the information loss inherent in model selection for our specific simulations purpose, neither for different number of parameters. Consequently, we prioritise the AIC_c as the primary selection metric, as it rigorously penalises model complexity while preventing over-fitting, especially in datasets with limited sample sizes.

In practice, these metrics are read in synergy: a chosen model by the lowest AIC_c is only considered **robust** if accompanied by a physically sensible RMSE (low enough to be within experimental precision) and a stable RSE.

Considered metrics are collated in Table 2.1.

Metric	Formula	Meaning
SSR	$\sum_{i=1}^n \left(\mathcal{D}_i - f(t_i, \hat{\theta})\right)^2$	Sum of squared residuals.
RMSE	$\sqrt{\frac{\mathrm{SSR}}{n}}$	Normalised average error.
RSE	$\sqrt{\frac{\mathrm{SSR}}{n-k}}$	Unbiased absolute physical error.
AIC	$2k - 2\log(\hat{\ell})$	Likelihood metric golden standard.
AICc	$\mathrm{AIC} + \frac{2k(k+1)}{n-k-1}$	Weighs AIC by number of points n.

Table 2.1: Goodness-of-fit metrics and model selection criteria.

Monte-Carlo Simulations

Previous sections described models responses to different settings and fitting needs. We now prove their robustness through **Monte-Carlo** (MC) simulations, consider a number of M iterations for each step of point removal, while applying a random points selection. As an introductory explanation, we first showcase what a single step of sub-sampling looks like. We refer, again, to the experimental data of Figure 1.4 . Since we have established that the exponential curve does not contribute significantly and meaningfully to curve fitting, we will not include it in our analysis from now on.

[4]In the general case of a multiple linear regression with p features RES $:= \sqrt{\frac{SSR}{n-p-1}}$, such that the denominator is the degree of freedom of the system. For our case, $k = p + 1$.

Take, for instance, the *PBS* dataset, after neglecting outliers. Let us consider $M = 500$ iterations, which turn this problem into a computationally demanding task [5]. The original best model for the considered set is Richards, according to AIC_c and parameters interpretability, and we investigate whether this trend continues when reducing the number of points. We focus on two removal step consisting of, respectively, 15% and 30% of total number of points (translating to 5 and 11 removed points). As a result, Richards gets chosen in 71% of simulations, ahead of the logistic curve. The former percentage then reduces to 42% in the next case. We remark that AIC_c between the two models differs, on average, by less than 1%, and it is lower than the baseline value, meaning an overall fit improvement, as a stable RSE value shows.

Moreover, DW values are shifting from an auto-correlation value of one towards smaller AR(1) values, reaching an average of 1.5. A possible hypothesis arises here: by not considering an excess of points in the original interval that might have been interpreted as noise rather than information itself, models perhaps are now capable of more meaningful fit. This would mean that the risk of removing chunks of growth is far out-weighted by the random distribution of points in most other iterations, as the average winning AIC_c drops from 200 down to 160. By studying parameters values distribution - for which only the best ones get plotted - we see how different ranges of r, K are gaining momentum, related to the Logistic model being chosen more often. Interestingly, σ appears to distribute more into lower values of the spectrum, implying lower absolute fitting errors. The same testing on V dataset results could be hereby labelled as not relevant from a sigmoidal point of view, with Richards getting chosen in every iteration, overtaking Logistic and Gompertz (with averages ΔAIC_c of up to $15, 30\%$). This is possible by β values with a not realistic average of 10. Looking at Figure 2.14 we get an idea of how uneven the comparison works out.

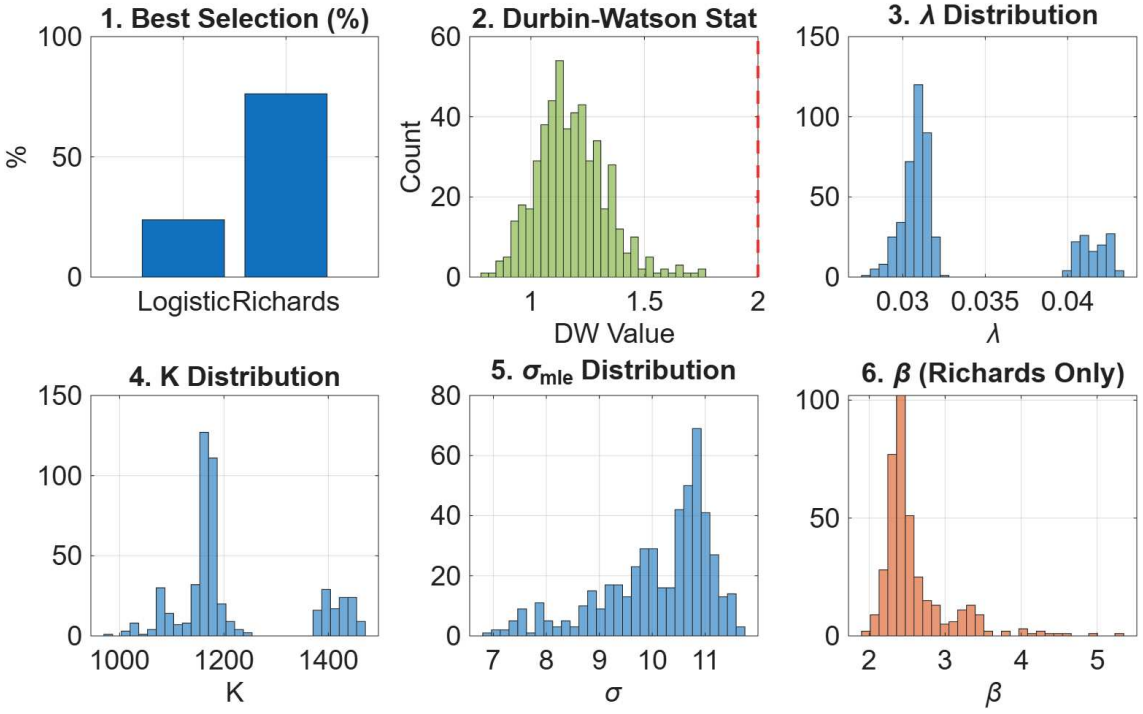
This snapshot make us wonder if such a coherent trend in model selection appears over a complete list of removal steps. However, from this type of data visualisation, parameters explicit representation lose importance due to the intrinsic mechanism of random selection of points[6]. Hereby, by interrogating AIC_c and RSE evolutions we are able to distinguish if the majority of fits keep a meaningful "message" across the removal steps. By showing a hopefully stable pattern, we can push the analysis to measure how much qualitative information is available in worse data scenarios, even for simple models, before results become unreasonable.

Since parameters seem to remain coherent, we will trade their explicit plotting for a concise but dense collection of fit metrics[7], and commit our efforts in measuring the mentioned property.

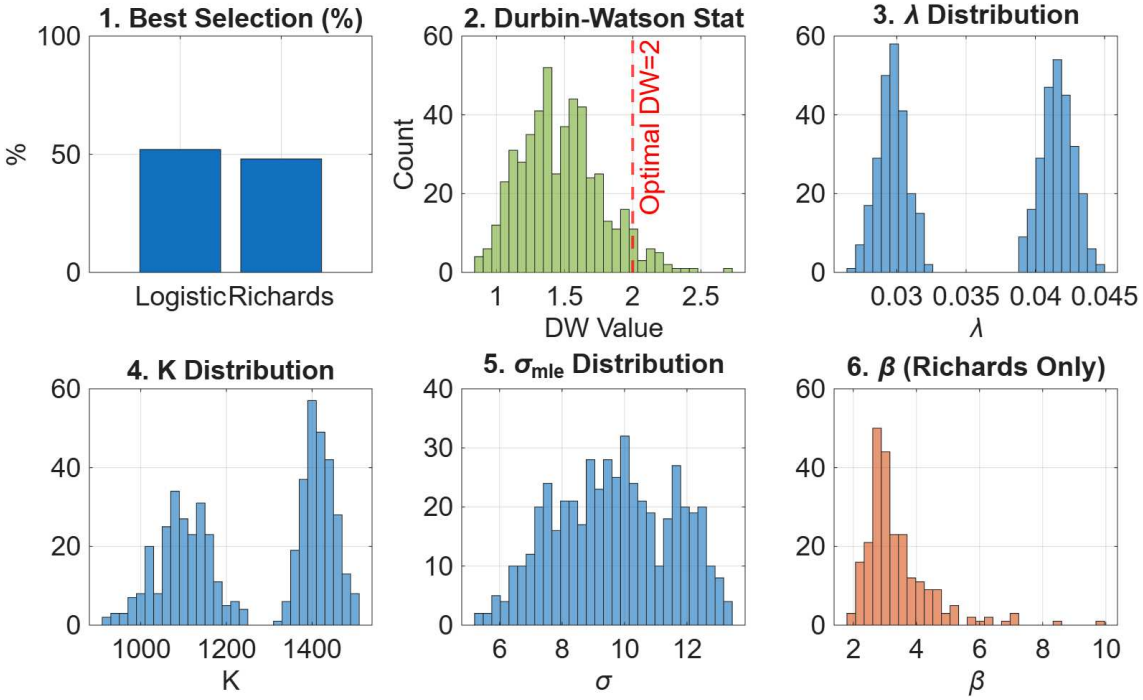
[5] To best optimise the running process, the implemented MATLAB code leverages parallel execution, achieving a *speed-up* of roughly 2.0 times with a 4 cores i7 8th gen laptop. Total time of execution was about 2 minutes for each step, for all datasets together.

[6] We point out that, in order to allow results to be reproducible, we impose a MATLAB *seed(42)* as a fixed value throughout our code. With this inclusion we standardised the approach that otherwise would depend on the parallelisation settings of local machine.

[7] In the MATLAB *MLE_Persistence* implementation, we save into a *log* file every step average values for each parameter and metric.

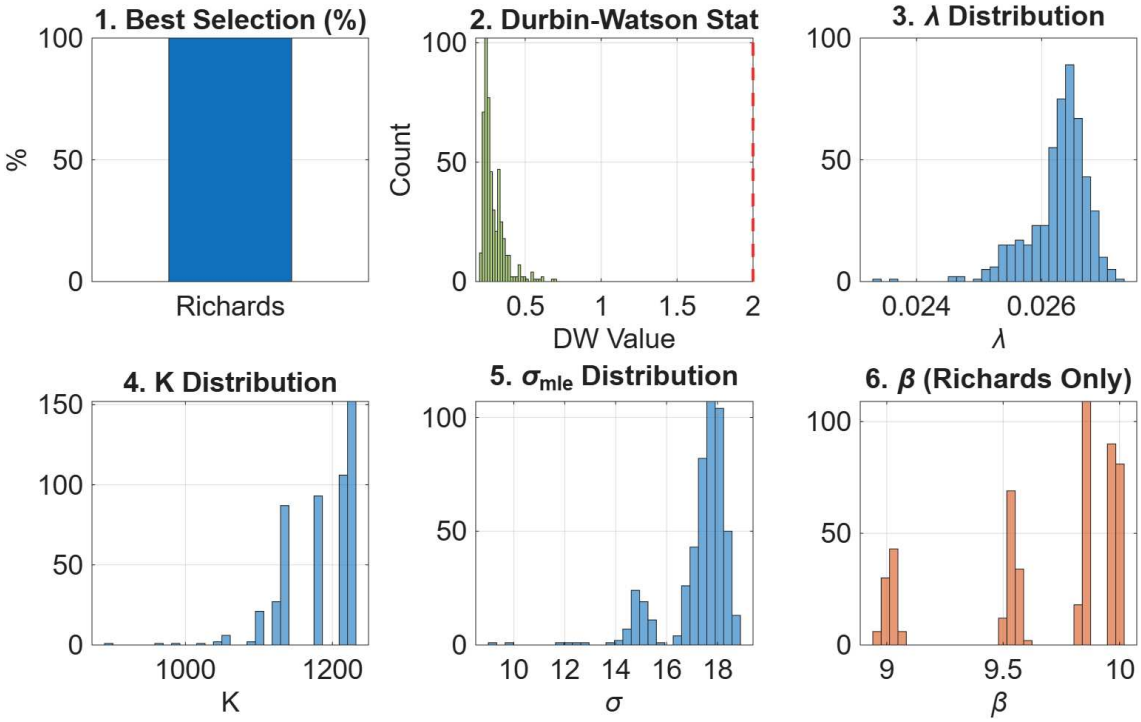


(a) 15% of points removal.

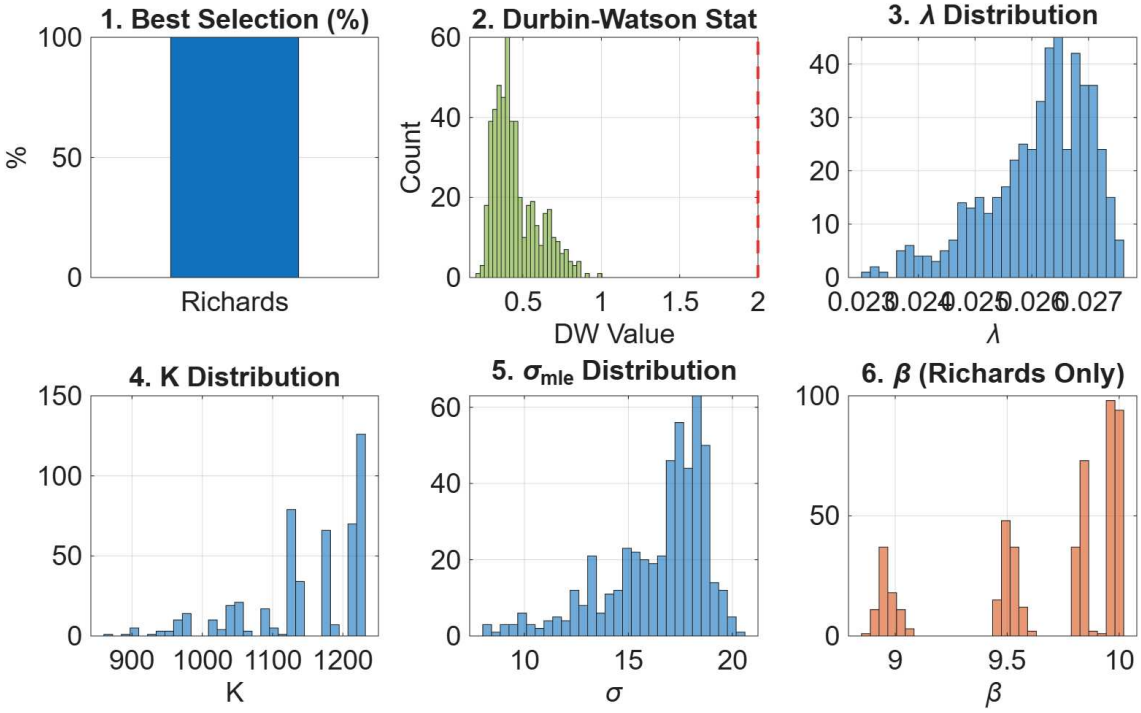


(b) 30% of points removal.

Figure 2.13: Monte-Carlo iterations for the removed outliers *A549 PBS* dataset. We remark that the Gompertz model was a possible choice that did not "win" in any iteration.

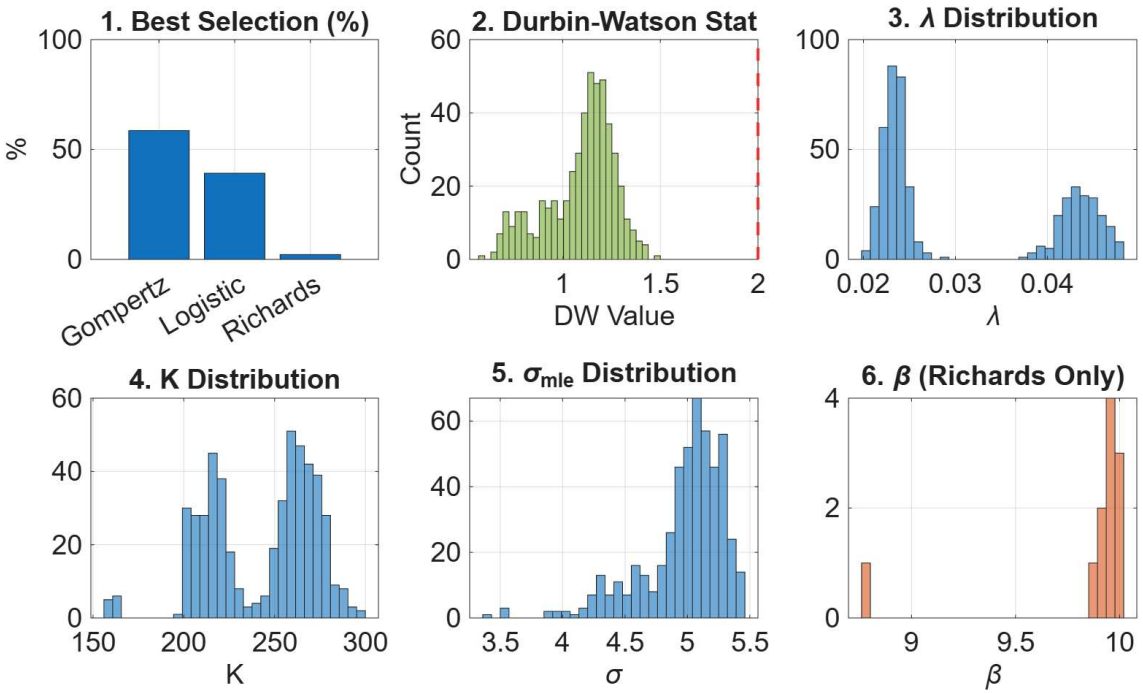


(a) 15% of points removal.

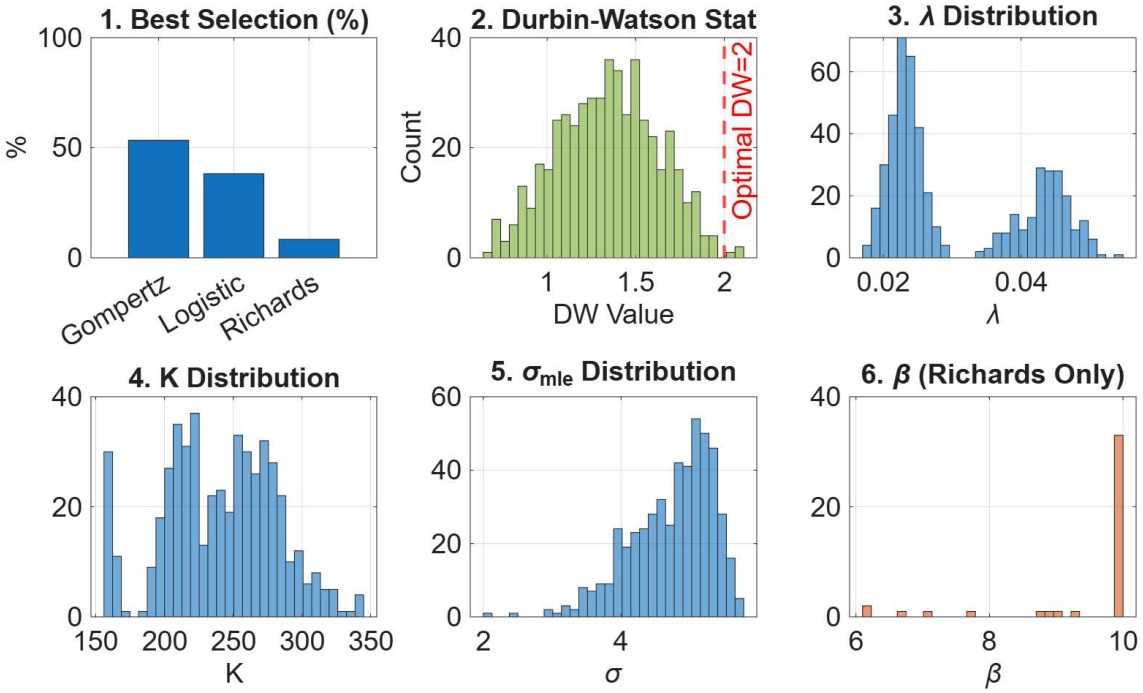


(b) 30% of points removal.

Figure 2.14: Monte-Carlo iterations for the post-nadir *A549 V* dataset.



(a) 15% of points removal.



(b) 30% of points removal.

Figure 2.15: Monte-Carlo iterations for the post-nadir *A549 VE* dataset. We remark that Gompertz model was indeed a possible choice that was never chosen, in any iteration.

Persistence analysis

We deep dive into each model's robustness by addressing the persistence framework[8]. By considering both figure 2.16 and 2.17 plots we reported our complete persistence analysis. As we discussed previously, we recall that, at the baseline level, the model selection follows a distinct pattern across datasets. For the *PBS* and *V* cohorts, the Richards model provides the most parsimonious description, achieving the lowest AIC_c values (244.66 and 277.66, respectively). In these cases, the Richards model also minimises the RSE, with values of 11.19 for *PBS* and 19.06 for *V*, significantly outperforming the Logistic and Gompertz alternatives.

Conversely, in the *VE* dataset, the Gompertz model is selected as the optimal descriptor ($AIC_c = 198.24$), followed closely by the Logistic model ($AIC_c = 198.37$). Here, the Richards model ($AIC_c = 201.09$) is penalised for its additional complexity, as the reduction in RSE does not statistically justify the inclusion of the shape parameter β, especially considering how much it becomes not meaningfully detectable in final removal steps.

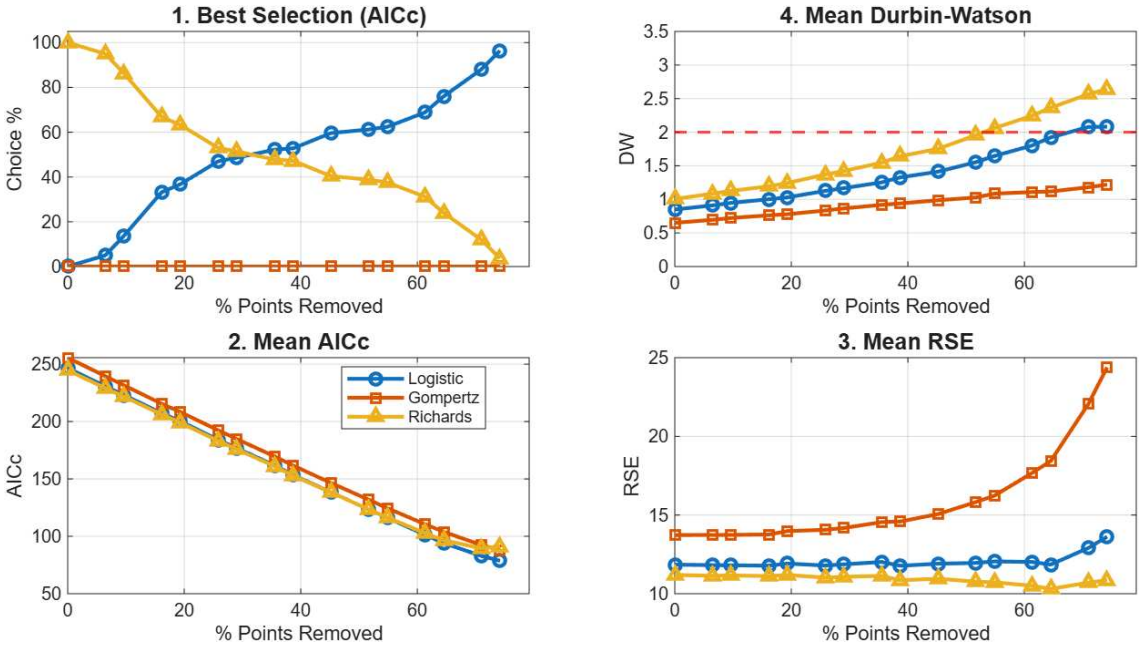
The persistence analysis reveals a critical **Crossover Point** (CP) where data sparsity forces a shift in model selection due to the lack of data in the model itself. In the *PBS* dataset, the Richards model maintains its statistical superiority until more than half of points are removed ($n = 14$). Beyond this threshold, the Logistic model becomes the best also for averages AIC_c values[9]. For the *Virus* dataset, the Richards model demonstrates higher resilience, remaining the optimal choice until 64.5% data removal, after which the Logistic model takes the lead by a 5% accuracy. Notably, while the Richards model consistently yields a lower RSE across all steps, the AIC_c correctly identifies overfitting as the sample size decreases and the quadratic penalty term increases sensibly.

The DW statistic further characterises this transition, but one should be aware of its main risk of use. At baseline, the *V* dataset shows a great AR(1). As data density decreases and simpler models are selected, the DW statistic for the selected models approaches as optimal range, but is not to be considered positive. This shift indicates that in sparse datasets, the residuals transition from systematic structural error to random Gaussian noise and confirms that the additional information provided by the Richards model is no longer relevant. Most importantly, data is treated more as noise than a complex biological phenomenon. This is interesting because it provides an indirect measure of the minimum amount of data points to be collected in these types of experiment to maintain a scientifically relevant structure.

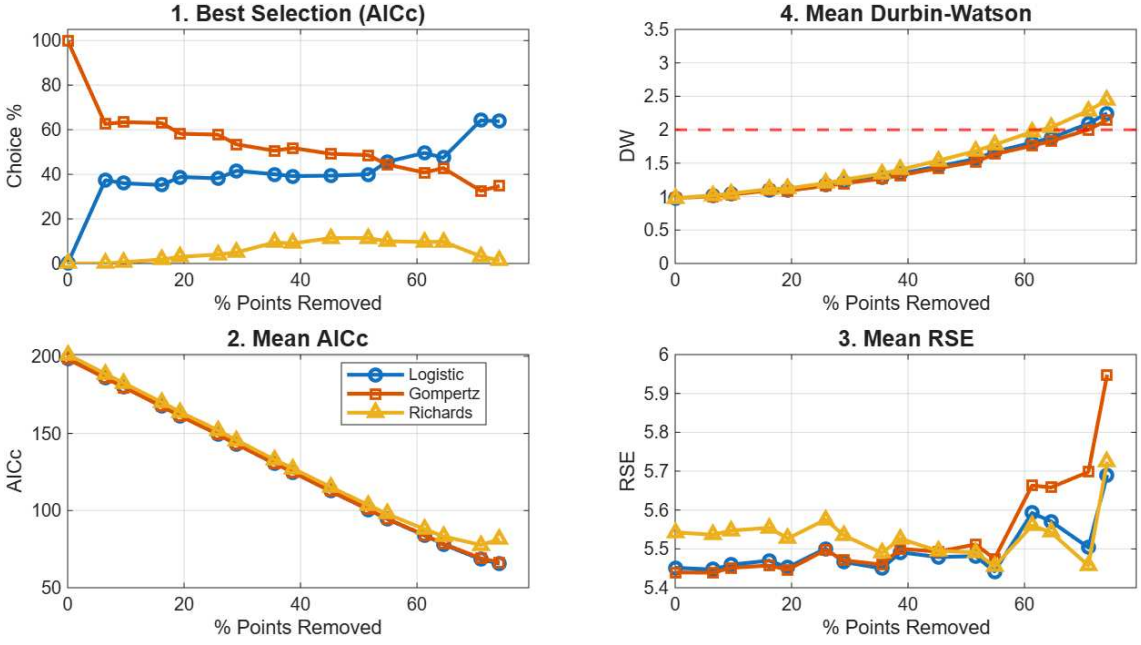
Moreover, we observe that CP is taking place "late" in this sequential analysis. Being able to cut the interval, with less than 15 points left and spread over 2 months (hence, with most iterations showing clusters of points instead of a uniformly discretised interval), while still capturing main dynamics behaviour is to be considered quite remarkable.

[8] As an extension of the MC approach, this task is dimensionally more computational demanding. In order to complete the list of steps point by point until reaching the least number of points possible, a run lasts for 30-40 minutes depending on dataset size.

[9] We remark that the winners graph is not contradicting the mean AIC_c plot; it simply implies that on average Richards holds for lower values, but gets selected less frequently. It might be due to a more consistent Richards behaviour, while Logistic tends "loose perception" in some circumstances.



(a) *PBS* dataset.



(b) *VE* dataset.

Figure 2.16: Persistence analysis conducted on the post-nadir *A549* relevant datasets. We remind the reader that all considered cleaned datasets here presented are composed of 30 points.

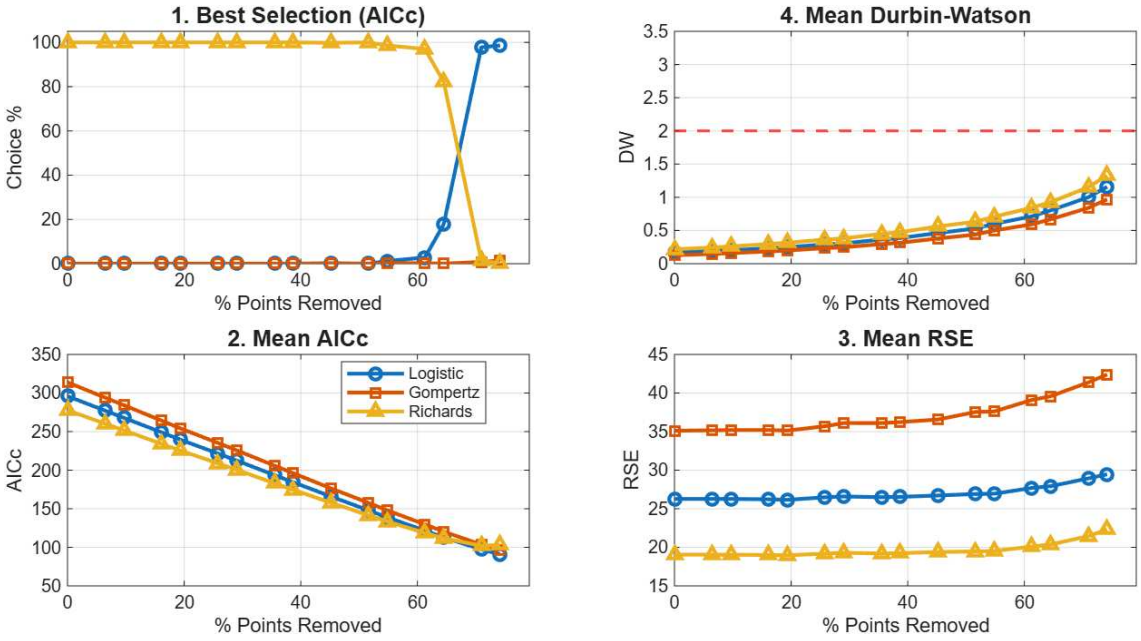


Figure 2.17: Persistence analysis conducted on the post-nadir A549 less meaningful V dataset.

Conclusions on *A549* data

Finally, estimated parameters across all datasets exhibit strong biological consistency and numerical stability, coherent with experimental settings and previous analysis. That is, at least, before the CP of information loss over removed points. Indeed, carrying capacity K correctly reflects the observed growth limits, with mean values for the control group (*PBS*, $K \approx 1400\ mm^3$) being significantly higher than those for the treated groups (*VE*, $K \approx 250\ mm^3$), capturing the macroscopic effect of viral treatment. Throughout the entire analysis it remains the most stable. Instead, shape parameter β identifies distinct growth regimes depending on the considered cohort. A class of high asymmetry ($\beta > 1$) in the *PBS* and *Virus* groups, β values (2.5 and 9.5, approximately) indicate a late-stage deceleration, where the tumour maintains an accelerated growth phase until it is near the carrying capacity: this is not biologically consistent. Then, a second class of strong deceleration ($\beta < 1$) in the *Virus enhanced* group, where the low value of $\beta \approx 0.1$ suggests a shift toward Gompertz-like dynamics. The IP occurs much earlier, reflecting a more immediate impact of the therapy on the growth rate. However, for *PBS* and V dataset, the CP point change parameters significance. Fits attain "poor" quality as metrics worsen noticeably, while β and r starts "oscillating".

The apparent discrepancy between RSE and AIC_c is fundamental to our model selection logic: while the Richards model usually achieves the minimum RSE - prior to the CP - its purely descriptive advantage is weighted by the AIC_c penalty. We conclude that Richards tends to be the preferred biological descriptor only when the reduction in RSE is substantial enough to overcome the information cost of the shape parameter β. Where $\hat{\ell}$ favours simpler models despite a higher RSE, we interpret the Richards fit as a case of over-parametrisation due to limited sample size.

Limitations

Our intuition is that while the Richards model often provides a superior qualitative description for dense datasets, its performance is highly sensitive to sample size. A critical finding from our persistence analysis is the systematic divergence between the standard AIC and its corrected counterpart in sparse data regimes. In several high-removal steps, both RSE and AIC favoured the more complex Richards model, whereas the AIC_c strongly penalised it in favour of simpler models[10]. This reveals that this reduction in residual variance is an artefact of over fitting. We therefore conclude that in these sparse regimes, the data lacks sufficient structural information to justify the estimation of the shape parameter β. Consequently, the AIC_c serves as the only robust and reliable metric for model selection in this study, successfully preventing the misinterpretation of random experimental noise as biological asymmetry.

We remark that, although it would sound re-assuring, an enthusiastic and biased approach does not provide any benefit to the research itself. One shall be honest in pointing out limitations for each model; in case of the generalised Logistic, its redundant and complex structure might not fit in sparse datasets. However, another metrics confirms that fit quality is not degrading unless surpassing CP. Thus, we can conclude that this persistence property is present as a "shared" trait between the sigmoidal family. Indeed, as we will see in other examples, even if there is no such thing as a perfect model, whenever one is the best fit for a cohort, it tends to hold up over several points removal.

We conclude this chapter on model evaluation by measuring the predicting power of sigmoidal curves.

2.7 Forecasting

This section could illustriously be resumed with a famous quote attributed to Niels Bohr (or perhaps descending from a 20th century Danish proverb) that goes: "It's tough to make predictions, especially about the future". In order to understand why, we finally address the elephant in the room: *does a better fit necessarily imply a better long-term prediction accuracy?* Before being able to answer, we first need to define how the evaluation process might tackle this question.

A549 Testing

The first step is divide the dataset into a first **train set** and a right-hand portion of data called **test set**, the two varying by size. As a first idea, we can think of the former to be around 80% of total points and the latter covering the rest of the considered cohort. Once we have separated the interval, the technique requires to fit the curves according to the training set and, using the obtained θ_{mle}, simply plot model results for the remaining days. Clearly, any prediction beyond known points is a guess with no real comparison and might be used to visualise the order of

[10] With simple calculation, the penalty correction term impacts Richards with $c_R = \frac{60}{n-6} = 15$, while $c_L = c_G = \frac{40}{n-5} = 8$, for $n = 10$. By hypothesising a baseline AIC=200 for each model, the quadratic penalisation would account for an extra 4% difference, while linear penalty already affected Richards for the β parameter inclusion.

magnitude of the final dynamics, for which the IP position already gives a qualitative estimate.

Once the interval has been pruned from outliers (or nadir delays), focus shift onto measuring the accuracy of the estimated prediction and comparing residuals between models. Since in this case there is no distinction in interval length, we can use the physical error calculation of the RMSE metric as an accurate average predictor instead of dimensionless values. We calculate the metric explicitly for the train set, as well as the test set, and we consider a forecast accuracy ratio of training error over test error:

$$\mathrm{r}_f \coloneqq \frac{\mathrm{RMSE}_{tr}}{\mathrm{RMSE}_{te}},$$

with the main purpose of quantifying a possible answer about prediction accuracy. It goes without saying that values of r_f closer to one are preferred.

Again, rather than aseptically looking at numbers, we are seeking for qualitative information. Another clue lies in the residual analysis: since the model is not training in the second part of the plot, we limit the analysis on whether it is over or under-fitting. The interesting aspect here is that models "see" the same set of points, and we may ask which growth slopes or section influence more each of them. For this case, we propose a few different train set selection, while maintaining test size of the last 15 days ($\approx 20\%$ of points)

An initial remark has to be made: due to the presence of strong shifts in some region of data, by selecting a specific point as the start or end of the training set may heavily influence outcomes. This should not be considered as an Achilles' heel, but rather as an intrinsic characteristic of this model training strategy. Our approach currently relies on the naïve idea of cutting the main dataset, exposing the risk of misleading local trends compared to a more generalised sight on the interval. To address this, we tested a few combinations and compared them, but more general and reliable methods could be deployed. Nevertheless, some results look promising with the right mathematical lens. Figure 2.18 shows results relative to a mature phase of growth, happening over 2 weeks, 40 days after the experiment begins. Interestingly, for each cohorts, the corresponding preferred model according to AIC_c and persistence analysis accounts for the lowest RMSE_{te} and overall ratio.

If, instead, we take as train set the days between nadir ($n = 11$) and day 55, the past data density yields a more accurate fit. While both figures 2.19 and 2.21 confirm the limitations already discussed for which only Richards maintain some significance, forecasting performance improves noticeably with the study of the full interval, except for the generalised Logistic case; it appears that it is more susceptible to minor but sudden changes, leading to a strong response by anticipating or delaying the IP location. On this specific dataset, to draw positive qualitative conclusions would mean to force a biased interpretation.

2.8 Future Improvements

By combining ideas from previous sections, one could come up with a perhaps more accurate measurement of forecasting potential. If we were to put ourselves in the biologist's shoes, beside knowing which model has been identified as more meaningful depending on classified markers, and its ranges of significance for parameters, we would seek for a better tool. Such a technology should incorporate a one-size-fits-all structure. Clearly, a magic answer does not exist, but we could potentially get

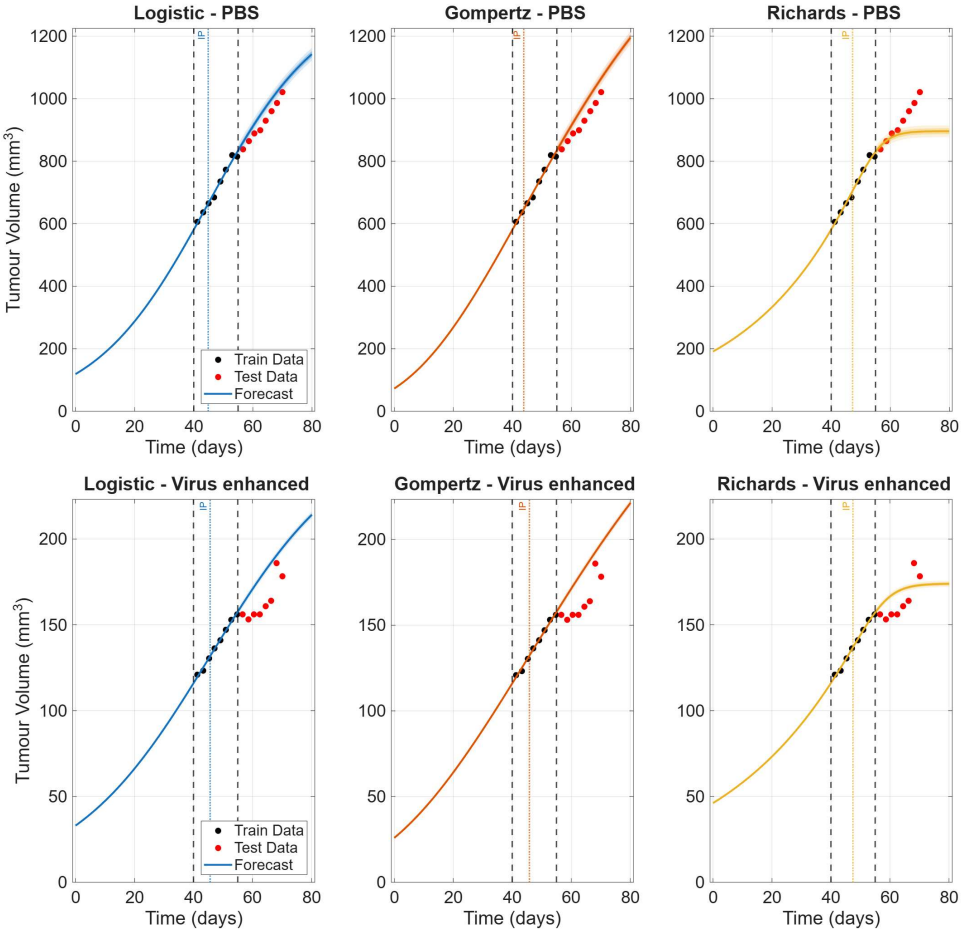


Figure 2.18: After cleaning the *A549* data from the nadir reach, we trained models over a 40-55 days time frame.

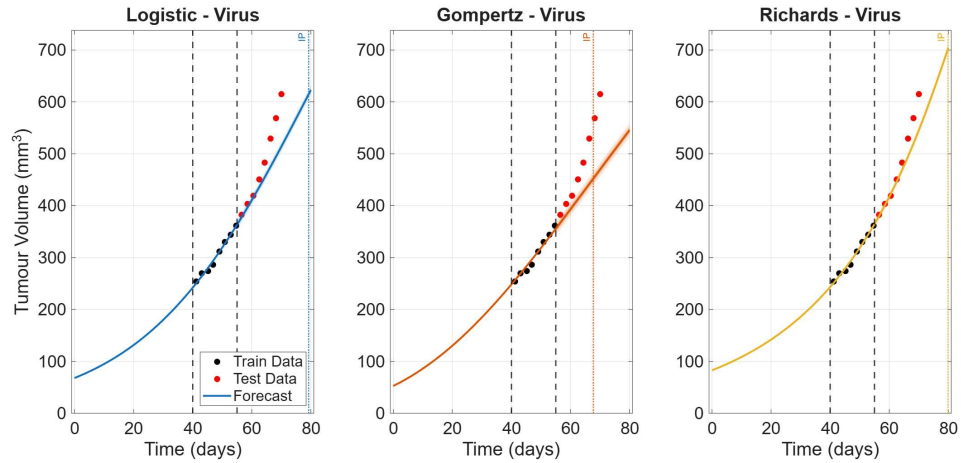


Figure 2.19: *A549 V* cohorts forecast analysis for the same train set choice. Richards is the only model not showing an IP and hence incorporating a steeper growth, more adherent to measured data.

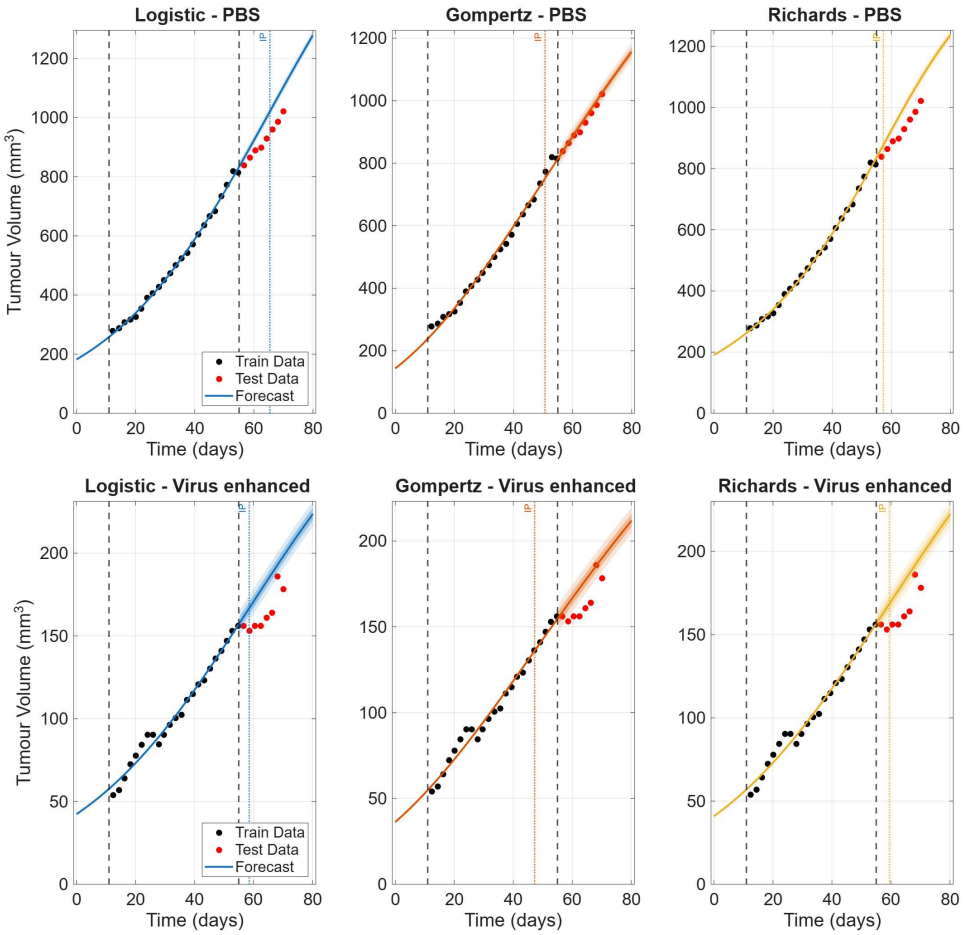


Figure 2.20: After cleaning the *A549* data from the sets close to the nadir, we trained models over the complete time series until day 55.

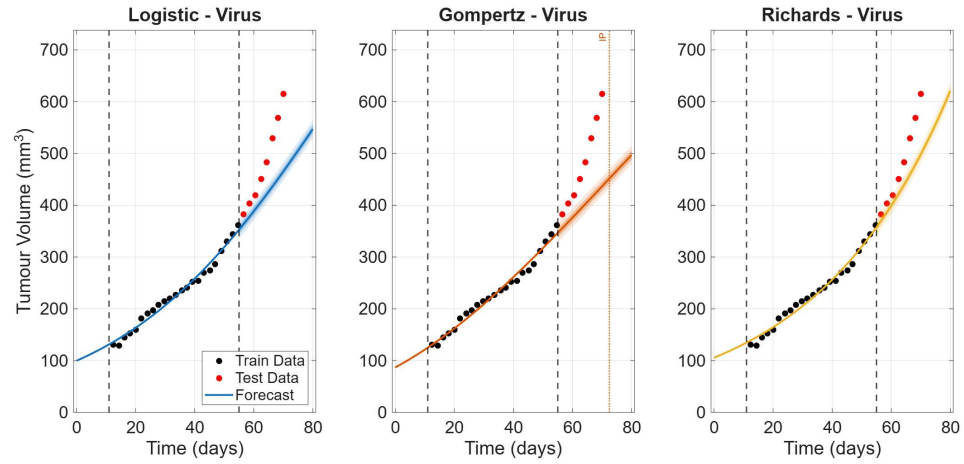


Figure 2.21: V cohorts forecast analysis of the complete training time series

closer by implementing the following idea. Since we chase predictive power and we stressed trends capture, a synergetic analysis could rise.

Hence, we suggest a persistence analysis in which, for every step of points removal, a number of MC iterations arrange the training set. This represent a forecasting measurement comparing, for each model, the predictability over the test set (ideally for different ranges, e.g. by means of 5% increase). Measuring on average the forecast accuracy between the M iterations in each step and plotting the best performing model would result in a more practical use of sigmoidal curves. We hypothesise that the randomness of the points choice would converge to a statistically significant measure of forecasting over various datasets, ideally providing clinicians with a tool able to employ statistically sound uncertainty levels, for short to medium term forecasting.

Uncertainty could be measured from combining a range of values (actually a generalised grid) for each of the "hot" parameters and measure how "playing" with a specific range results in different outcomes. This is a form of sensibility analysis.

Moreover, the process might be applied in the same fashion to more complex models and could be compared to truly estimate how much overall dynamic understanding is possible with single ODE models running over 5 (or less) parameters.

Next steps

Other methods can still be employed or devised to improve the analysis. By measuring fit quality, we stressed the importance of residuals, because they can carry hidden information and express what a model is lacking. By investigating their AR(1), we observe how, often overlooked, noisy data perturbations may alter results themselves. As we seen, metrics might seduce one's perception on a topic, but fit comparisons left space for doubts. Indeed, we pose the focus on the intrinsic statistical optimisation strategy: from the definition of MLE and in every optimisation step, $\hat{\theta}$ has been identified from the standard assumption of Gaussian distribution. Since, as we seen, asymmetry is an important component in complex tumour dynamics, it could be appropriate to look at different distributions. We might look for a change in the *normal* paradigm and consider a less regular and asymmetrical probabilities for explaining studied phenomena.

It is also relevant to emphasise that some tumour-virus dynamics are known to produce oscillations. This element cannot be addressed by 'S' shaped curves since they allows for a flection at the IP. It is instead a possibility to consider numerical solutions from ODEs that capture the essence of viral and tumour dynamics. In the next chapters, we will look at the effect of a log-Normal distribution on MLE and how predictability for oscillatory dynamics can be treated successfully.

Chapter 3

Normality Scepticism

Data Science is a vast realm, as the term itself incorporates a broad range of interpretations. By undergoing an analysis over a set of experimental observations, one shall intersect knowledge from different scientific fields aiming at better elaborate information. If we think about mathematical modelling for biological systems, we acknowledge that the underlying mechanism governing how we perceive reality is influenced by Statistics. Indeed, each time we chose and implement a model, it is crucial to address issues revolving around data distribution and, consequently, residuals related noise. Whether electing the null hypothesis for an experiment or detecting frequency of events, we are mandated to address randomness and uncertainty.

3.1 Statistical Assumptions

The main contributor determining the statistical analysis of datasets that can be validly carried out is the **Distribution**. Interestingly, across scientific literature over decades of empirical results, some actors took the scene more often than others: it is the case for the **Normal** distribution

$$X \sim \mathcal{N}(\mu, \sigma^2).$$

The nature of the bell-curve uses the **mean** and **standard deviation** parameters to express variability of points, such that they are generally depicted as

$$\mu \pm \sigma.$$

We remark that in this formulation, every point is considered as an additive shift from the mean value. This aspect determines its symmetry: the odds of being far off the chart on one side is obviously the same as on the other side, and both tails incorporate the same probability. The corresponding probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Potentially biased interpretations

A vast number of research takes the Gaussian distribution for granted, electing it as the undiscussed leader at explaining all sorts of phenomena. Whilst it is true

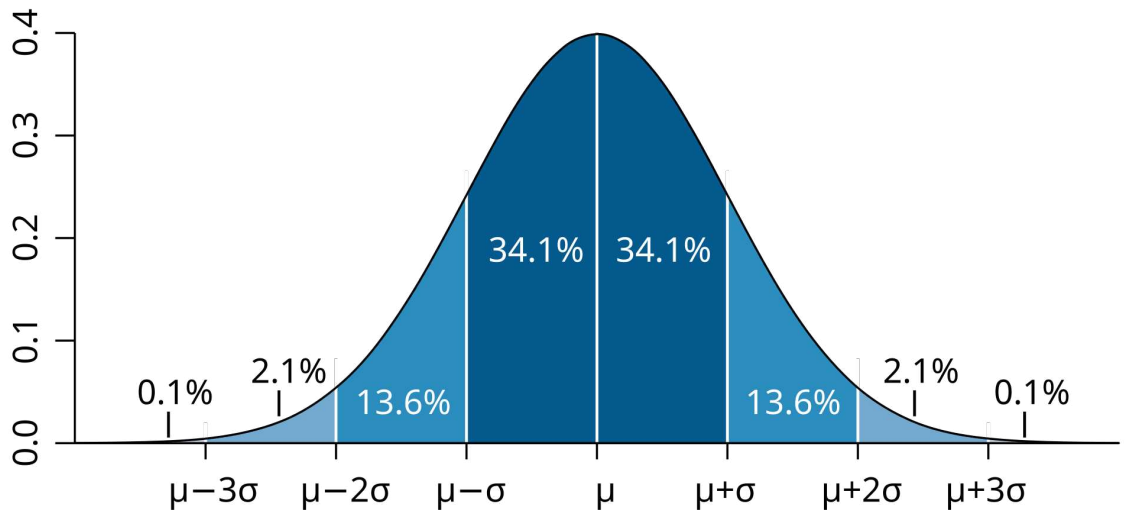


Figure 3.1: Normal distribution segmented by standard deviation additive sectors. In particular, values less than one σ from the mean account for 68.27% of the set.

that many measured events occur with a valid Normal configuration, such as random deviations from target values in industrial processes, some others happen in patterns that must be described in terms of asymmetry (**skewness**). To get an idea, simply think of situations in which a momentum-related dependency takes place; examples are to be found in species abundance or infectious diseases development to name a few. Moreover, a crucial aspect of a well-fitted Normal distribution is the variance evolution for which data should follow a **homoskedastic** trend: since the bell-curve statistic accounts for additive effects, it will perform poorly in contexts such that variability of points is changing over time. Throughout our analysis considered datasets containing qualitative shifts in tumour growth. As a result, *a posteriori* discussions over residuals often showed partially unexplained evolution patterns.

Hence we denote two of the most common ways for which data interpretation is forced, perhaps unintentionally, over Normal explanations. Profiles showing skewness tend to have large outliers; common practice is to omit such observations, shrinking shapes but introducing a bias. Another way is by averaging clusters of asymmetrical, irregular data, so that dealing with mean values easily tends to a more Gaussian fit. The main limitation of the bell-curved statistic is that is impossible to address *multiplicative* problems by matters of plain sum of standard deviations.

Abraham Maslow wrote "*it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail*", expressing a surging bias potential in rigorous approaches of any kind. A narrow sight on things might result in worse results, due to misspecified and meaningless models. We thus introduce new metrics to try and detect possible differences in data statistic. More specifically, we point out a crucial aspect of the SSR - and therefore RMSE - metric. By looking at the second value, we get an idea of the average physical error, but it is not always impartial at comparing two fit examples. Since each term of the sum is the squared distance between model estimation and the real value, an error occurring on high values weights sensibly more than one from smaller ones. Consequently, the model prioritises the fit at high tumour volumes, where absolute errors are largest[11]. In the present case, this

[11]We recall that the MLE method operates proportionally to the minimisation of the SSR.

approach sacrifices the accuracy of early-phase growth dynamics, as the left tail of the distribution represents the initial stages with lower tumour volumes. One might argue that a poorer fit in the early stages is acceptable, assuming that capturing the later dynamics is more critical for long-term projections. However, while there is no mathematical imperative to prioritise the tails of the distribution, a robust analysis requires a totally unbiased evaluation. By shifting the perspective from absolute to relative errors, we can treat the time series as a sequence of uniform weights, effectively mitigating the bias toward larger values and addressing the growth as a consistent process rather than an ever-increasing risk.

To resolve this discrepancy, we introduce the **Mean Absolute Percentage Error** (MAPE) as a more equitable metric.

$$\text{MAPE} := \sum_{i}^{n} \left| \frac{\mathcal{D}_i - f(t, \hat{\theta})}{\mathcal{D}_i} \right|.$$

Unlike the RMSE, which weights errors by their square in linear units, the MAPE evaluates the average relative deviation. Surging from the intrinsic multiplicative nature of biological growth processes, we should not ignore possibility for variance to change over time. As discussed in the literature (Ref. [17]), biological variation is often proportional to the mean leading to heteroskedasticity, a phenomenon that additive models fail to capture.

Log-Normal distribution

Definition 1. A dataset $\mathcal{D}$ follows a *log-Normal* distribution if its corresponding log-transformed set of points follows a Normal distribution.

From it, we derive that $\log(x) \sim \mathcal{N}(\mu, \sigma^2)$ has mean μ and variance σ^2. We take advantage of the "back-transformed" - to data level - values:

$$\mu^* := e^{\mu}, \quad \sigma^* := e^{\sigma}; \tag{11}$$

we then use

$$X \sim \Lambda(\mu^*, (\sigma^*)^2). \tag{12}$$

An immediate consequence of the definition is that the dataset presents symmetry at the log level, but does not resemble a Normal shape in the original space. We remark that if data fit a Normal distribution, they also fit a log-Normal one, while the opposite is not true; this perspective allows to see the log one as a generalisation of the standard Gaussian [17]. The probability density function of a log-Normal distributed random variable x is given by:

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(\log(x) - \mu)^2}{2\sigma^2} \right).$$

The median of the log-Normal distribution is μ^*, since μ is the median of $\log(X)$. Thus, chances for a value to be greater than μ^* are 50%, just as they are for it being less than that. Instead, the **multiplicative standard deviation** σ^* accounts for the aforementioned multiplicative effect of this distribution, determining its shape. Whilst the former is a scale parameter responding accordingly to changes in physical units, the latter does not get affected.

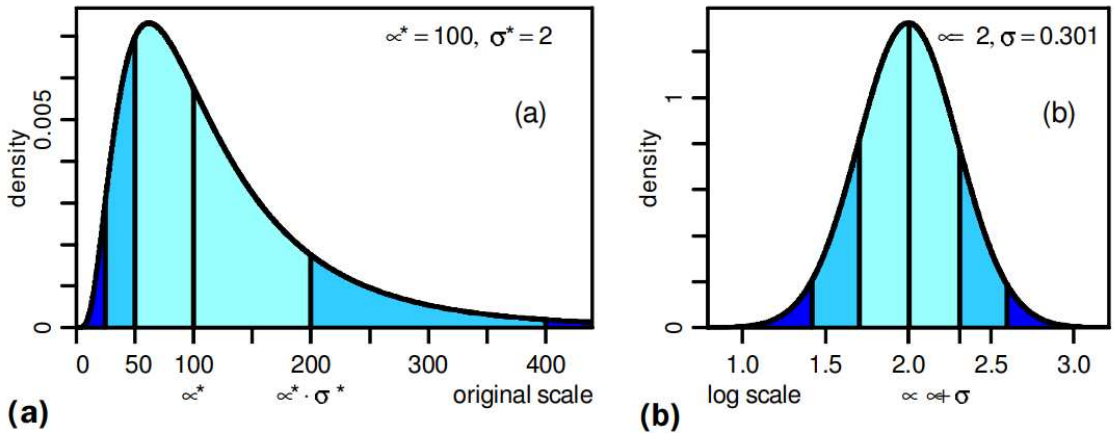


Figure 3.2: A log-normal distribution with original scale (a) and with logarithmic scale (b). Areas under the curve, from the median to both sides, correspond to one and two standard deviation ranges of the normal distribution.

An interesting feature of Gaussian distribution is that

$$\mu \pm \sigma \simeq 68.3\%, \quad \mu \pm 2\sigma \simeq 95.5\%,$$

whereas corresponding statements for the log-normal quantities are

$$\mu^* \circledast \sigma^* \simeq 68.3\%, \quad \mu^* \circledast (\sigma^*)^2 \simeq 95.5\%$$

with the introduced notation[12]. In order to stay coherent with previous analysis, we express data in their original form whenever possible to promote comprehension on the topic.

This distributions holds interesting results, one of which is a multiplicative generalisation of the Central Limit Theorem[17] - stating that under appropriate conditions, the normalised version of the sample mean converges to a standard normal distribution. Since the logarithm of a product is simply the sum of the logarithms of the factors. Therefore, when the logarithm of a product of random variables that take only positive values approaches a normal distribution, the product itself approaches a log-normal distribution.

Whilst the Gaussian curve yields an additive effect on values, log-normality accounts for a multiplicative influence. This is particularly relevant for asymmetrical processes, where two or more events can happen with different probabilities. In a general case, the possibility of death and birth of cancer cells in the context of oncolytic virotherapy can be regarded as highly asymmetrical.

To show the relevance of such an effect, we explicitly computed the σ^* by simply transforming back the corresponding log value; this number determines how skewed the curve is. Across scientific literature, analysis have been conducted addressing this number, suggesting that in many Normal-labelled cases, hetero-skedasticity was present: ranging from phyto-medicine to social sciences, from food technology to human medicine, where latent period of diseases fitted low values as well as being greater than 2 [17]. Authors state that values below 1.2 "*may even be common and therefor of great interest in science*".

[12]The $\pm$ symbol means "plus/minus"; we denote the symbol $\circledast$ as "**times/divide**". Explicitly, $\mu^* \circledast \sigma^* := [\mu^*/\sigma^*, \mu^* \cdot \sigma^*]$.

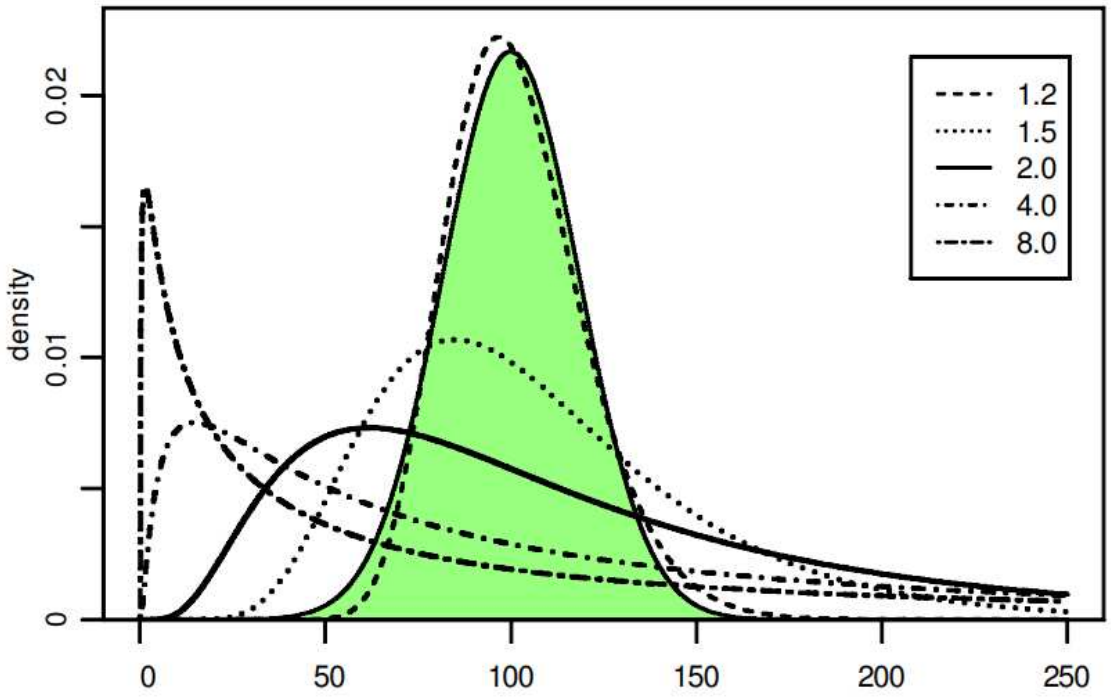


Figure 3.3: Density functions of selected log-normal distributions for various σ^* values compared with a normal distribution. Changes in μ^* affect the scaling in horizontal and vertical directions, but the essential shape σ^* remains the same.

Moreover, we can consider the **Coefficient of Variation** (CV), a measure for relative variability across the dataset,

$$\mathrm{CV}_\Lambda \coloneqq \sqrt{\exp(\sigma^2) - 1},$$

which should align with MAPE's results.

Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) for the log-Normal model is performed by maximising the ℓ function in the logarithmic space. However, a direct comparison between the AIC of a Normal model and a log-normal one is mathematically invalid without a proper change of variables. Since the fitting is conducted on the log-transformed data, we must account for the stretching of the probability space through the *Jacobian term* of the transformation.

To compute this term from a set of observations y, we apply the transform $x = \log(y)$, so that $X \sim \mathcal{N}(\mu, \sigma^2)$. Hence

$$f_Y(y) = f_X(x) \cdot \left| \frac{dx}{dy} \right|,$$

thus substituting $\dfrac{dx}{dy} = \dfrac{1}{y}$,

$$f_Y(y) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(\log y - \mu)^2}{2\sigma^2} \right)}_{\text{Normal}} \cdot \underbrace{\frac{1}{y}}_{\text{Jacobian}}.$$

In the same fashion as in chapter 2, we apply the logarithm to the product of the total probability; this results in the addition of the term $\sum \log(y_i)$ to the $-\ell$ obtained from the Gaussian fit of the log-transformed data. Thus,

$$-\hat{\ell} = \frac{n}{2}\log(2\pi) + n\log\sigma + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log y_i - \mu)^2 + \sum_{i=1}^{n}\log y_i. \tag{13}$$

This Jacobian correction rescales the log-Normal likelihood to the same units as the Normal one, allowing the AIC_c to function as a legitimate coefficient for of model parsimony.

3.2 Fit and Forecast

Beside the calculation differences dictated by the log-transformation of data, every step in the optimisation process remains the same. It is our assumption that a model performing an optimisation over an equal weighted interval, instead of attributing more importance to end tails, would result in residuals closer to white noise DW range and a lower AIC_c value.

Moreover, by undergoing forecasting analysis, we show the impact of log-Normal σ^* in contrast to Normal σ. In particular, we plot CI regions accounting for both one and two (multiplicative) standard deviation as the uncertainty region for each model. The Normal case describes a "*rectangle*" with constant width and the curve in the middle, whereas the log-Normal distribution creates an asymmetrical "*cone*", since the depiction of data is $\mu^* \divideontimes \sigma^*$. Obviously, greater values of multiplicative standard deviation imply a broader spectrum of uncertainty.

We adopt the convention that models following a Normal distribution are denoted with the subscript $model_N$, whereas those based on a log-Normal specification are marked with the subscript $model_\Lambda$.

Datasets A549

For each datasets of figure 1.4 we compare the two distributions for the best performing model. For both *PBS* and *V* cohorts, Richards achieved lower AIC_c values, with a slightly better residuals profile as we can see from figure 3.4. In particular, it appears to have less AR(1). However, for the last cohort there is no improvement. Parameters and metrics are reported in table 3.1

Due to the shift from absolute errors to relative ones, log-Normal distributed models tend to reduce late values' estimation, thus anticipating the IP. Looking at σ^*, values ranges between 1.02 and 1.05, which does not translate to a strong skewness. We conclude that *V* is the only cohorts for which curves show a noticeable improvement, probably due to the steep final growth, which forces Normal distributed models to adapt; interestingly, $\mathrm{Richards}_\Lambda$ achieves better fit than the Exponential, both AIC_c-wise and on the qualitative side. By considering the profiling process, results show slightly narrower profiles for both *PBS* and *VE* cohorts, while maintaining the same qualitative shapes in every dataset as they showed for the Normal distributed case. See figure 3.5.

It has been proven by literature (Ref. [16]) that the IP is crucial for meaningful forecasting. However, in situations where the tumour overtakes the effects of OVT and the experiment is terminated early, we cannot be sure where it would locate. Learning from this, we tested different interval choices: referred as **middle stage**

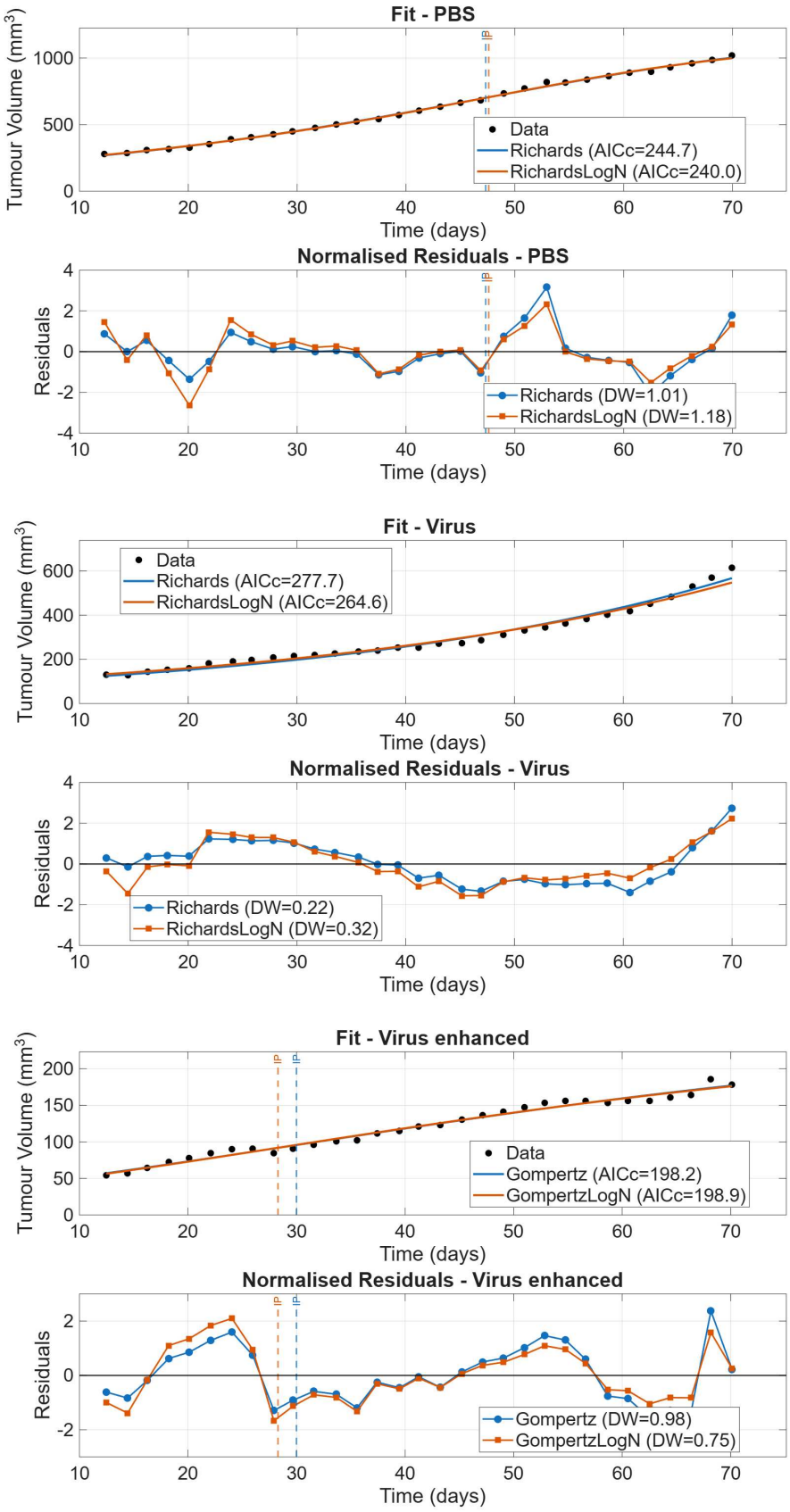


Figure 3.4: Normal and log-Normal fits in *A549* cohorts, showing different trends in the residuals.

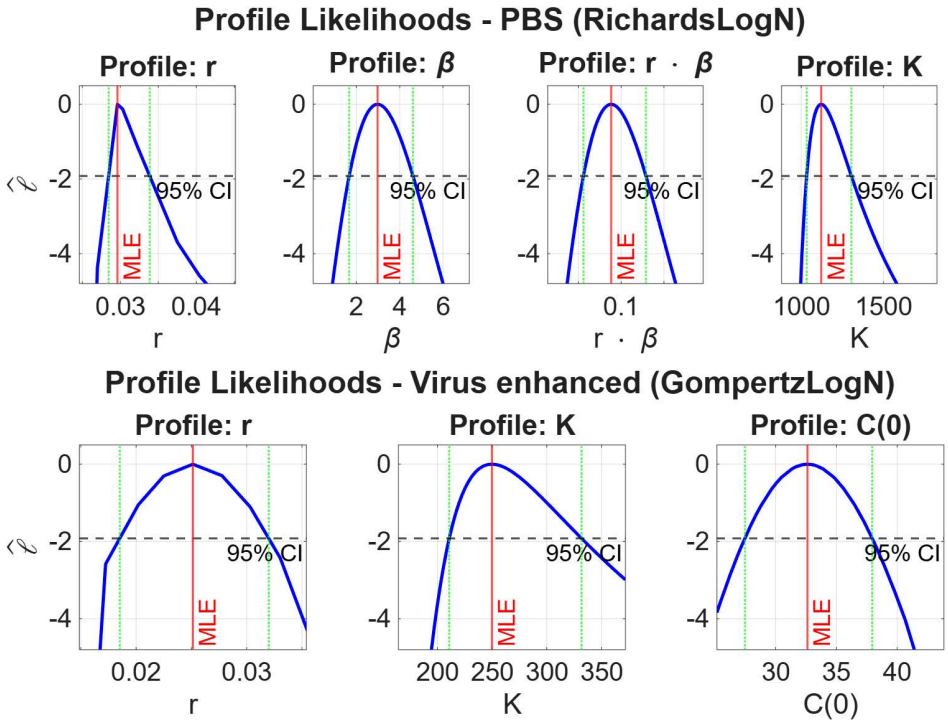


Figure 3.5: *A549* dataset profiles.

Table 3.1: Fit metrics comparison across *A549* cohorts. The ΔAIC_c column refers to the distance between the two models.

Model	**Type**	ΔAIC_c	AIC_c	RMSE	MAPE	DW	σ/σ^*	IP
			Dataset: *PBS*					
Richards	Normal	2.0%	244.7	10.25	1.2%	1.01	10.25	47.3
	log-Normal	0.0%	240.0	10.38	1.3%	1.18	1.02	47.6
			Dataset: *Virus*					
Richards	Normal	5.0%	277.7	17.45	5.2%	0.22	17.45	90.1
	log-Normal	0.0%	264.6	19.16	4.3%	0.32	1.05	93.2
			Dataset: *Virus Enhanced*					
Gompertz	Normal	0.0%	198.2	5.08	3.9%	0.98	5.08	30.0
	log-Normal	0.4%	198.9	5.09	3.9%	0.75	1.05	28.3

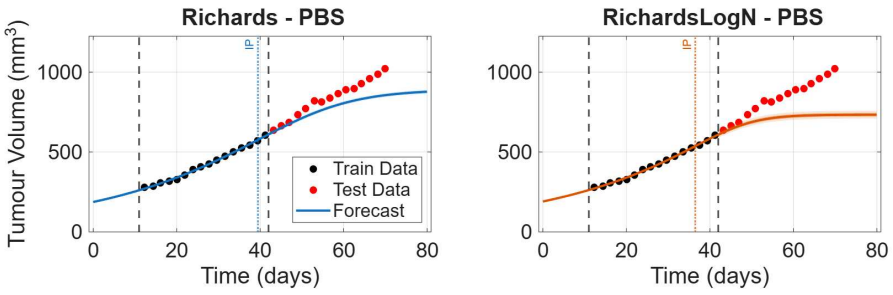
and **late stage**, they account, respectively, for the *train_end* situated at (IP) and (IP+10%). We report results for both PBS and V cohorts in figure 3.6.

We remark that the general underestimation trend performed by log-Normal models leads to lower predicted values. Interestingly, IC bands allow to take this effect into a better account than the Normal case by a broader, yet more adherent to reality, forecast horizon. Interestingly, it is of note that we encounter the same pattern over several datasets.

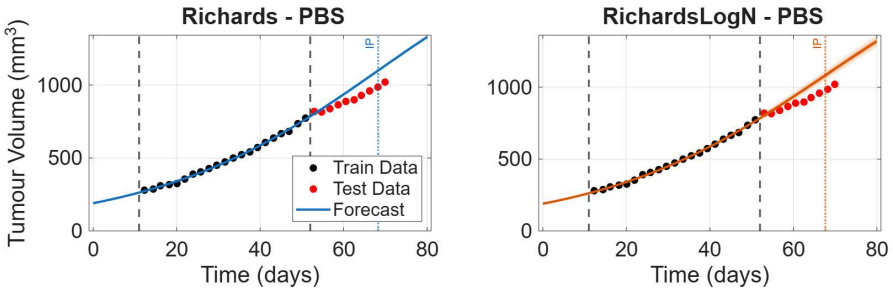
Datasets *rA549*

We now consider Ref. [10] and refer to it as the *rA549* datasets, in figure 1.7. *VE* group faces a stronger anti-tumour reaction after nadir, reached at day 5.

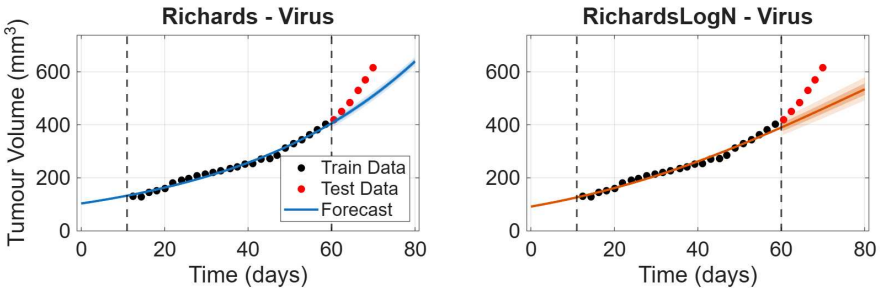
Figure 3.7 shows a comparison between the three sigmoids, both Normally and log-Normally distributed. Interestingly, adopting a skewed statistic yields significant



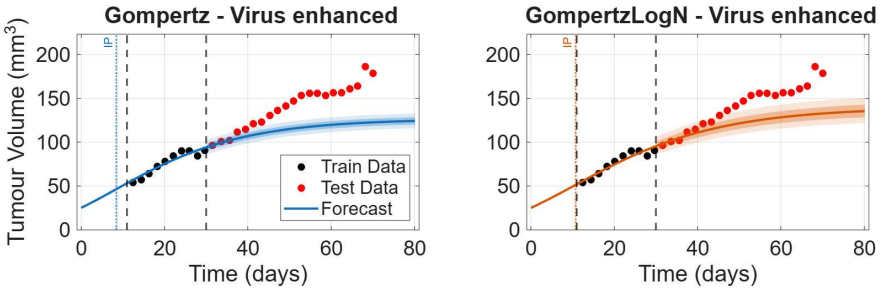
(a) Middle stage *PBS*



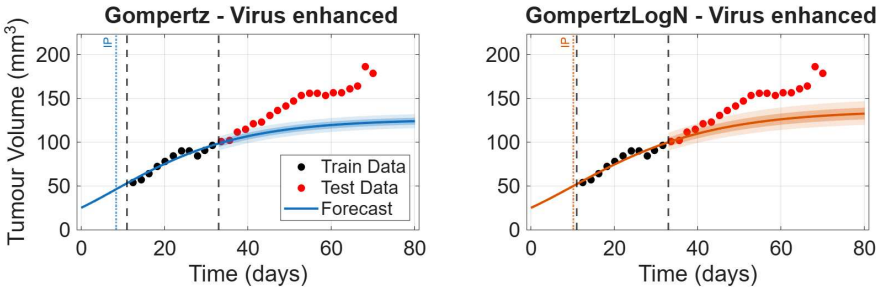
(b) Late stage *PBS*



(c) V dataset with $train\_end = 60$.



(d) Middle stage *VE*



(e) Late stage *VE*

Figure 3.6: Forecasting of *A549* cohorts. ICs are plotted for confidences of 1σ and 2σ.

Table 3.2: Fit metrics comparison across all *rA549* datasets.

Model	Distribution	AIC_c	RMSE	MAPE	DW	σ/σ^*	IP
Dataset *PBS*							
Logistic	Normal	196.5	41.74	7.5	1.91	41.74	44.2
	log-Normal	177.6	44.95	6.0	2.19	1.07	40.3
Gompertz	Normal	194.5	39.51	5.5	2.03	39.51	48.4
	log-Normal	176.0	40.59	5.2	2.20	1.07	52.4
Richards	Normal	198.3	39.37	5.2	2.06	39.37	46.0
	log-Normal	178.1	39.62	5.0	2.49	1.06	43.8
Dataset *Virus*							
Logistic	Normal	181.3	27.37	9.0	0.98	27.37	37.3
	log-Normal	169.7	28.18	6.9	1.34	1.08	36.5
Gompertz	Normal	182.0	27.93	13.6	0.77	27.93	39.2
	log-Normal	187.8	38.25	10.2	0.56	1.13	51.0
Richards	Normal	181.5	24.69	8.5	1.05	24.69	36.6
	log-Normal	173.4	29.54	6.9	1.36	1.08	36.2
Dataset *Virus Enhanced*							
Logistic	Normal	177.8	19.59	10.2	1.10	19.59	43.4
	log-Normal	170.5	20.04	10.2	0.93	1.12	45.9
Gompertz	Normal	176.3	18.80	11.9	1.05	18.80	52.2
	log-Normal	177.0	19.86	11.5	0.73	1.15	56.5
Richards	Normal	179.7	18.65	10.9	1.11	18.65	46.3
	log-Normal	172.1	23.04	9.4	1.05	1.12	42.4

improvements, ranging, on average, from $4-10\%$ by AIC_c for each model, as reported in table 3.2. Whilst the measured drop in this metric is greater in cohorts including highest values, σ^* values increase in V and VE datasets, even though their right-hand tails have smaller volume counts. Our guess is that in situations where OVT is successfully limiting the steepness, there is the emergence of a more pronounced multiplicative profile and the log-Normal hypothesis better suits such context. If this pattern occurs in several scenarios, it would attribute a considerable importance to the distribution in cancer therapies' research.

Other common aspects among log-Normal distributed models, is a lower MAPE and a generally better looking residuals profile. As we explained before, RMSE may differ from the relative errors consideration due to absolute larger final values. Furthermore, it appears that in the symmetrical statistic case, models tend to agree more, while in the log-Normal distribution curves manifest dissimilar trajectories.

We conducted a forecasting analysis with training intervals consisting of all points between nadir and IP of each dataset; results are shown in figure 3.8. Unfortunately, Richards often appears to be misspecified, with β values greater than 2; we suspect that the fewer number of points compared to previous datasets somehow induces the generalised logistic to trade conceptual understanding for "artificial" plot adherence, resulting in a the plateau being delayed until reaching the last training point. This suggests that, in some cases, the cost of including the shape parameter is not justified. Profile likelihoods from figure 3.9 confirm this guess: since profiles are wider for V and VE datasets - the last one even holding a not identifiable profile for β - the models possesses the freedom to adapt over a broad interval of parameters but achieve meaningless fits.

Nevertheless, it is difficult to draw quantitative correlations for MAPE values correspondences from train set to test set and it is not obvious whether the best fitting models may or may not be also the best at predicting future values. What emerges qualitatively is that the multiplicative standard deviation carries a strong

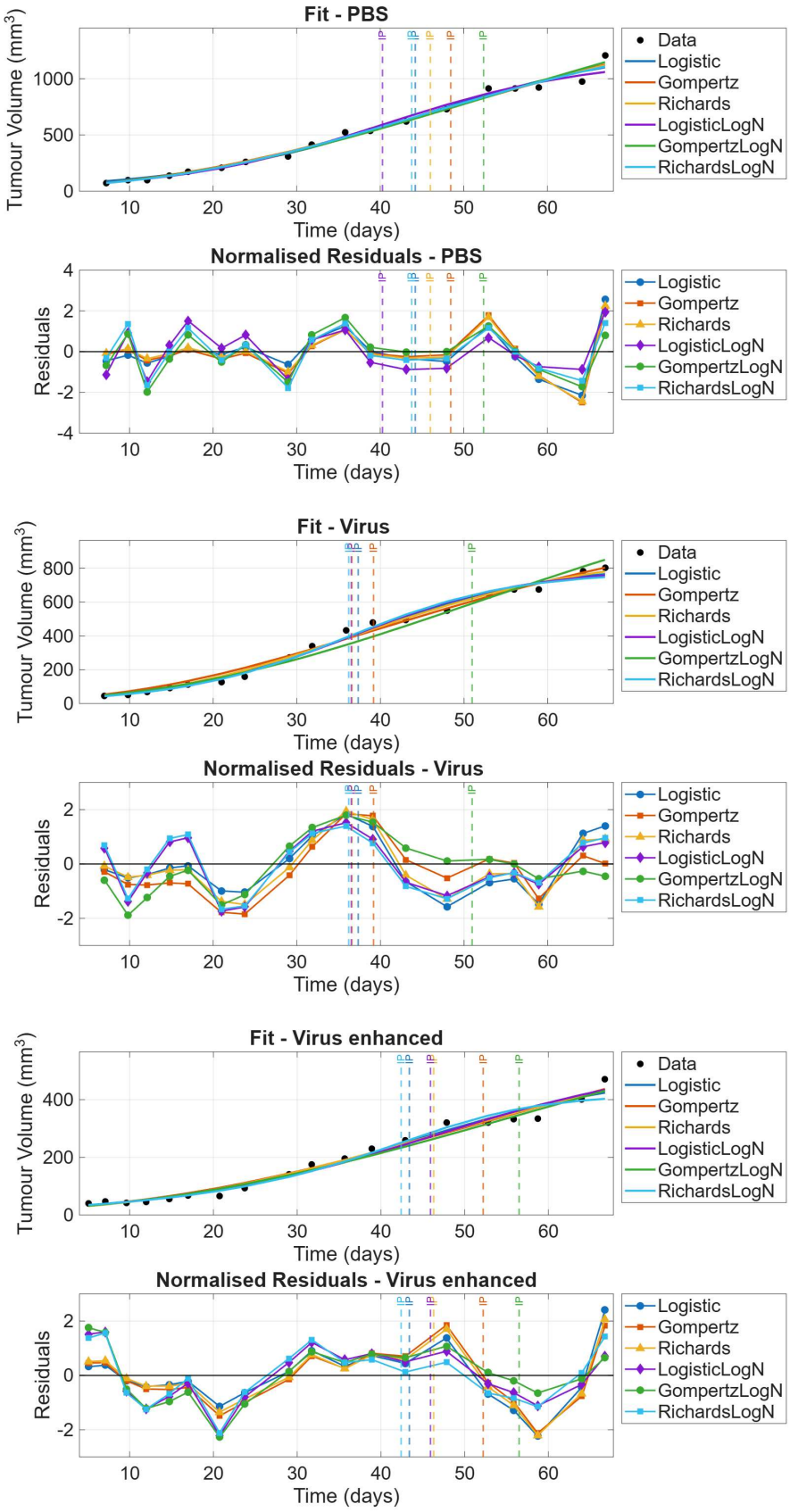


Figure 3.7: Normal and log-Normal models fit in *rA549* cohorts from figure 1.7.

potential for forecasting purposes, possibly due to the intrinsic asymmetry of the processes involved in oncolytic virotherapy.

Other examples

Datasets *U343* We now refer to the *U343* cohorts of figure 1.8 (see Ref. [10]). Figures 3.10, 3.11 and 3.12 shows results from the multi-step analysis of each of the *U343* cohorts from Ref. [10]. Goodness-of-fit metrics are reported in Table 3.3. We set $n = 5$ as the pre-nadir reach.

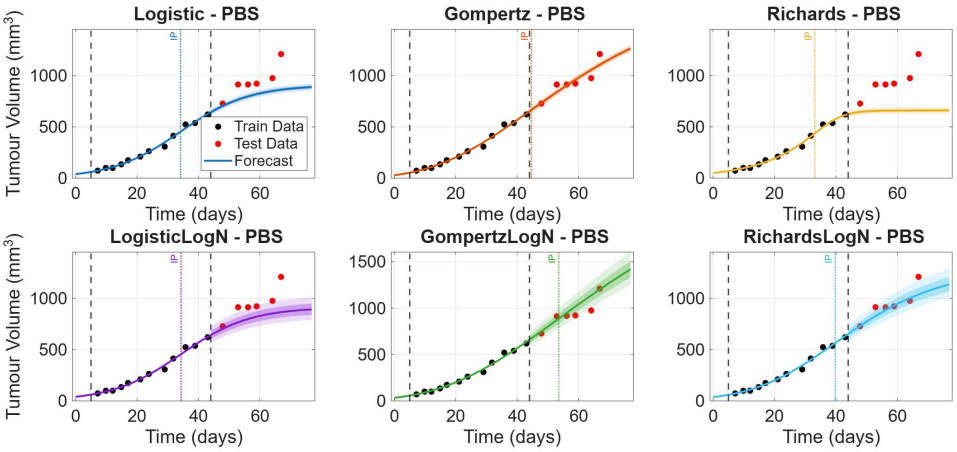
In every performed fit, the best model, according to both AIC_c and RMSE, is the Logistic curve with the log-Normal distribution. In general, for every curve the skewed statistic inclusion performs sensibly better than the correspondent Normal case. Coherently to previous examples, MAPE precision favours the log-Normal distribution in every case, whilst the RMSE metric "warns" about the physical distance caused by end-values underestimation. Values of σ^* are larger for *V* and *VE* cohorts, further justifying the multiplicative explanation of the underlying biological dynamic. Finally, we observe that forecasting shows interesting results, accounting for realistic predictions in most of the log-Normal distributions.

Persistence property's hypothesis is further reinforced by resulting, for each dataset in a form of resilience adopted by the Logistic curve with the skewed statistic. The only exception is an emerging trend of Normal Logistic choice that seems to appear along considerable points removal and the consequent rise in DW, reaching white nose, no AR(1) value of two. Parallel to such trend, we observe a sudden increase of the RSE metric, probably due to the risk for models to misinterpret the "vanishing signal" from datasets.

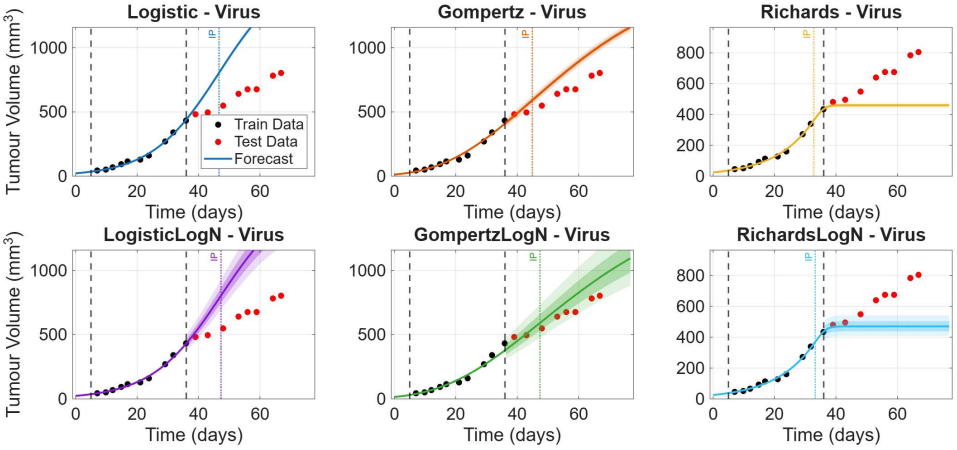
We explicitly point out the major benefit of dealing with relative error metrics and multiplicative standard deviation factors. While some experimental data may stretch over similar scaling, it is usually the case for cohorts to resize over magnitudes of difference with respect to an almost identical settings, but with other OVT techniques involved. In the current example, in addition to the different shape of the *PBS* dataset, tumour growth soon overtake others by 2 and 3 times respectively. If we ask whether a curve is fitting properly, or we investigate to which extent the measured error is meaningful, dealing with a standardise measure such as MAPE and an asymmetry shape term σ^* is much more meaningful and convenient, while also easier to discuss, compare and generalise. For example, in this same experimental case, the *PBS* dataset results in better accuracy than the *V* cohort. If we were to aseptically look at RMSE values between such cases, they would look almost identical, but we would not be being aware that the latter accounts for roughly twice the MAPE value compared to the former. Finally, by looking at σ^* we expect a broader forecast cone that, in almost each log-Normal distributed model, leads to a more accurate prediction of the test set; the only exception is a misspecification trend -that we already discussed - interesting the Richards curve.

Datasets *sA549* We now consider data from Ref. [11], for which figure 1.9 presents an overview. Following the same procedure as in previous examples, we plot fit results for each cohorts, named as usual and referred as *sA549* datasets. We get an insight of fit quality through figure 3.13, while goodness-of-fit metrics are reported in Table 3.5.

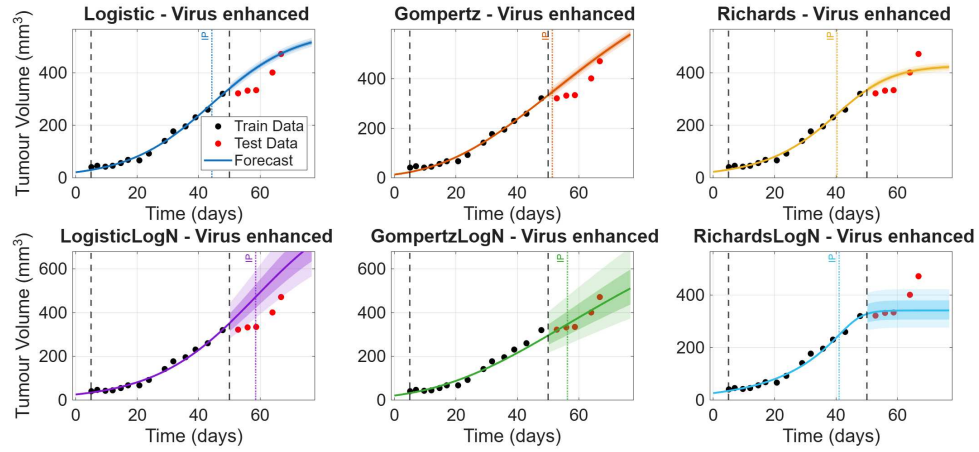
Fits difference, here, is less noticeable as in *V* and *VE* datasets every curve is quantitatively close to each other by AIC_c. The main remark on this step of



(a) *PBS* cohort.

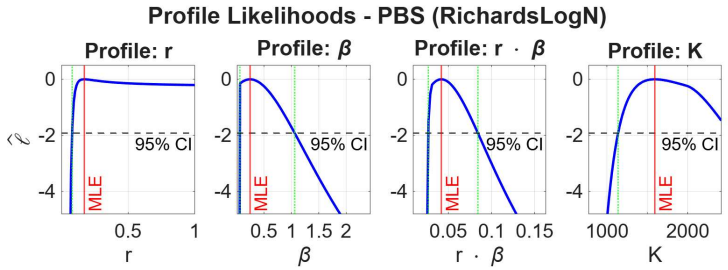


(b) *V* cohort.

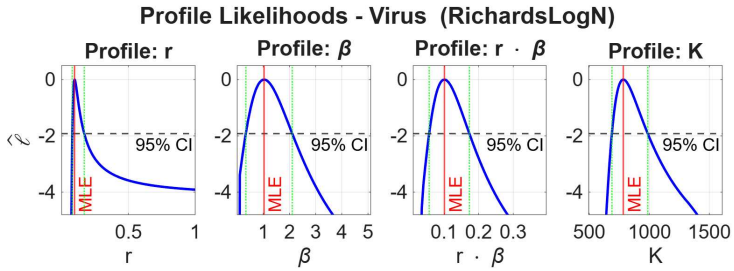


(c) *VE* cohort.

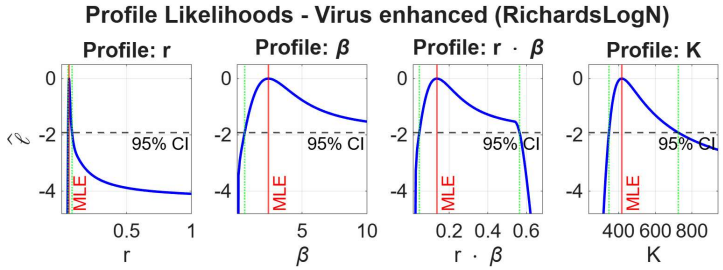
Figure 3.8: Normal and log-Normal models forecasting in *rA549* cohorts.



(a) *PBS* cohort.



(b) *V* cohort.



(c) *VE* cohort.

Figure 3.9: Profiling conducted on *rA549* datasets for Richards_Λ.

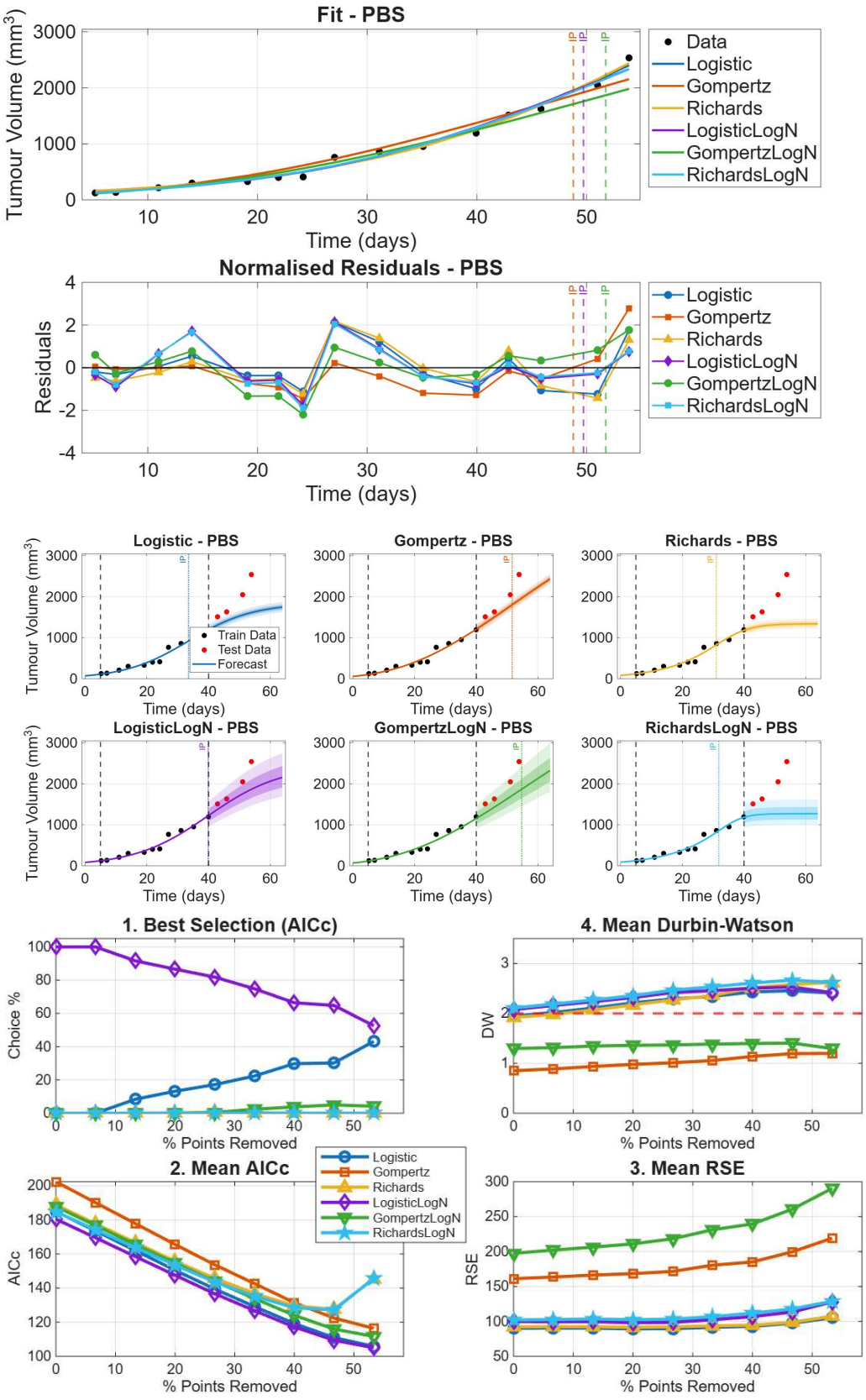


Figure 3.10: Normal and log-Normal fits, forecast and persistence analysis conducted on the *U343 PBS* dataset from figure 1.8.

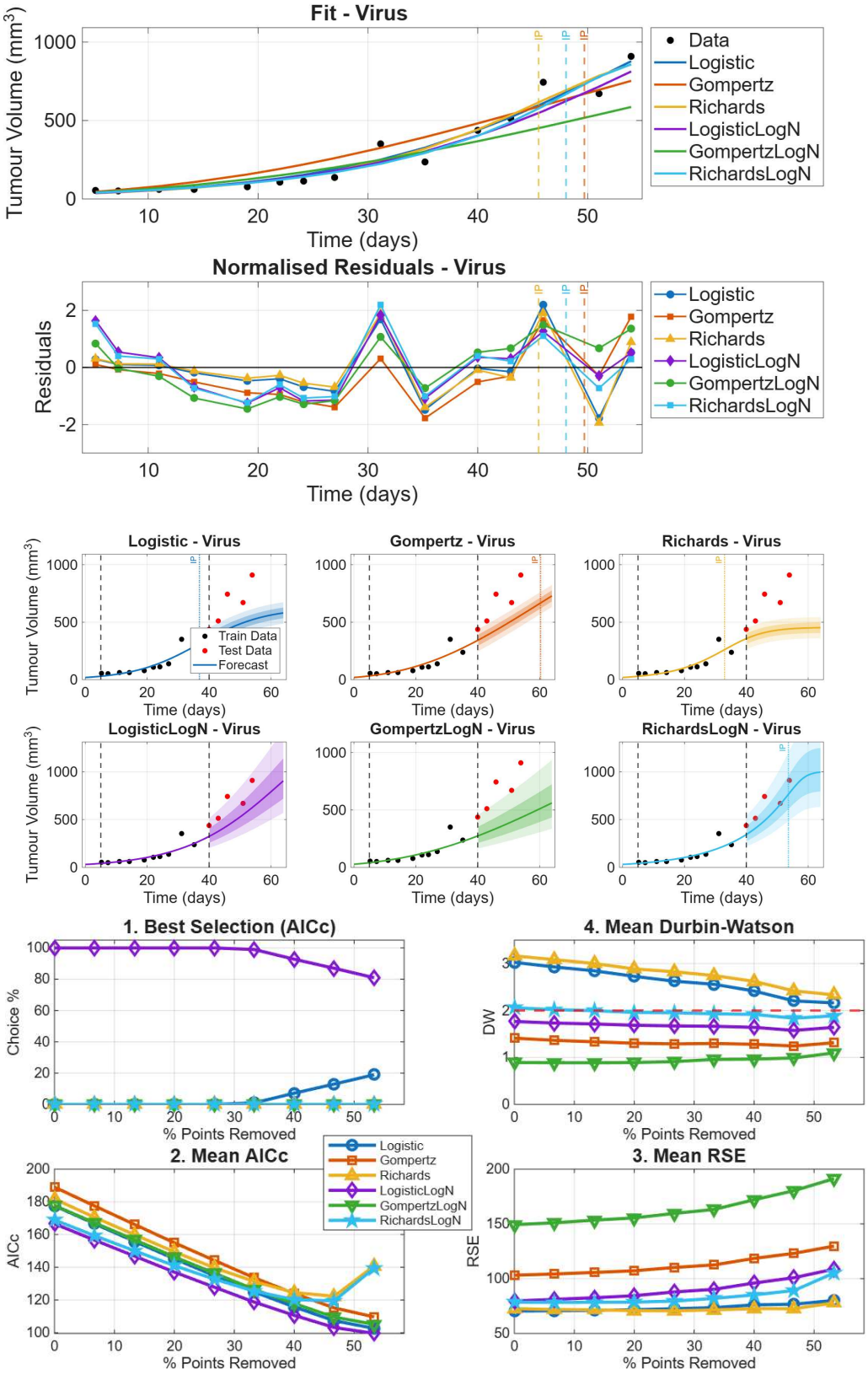


Figure 3.11: Normal and log-Normal fits, forecast and persistence analysis conducted on the *U343 V* dataset from figure 1.8.

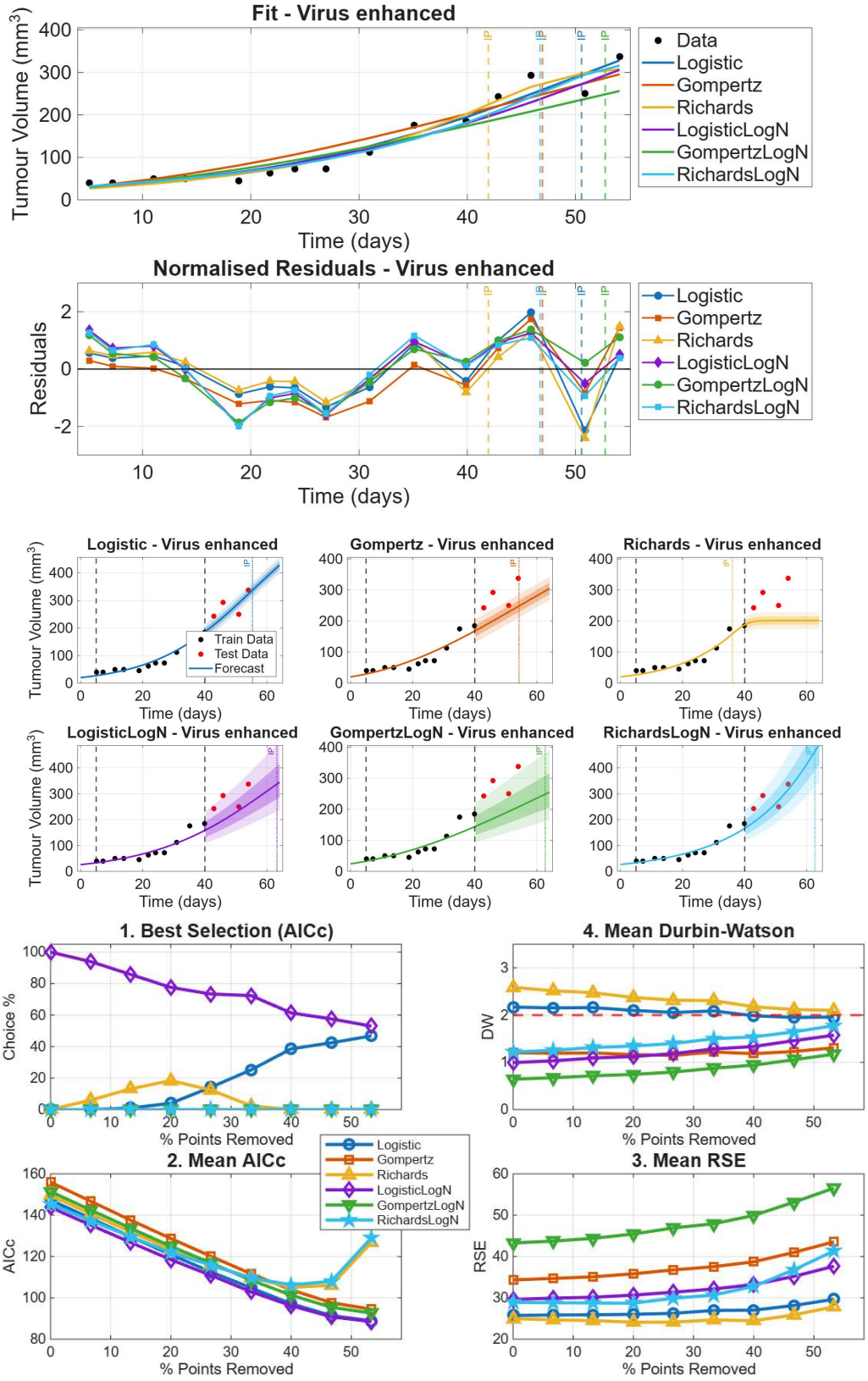


Figure 3.12: Normal and log-Normal fits, forecast and persistence analysis conducted on the *U343 VE* dataset from figure 1.8.

Table 3.3: Fit metrics comparison across all *U343* cohorts. The ΔAIC_c column refers to the distance between specified model and the lowest achieved value for that dataset.

Model	**Type**	ΔAIC_c	AIC_c	RMSE	MAPE	DW	σ/σ^*	IP
Dataset: *PBS*								
Logistic	Normal	2.5%	185.0	77.23	9.4%	1.95	77.23	55.4
	log-Normal	0.0%	180.5	85.52	8.9%	2.08	1.12	49.7
Gompertz	Normal	12.2%	202.4	138.03	13.1%	0.85	138.03	48.8
	log-Normal	4.0%	187.7	169.07	11.6%	1.30	1.15	51.7
Richards	Normal	4.7%	188.9	75.30	12.0%	1.92	75.30	57.1
	log-Normal	2.4%	184.7	83.43	8.8%	2.11	1.11	54.2
Dataset: *Virus*								
Logistic	Normal	6.6%	177.6	60.36	20.6%	3.03	60.36	54.0
	log-Normal	0.0%	166.6	68.17	19.5%	1.77	1.25	56.8
Gompertz	Normal	13.5%	189.0	88.32	41.7%	1.41	88.32	49.7
	log-Normal	6.7%	177.8	127.83	30.7%	0.89	1.38	59.0
Richards	Normal	9.1%	181.7	59.34	19.0%	3.16	59.34	45.5
	log-Normal	1.4%	168.9	64.19	17.5%	2.06	1.23	48.0
Dataset: *Virus Enhanced*								
Logistic	Normal	2.5%	147.4	22.05	18.3%	2.17	22.05	50.6
	log-Normal	0.0%	143.8	25.37	16.9%	0.99	1.21	57.3
Gompertz	Normal	8.5%	156.0	29.41	25.7%	1.21	29.41	46.9
	log-Normal	5.4%	151.5	37.06	22.4%	0.64	1.28	52.7
Richards	Normal	4.1%	149.7	20.39	16.8%	2.59	20.39	41.9
	log-Normal	1.2%	145.4	23.59	15.2%	1.22	1.19	46.7

analysis is the σ^* value revolving around 1.09 range for most fits. We point out that persistence analysis shows result following the same "laws" already observed, with an emerging choice of Logistic for either Normal or log-Normal distribution; the only exception is represented by the *VE* cohort for which Gompertz leads in both statistic cases. We recall that in other experimental data, such as those of figure 1.4, Gompertz curve was the best at modelling slow, delayed, growth and that is enough to explain the behaviour of best model selection across points removal. Again, in the last two steps , once no residuals autocorrelation is left and contemporary to the sudden rise of RSE, both Logistic distributions gain roughly 10% of choice. Another self-repeating trend is the sensible increase of AIC_c metric for Richards model over last few steps.

We then undergo forecasting for each dataset. By choosing an early stage or a late stage training sets, we measure improvements for each model, see figure 3.15 and 3.16. The same pattern over metrics and significance on parameters previously describe occurs here, further justifying chosen methods and conclusions.

Further mentions In the context of OVT, more advanced GE OVs have an experimental advantage that is sometimes magnitudes of difference from original viruses. The formers, for example, incorporating mutations for improved behaviour or being encapsulated in order to hide from Immune System. Moreover, OVT now achieves better and more promising results by combining different therapies. As we categorised more complex tumour-virus dynamics with the name of *Category II*, we conduct a similar multi-step analysis.

In 1 we make reference to Ref. [9] datasets, plotted in figure 1.6. In the cited

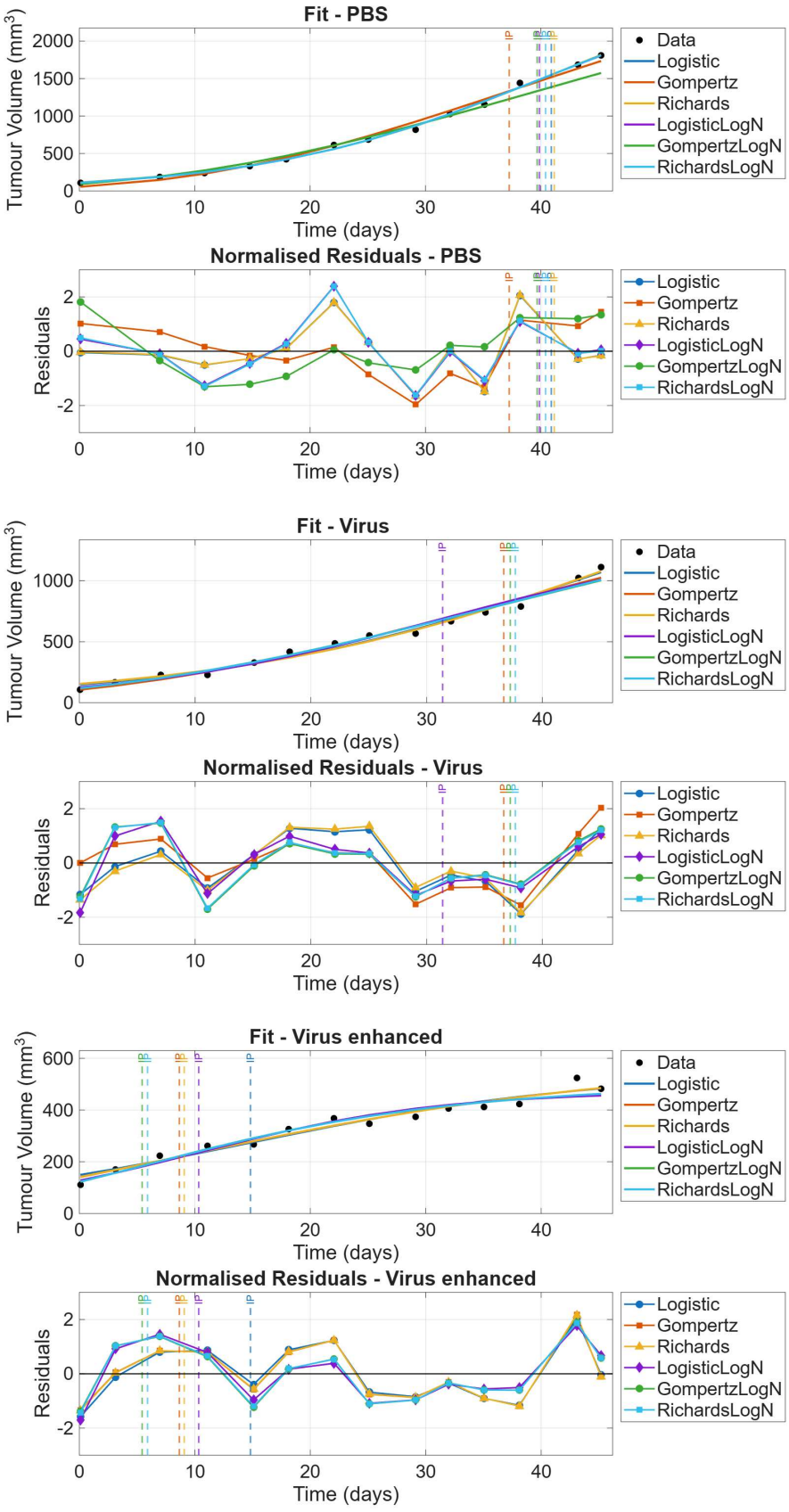


Figure 3.13: Normal and log-Normal models fit in *sA549* cohorts in figure 1.9 from Ref. [11].

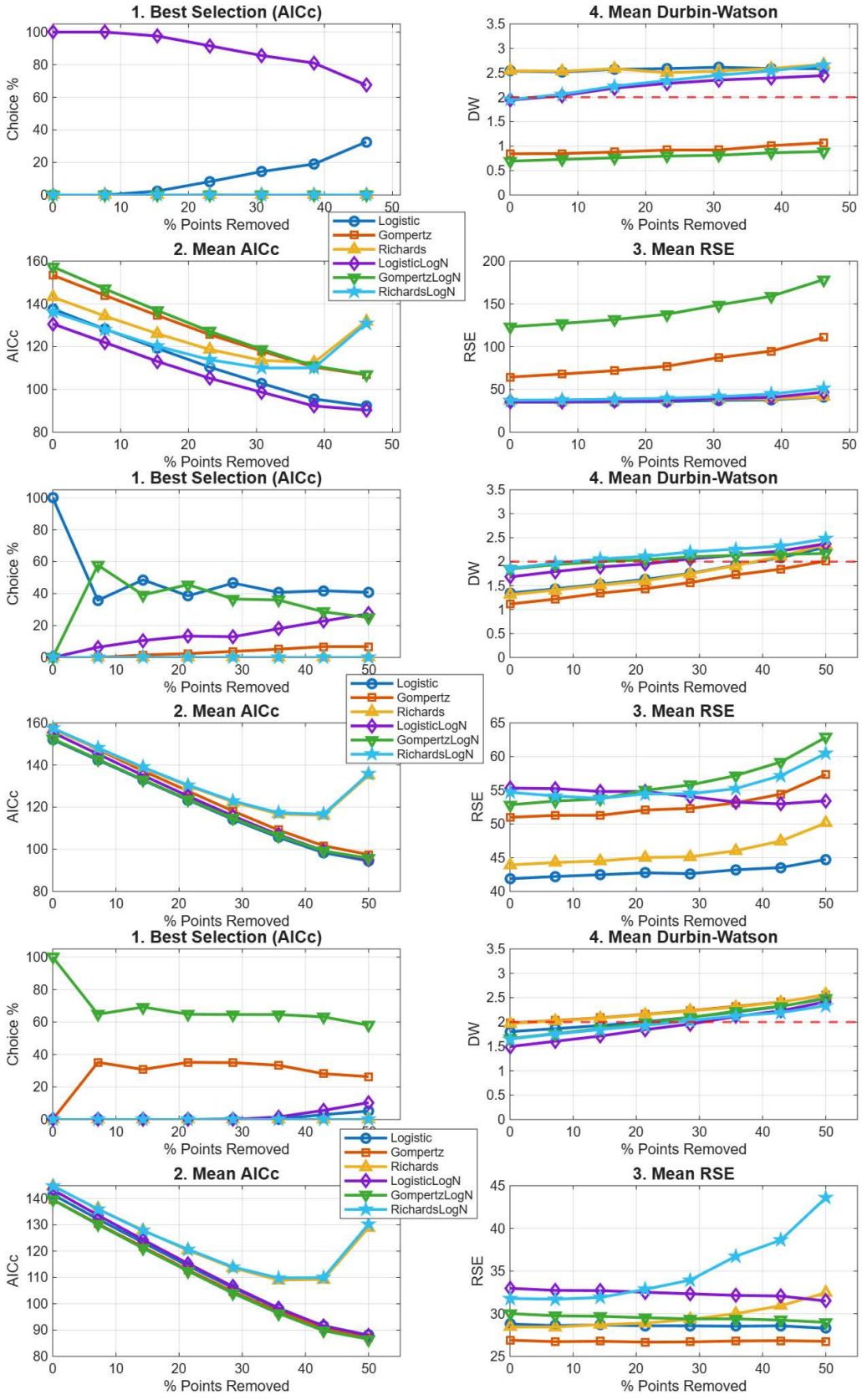
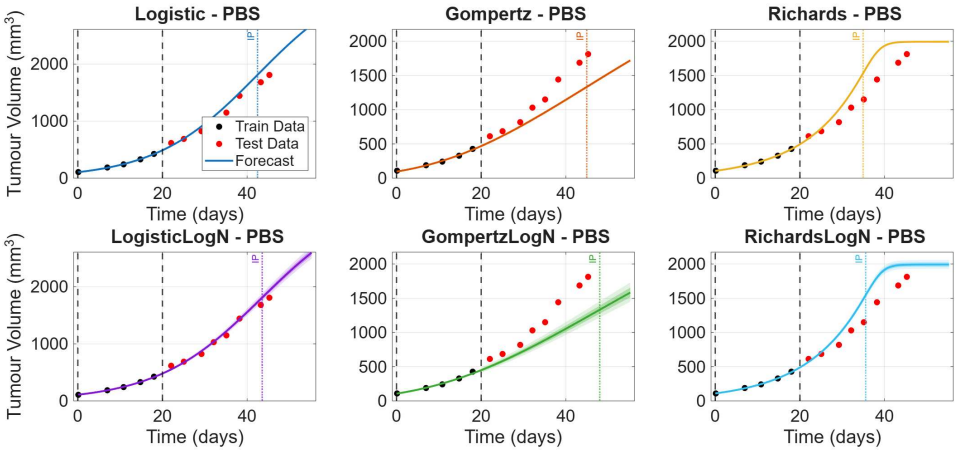
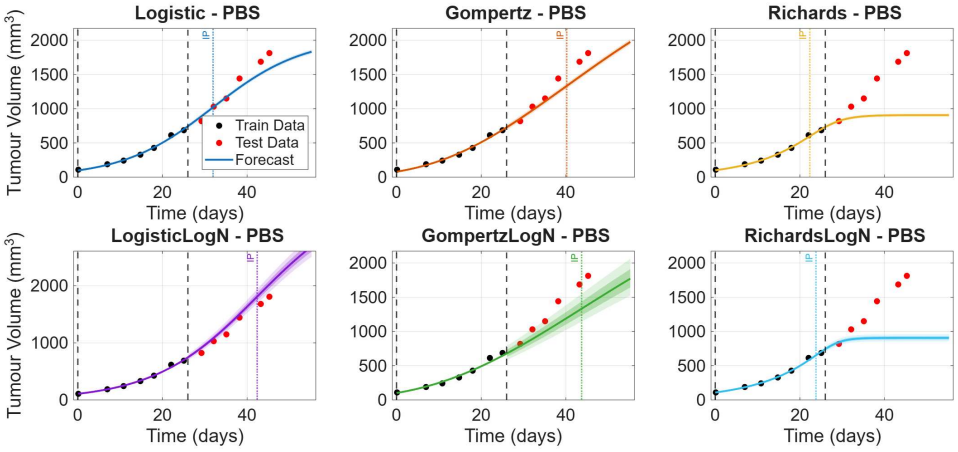


Figure 3.14: Normal and log-Normal models persistence analysis in *sA549* cohorts.

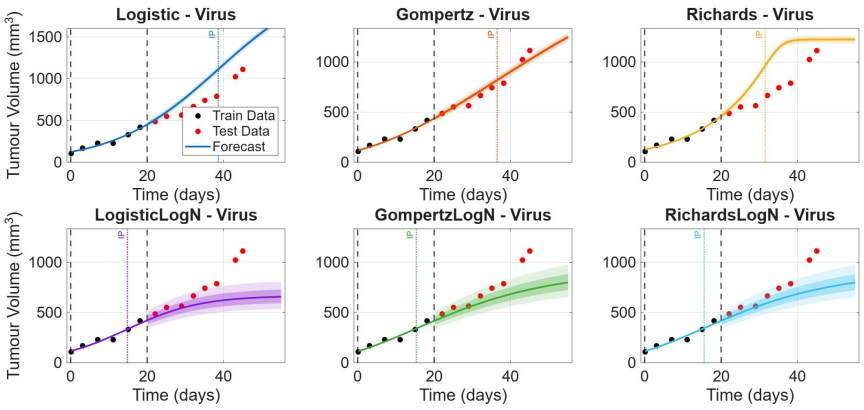


(a) train_end = 20.

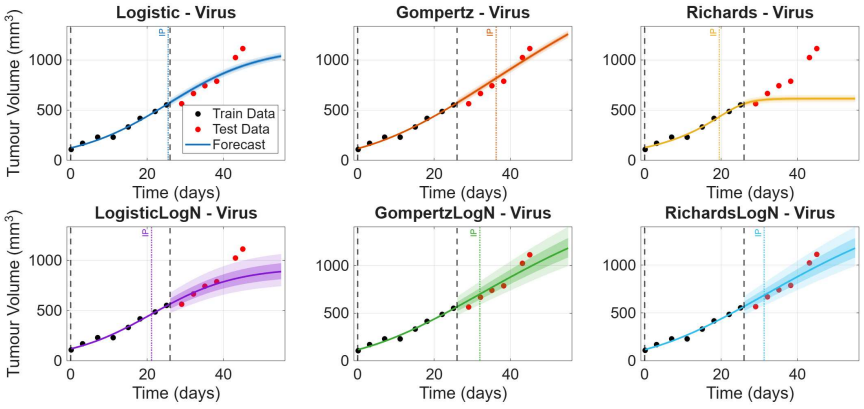


(b) train_end = 26.

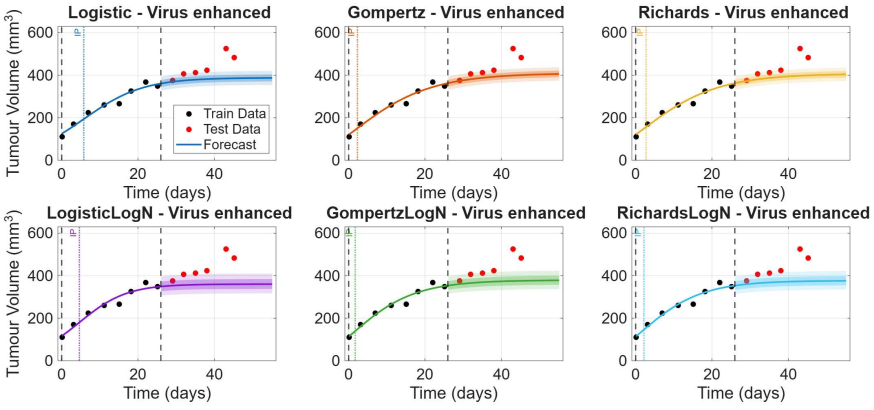
Figure 3.15: Normal and log-Normal models forecast in *sA549 PBS* cohort for early stage and late stage train sets.



(a) V cohort early stage.

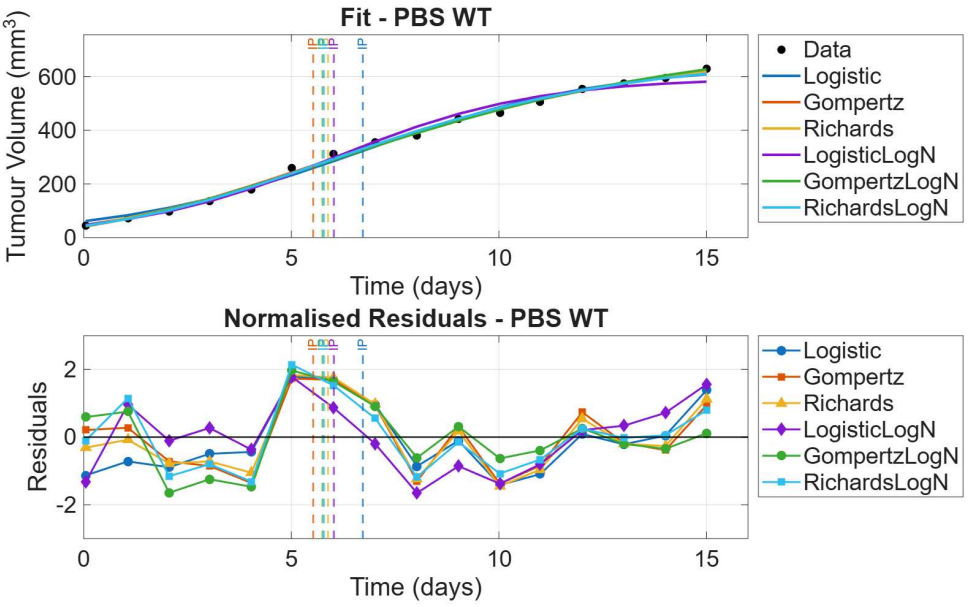


(b) V cohort late stage.

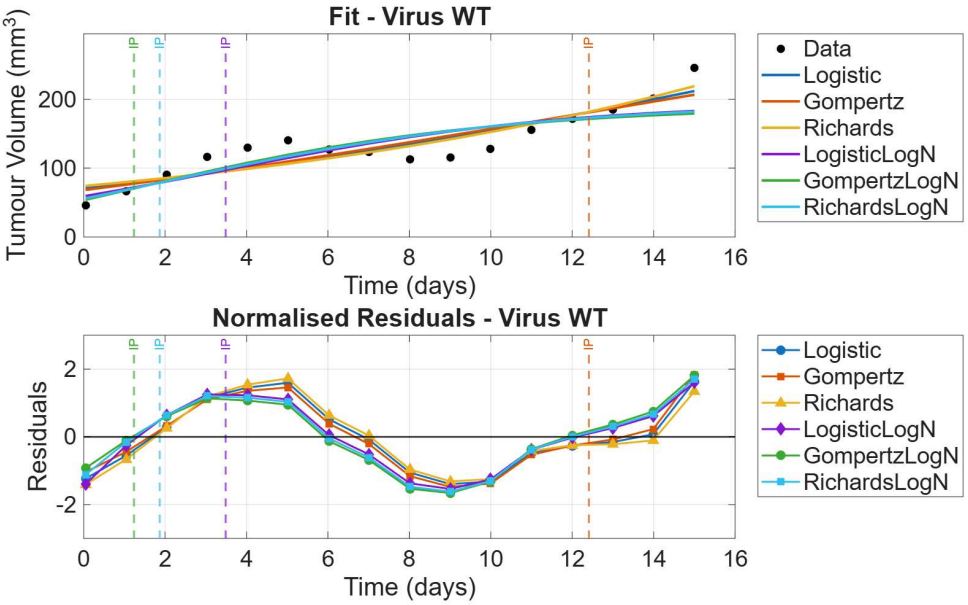


(c) VE cohort late stage.

Figure 3.16: Normal and log-Normal models forecast in *sA549 Virus* cohort for early stage train_end = 20 and late stage train_end = 26 train sets.



(a) *PBS* cohort fit acts as in others *Category I* examples.



(b) *Chimeric Category II* agent and residual trends. It appears that each sigmoid is just "cutting" through the data shape, as residuals show the same wavy structure.

Figure 3.17: Normal and log-Normal fit conducted on datasets from figure 1.6.

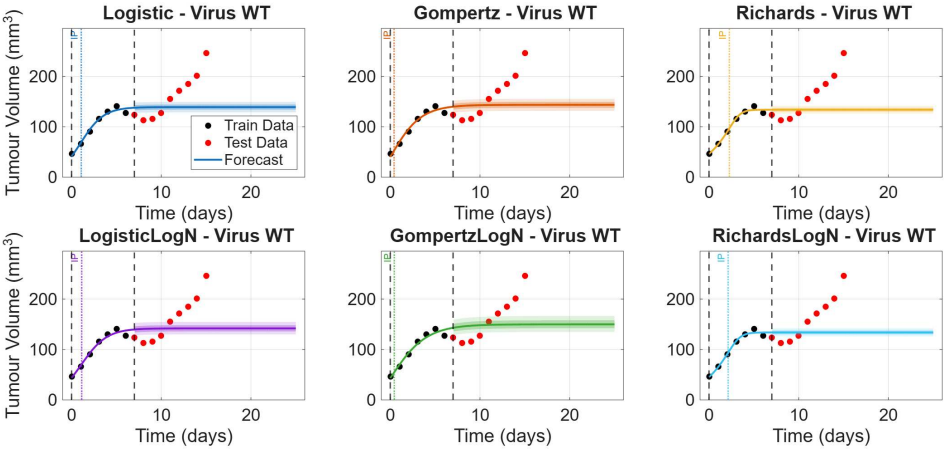
Table 3.4: Forecast metrics comparison across all *U343* cohorts. We recall that ratios closer to one account for more coherence between train and test accuracy. Again, MAPE metrics facilitates the evaluation of a model's performance. We remark that, since the train interval cuts the right-hand tail of each dataset, trained σ^* values are smaller, all ranging between 1.10 and 1.15, probably due to the asymmetrical momentum gain for this biological phenomenon.

Dataset: *PBS*							
Model	**Distribution**	**RMSE$_{Tr}$**	**RMSE$_{Te}$**	**Ratio**	**MAPE$_{Tr}$**	**MAPE$_{Te}$**	**Ratio**
Logistic	Normal	62.85	455.67	0.14	13.5%	25.0%	0.54
	log-Normal	65.34	197.99	0.33	11.7%	11.3%	1.04
Gompertz	Normal	69.45	260.47	0.27	15.4%	9.3%	1.66
	log-Normal	72.54	405.90	0.18	11.6%	18.3%	0.63
Richards	Normal	60.89	665.93	0.09	12.9%	38.6%	0.33
	log-Normal	63.27	480.05	0.13	11.7%	25.7%	0.46
Dataset: *Virus*							
Logistic	Normal	6.77	295.27	0.02	7.7%	44.7%	0.17
	log-Normal	7.42	331.50	0.02	7.9%	50.5%	0.16
Gompertz	Normal	7.67	355.61	0.02	8.6%	53.1%	0.16
	log-Normal	8.46	381.18	0.02	8.7%	57.6%	0.15
Richards	Normal	6.63	266.21	0.02	7.6%	41.1%	0.18
	log-Normal	7.25	310.52	0.02	7.8%	47.9%	0.16
Dataset: *Virus Enhanced*							
Logistic	Normal	5.12	125.52	0.04	6.6%	47.9%	0.14
	log-Normal	5.18	130.68	0.04	6.9%	50.0%	0.14
Gompertz	Normal	5.24	135.49	0.04	6.7%	51.4%	0.13
	log-Normal	5.30	139.75	0.04	7.0%	53.2%	0.13
Richards	Normal	5.09	120.38	0.04	6.5%	46.2%	0.14
	log-Normal	5.14	126.21	0.04	6.8%	48.5%	0.14

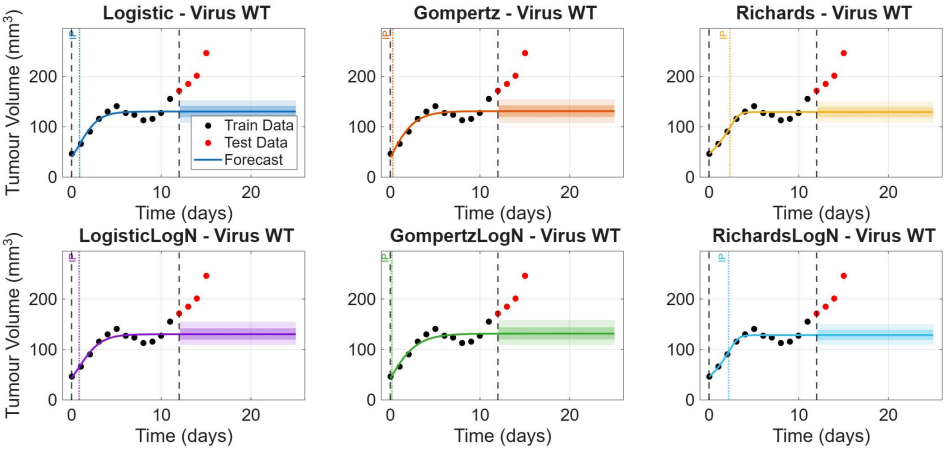
work, OVs are combined, forming a "*chimeric*" therapeutic agent and compared with a control, single treatment, group. We refer to the latter as *PBS* dataset, while the first as *chimeric Category II* cohort; we remark that it is hereby possible to distinguish two main behaviours across datasets. The *PBS* group treated with a form of immunotherapy follows a monotonic, ascending behaviour, whereas the *chimeric* agent achieves more successful results. Data collection begins with tumour proliferation for both group, but in the *Category II* cohort a shift in trend occurs: it reaches a drop in malignant volume at day 5, followed by a slower re-growth. This new kind of dynamic is more effective in limiting growth, yet it represents a challenge for Sigmoidal curves.

Fitting models to both dataset yields to results in figure 3.17. Moreover, we report forecast analysis of both considered datasets in figure 3.18. However, for the *chimeric* agent, fitting and forecasting results are far from being coherent to data. We hence suspend the analysis on such class of experiments since it would not produce any further improvement due to the intrinsic limitations of the considered models.

Both the Normal distributed models and the proposed change in distribution cannot surpass intrinsic limitations of the Sigmoidal shape, accounting for curves not bending anywhere except for the IP. Contrary to the previous examples of *Category I* where, even for complex behaviour, a common trend of reasonably accurate forecasting is possible, we observe that this type of strong, but temporary, tumour



(a) Forecast for *train_end*= 7.



(b) Forecast for *train_end*= 12.

Figure 3.18: Normal and log-Normal forecasting in *Chimeric Category II* cohorts. Every model describes data as plateauing, while not being able to "cope" with the different shape of data.

Table 3.5: Fit metrics comparison across all *sA549* cohorts. The ΔAIC_c column refers to the distance between specified model and the lowest achieved value for that dataset.

Model	**Type**	ΔAIC_c	AIC_c	RMSE	MAPE	DW	σ/σ^*	IP
Dataset: *PBS*								
Logistic	Normal	5.4%	137.7	29.24	2.9%	2.54	29.24	40.9
	log-Normal	0.0%	130.7	29.48	2.6%	1.94	1.04	39.9
Gompertz	Normal	17.4%	153.4	53.61	9.5%	0.84	53.61	37.3
	log-Normal	20.4%	157.3	102.56	8.7%	0.69	1.11	39.7
Richards	Normal	9.6%	143.2	29.24	2.8%	2.54	29.24	41.2
	log-Normal	4.3%	136.2	29.38	2.6%	1.95	1.04	40.4
Dataset: *Virus*								
Logistic	Normal	0.0%	152.0	35.39	8.5%	1.35	35.39	46.5
	log-Normal	2.2%	155.4	46.77	8.4%	1.68	1.10	31.4
Gompertz	Normal	3.6%	157.6	43.10	7.4%	1.12	43.10	36.7
	log-Normal	0.3%	152.5	44.66	7.3%	1.85	1.09	37.2
Richards	Normal	3.2%	157.0	35.21	9.0%	1.31	35.21	47.8
	log-Normal	3.6%	157.5	43.87	7.3%	1.86	1.09	37.7
Dataset: *Virus Enhanced*								
Logistic	Normal	1.4%	141.5	24.32	7.5%	1.80	24.32	14.8
	log-Normal	2.8%	143.4	27.87	7.3%	1.50	1.09	10.3
Gompertz	Normal	0.1%	139.6	22.73	6.8%	1.98	22.73	8.7
	log-Normal	0.0%	139.5	25.37	6.4%	1.66	1.08	5.4
Richards	Normal	3.8%	144.8	22.80	6.8%	1.97	22.81	9.1
	log-Normal	3.8%	144.9	25.48	6.5%	1.65	1.08	5.9

suppression dynamic shall not be modelled by sigmoids. We conclude that this set of advanced GE anti-cancer agents follow more complex behaviours, for which other models are required.

3.3 *Lex Parsimoniae*: Less is More

As it appears from results, Richards model requires, for both the considered distributions, a minimum density of points to justify its shape parameter β, according to the AIC_c metric. Below this threshold, the biological "signal" is treated like noise. It emerges that, in various persistence settings, best-model selection strategy happens in a specific fashion. Indeed, as AR(1) tends to the optimal value of two - usually occurring around half of the elimination steps - the Logistic curve gets chosen more often. Whilst Richards achieves artificially "perfect" fits accounting for unreliable forecasts, sparse datasets are better suited for the simplicity of a less parametrised, Normal distributed, symmetrical sigmoid such as Logistic. This may also explain the reason for, in these not ideal conditions, its measured advantage over Gompertz models and, sometimes, even the log-Normal distributed version of both curves.

We suspect that the described behaviour across points removal's steps depends on the loss of that **biological asymmetry** - which is a fundamental aspect to bear in mind when approaching this field - that often justified the log-Normal superiority over the Normal distribution. Consequently, what is left in residuals is pure white noise, making it impossible for models to distinguish other information, such as the slowing down effect of the IP. We could state that the modelling of such dynamics follows an "*Occam's Razor*" structure, where parsimony is the key element to

consider when selecting a curve as the best-fitting one.

Nevertheless, when a dataset possesses a sufficient number of points, results have shown improvement with the inclusion of a more skewed statistic, and we can draw some qualitative and quantitative conclusions.

Conclusions on the log-Normal distribution

The multiplicative standard deviation σ^*, ranging between 1.05 and 1.3 across the datasets, confirms the consistency of experimental measurements we discussed. Unlike the additive standard deviation of the Normal model, which imposes an unrealistic decrease in relative error as the tumour grows, σ^* characterises a constant proportional uncertainty. This biological realism is reflected in the better AIC_c scores and lower MAPE values of the log-Normal fits, making the choice of a multiplicative error structure for characterising oncological growth dynamics quite interesting. Moreover, across forecasting plots, the IC bands better represent data behaviour; even though the uncertainty is broader, they often include test points with **surprising precision**. This is interesting and requires a systematic study, whose scope is too broad for the present thesis.

Results suggest that log-Normal approach provides a superior description of the growth trajectory across various fit metrics and predicting power, preventing the misconsideration of low-volume data points in the objective function and ensuring more identifiable and biologically grounded parameter estimates, whilst achieving more accurate predictions over test set compared to Normal distribution. It would be valuable to assess how this distribution assists a better understanding of data in other clinical and therapeutic contexts where tumours are treated.

Chapter 4

Three-dimensional ODE Growth Models

MODELS often provide tools for detecting patterns. In chapter 2 and chapter 3 we adopted one dimensional ODE Sigmoidal curves for studying growth dynamic in experimental data. Results have been promising, especially when shifting the focus on detecting possible biases surging from the underlying statistical assumption, allowing to prove the multiplicative effect of such cases. Nevertheless, intrinsic limitations prevent S-shaped models from fitting mixed or cyclical behaviours. That is the main reason motivating a further step in the analysis.

Let us display experiments for which **oscillations** of cancer growth take place; that is the signal of strong therapeutic and immune responses leading to tumour remission, but that are unfortunately not sufficient to completely eradicate the tumour. As a result, new cycles of proliferation and tumour diffusion follow, triggering again an immune system reaction that synergies with an exterior agents. Scientific hope lies in discovering reliable tools that can predict whenever growth might peaks again in order to plan more efficient dose delivery to patients.

4.1 Learning from Ecology and Epidemiology

As a first form of exemplification, we described the tumour-virus dynamic through the total malignant volume evolution, neglecting viral load representation, while also ignoring any distinction between infected and normal cancer cells. This is solid enough for one-dimensional ODE modelling, however scientific literature gives a wiser sight on the topic. *Is the change of mathematical structure and the increased complexity balanced by a generalised solution to both sigmoidal and cyclical dynamics?*

Lotka-Volterra

The mathematical foundation for modelling the interaction between competing populations is rooted in the **Lotka-Volterra** (LV) first-order, non-linear, ODEs. It was

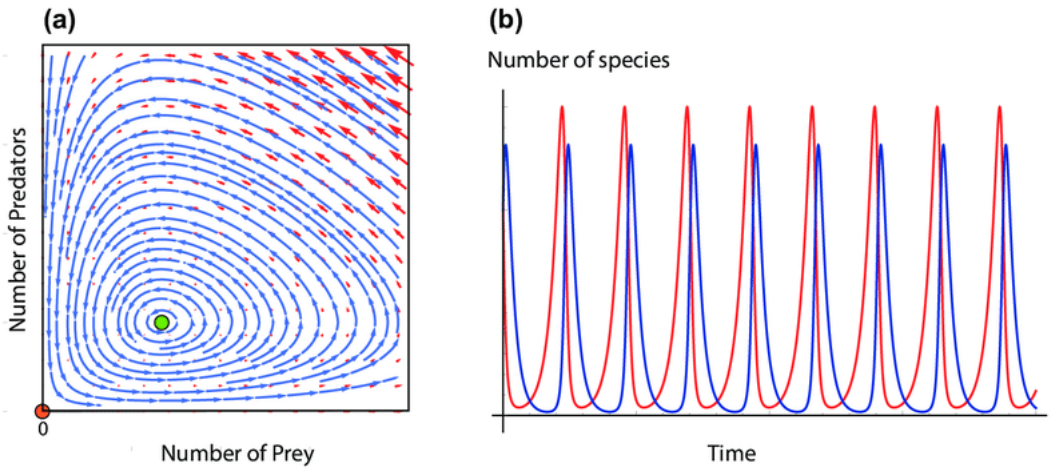


Figure 4.1: *LV*'s phase portrait and solutions. Prey population x in red, whereas the delayed blue curve describes predators count y overtime.

originally developed to describe a *predator-prey* system, and it is expressed as:

$$LV := \begin{cases} \dfrac{dX}{dt} = \alpha X - \beta XY \\ \dfrac{dY}{dt} = \delta XY - \gamma Y \end{cases} \tag{14}$$

and models a **predator** agent Y whose survival depends on the amount available resources, consisting of the **prey** population X. The model comes with some assumptions on environment and biology of the predator and prey populations, such as: X finds abundance food at all times, whereas Y only consumes prey and the rate of change of population is proportional to its size. Moreover, aggressors have limitless appetite and the system is isolated. Clearly, both densities can be described by a single variable.

Hunt for survival and reproduction of the former causes several lost in the latter. Consequently, aggressors' concentration peaks while X is dropping. As a result, competition effect hits Y density, whose individuals die of starvation parallel to prey reaching their minimum. While attacks become less frequent, prey starts to replicate again until reaching a maximum. Such an abundance, stimulates a delayed surge in the predator population, subsequently causing a collapse in the prey count; this behaviour is taking place infinitely many times in the hypothesised isolated system. More specifically, the solution to this system typically exhibits **periodic oscillations**. These solutions form closed orbits in the phase plane, indicating a neutrally stable equilibrium where the system never settles into a steady state but revolves around a centre, as seen in figure 4.1. We remark that the only steady possible configuration is, trivially, the origin $(0, 0)$ for which no species is present.

The parameters define the nature of the dynamic: α is the intrinsic growth rate of the prey, β represents the rate at which predators consume prey, δ describes the rate at which predators proliferate by consuming prey, and γ is the natural death rate of the predator. The rate of predation is determined by the number of encounters between the two groups and follows the *law of mass action*. This gives the **interaction term**, described by their product $X \cdot Y$, where obviously the sign reflects the food-chain direction.

However, this trivial, isolated, system is just a starting point. Obviously, OVs

action against the TME is far more complex; the main remark here is that the virions adopt - or, better said, are designed, engineered and delivered for - a predatory behaviour toward cancer cells. Another key difference is the number of actors itself, since a great effect of OVT consists of inoculated cells' lysis and apoptosis, with the consequent increase in viral load. Moreover, they stimulate an immune response. A first natural step from the *LV* system is the inclusion of a third actor, and related equation. We hereby reference to another famous example, accounting for this addition.

Compartmental models

Hence, we consider a class of models meant to model epidemiology contexts, such as studying the outbreaks of infectious diseases.

SIR Evolving from *LV* equations' structure, a first example is the **Susceptible-Infected-Recovered** (*SIR)* framework where the "predator" is the pathogen and the "prey" is the susceptible healthy population. Transition from general ecology to epidemiology is formalised by shifting the focus to the transmission and clearance of the infection:

$$SIR \coloneqq \begin{cases} \dfrac{dS}{dt} = \beta SI, \\ \dfrac{dI}{dt} = \beta SI - \nu I, \\ \dfrac{dR}{dt} = \nu I, \end{cases} \tag{15}$$

When a *Susceptible* (S) individual and an *Infected* (I) one come into contact, the former contracts the disease from the latter and transitions to the compartment of the latter. *Recovered* (R) individuals, are those who have been infected and have either remitted from the disease, therefore gaining **immunity**, or died. The main assumption is that the sum of the three terms vanishes

$$\frac{d(S+I+R)}{dt} = 0,$$

so that the population $N = S + I + R$ is constant. It is also assumed that the number of deaths is negligible with respect to the total population. During an epidemic, density of susceptible individuals falls rapidly as more of them are infected and thus enter the I and R groups. For the purpose of this research, time-frame does not allow for the disease to break out again because the susceptible population cannot build back up.

A quick overview on parameters: those who carry the disease infect on average β individuals each per day, for an average time of $1/\nu$ days. The ν parameter is called *recovery rate*, while β is the *disease transmission rate* . The average number of infections arising from an infected individual is then given by the *basic reproduction number*

$$R_0 \coloneqq \frac{\beta}{\nu}.$$

The model reduces to a two dimensional dynamic as we consider R a straightforward computation from I. With simple calculations equilibria are given by the S-axis (I=0), for which stable ones are to be find at $S < \frac{\nu}{\beta} = (R_0)^{-1}$, whereas

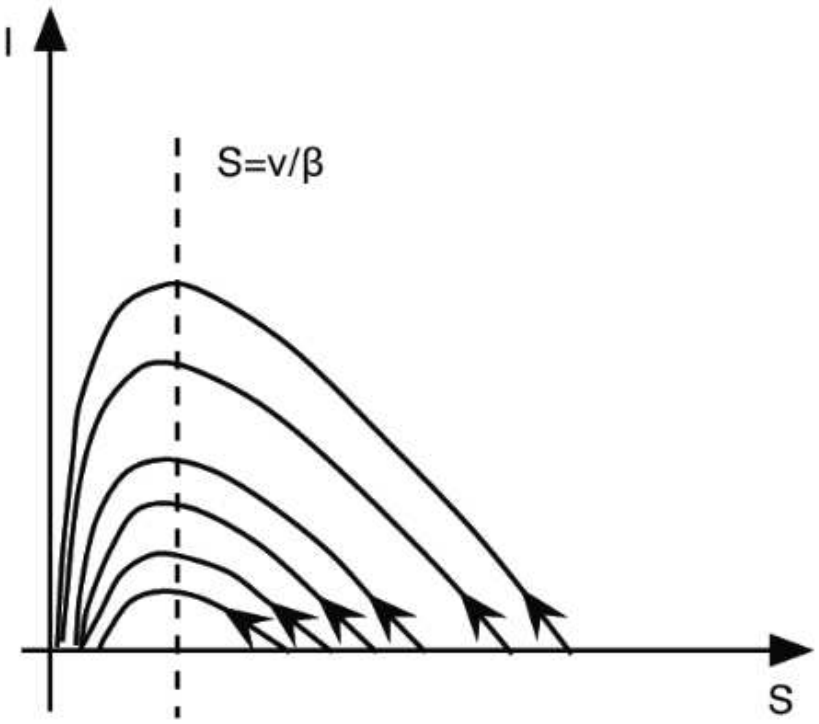


Figure 4.2: *SIR* model's phase portrait

the remaining interval is unstable. Thus, by studying phase portrait in figure 4.2, system yields to a unique solution curve connecting each pair of equilibria. Unlike the persistent, periodic cycles of aforementioned predator-prey system, a basic SIR model typically leads to a temporary outbreak; the infection peaks and then decreases as the susceptible group is left below the critical threshold. In this scenario, the system reaches a stable equilibrium where the disease dies out or remains at a negligible level. Also, $R_0 > 1$ implies the propagation of the infection until all the population dies.

This model is thought to be reasonably predictive for infectious diseases that are transmitted from human to human, and where recovery confers lasting resistances such as Measles and Rubella (Ref. [16]). Interestingly, an **analytical solution** has been found for such a model, please refer to Ref. [18] for more details, although it is implicit.

While the classic *SIR* model provides a fundamental framework for infectious diseases, its standard behaviour differs significantly from the sustained oscillations often observed in biological microscopical systems. A further step towards OVT direction is required, allowing for more various dynamics.

SIRS To better capture the long-term, cyclical behaviour of certain tumours - where periods of remission are followed by regrowth - we can first, focus on a more elaborate epidemiological framework, *SIRS*, by contemplating loss of immunity. Indeed, the added S alludes to the possibility for recovered individuals to fall back into susceptible cohorts, quantified via the parameter μ. With the addition of this term, the system becomes:

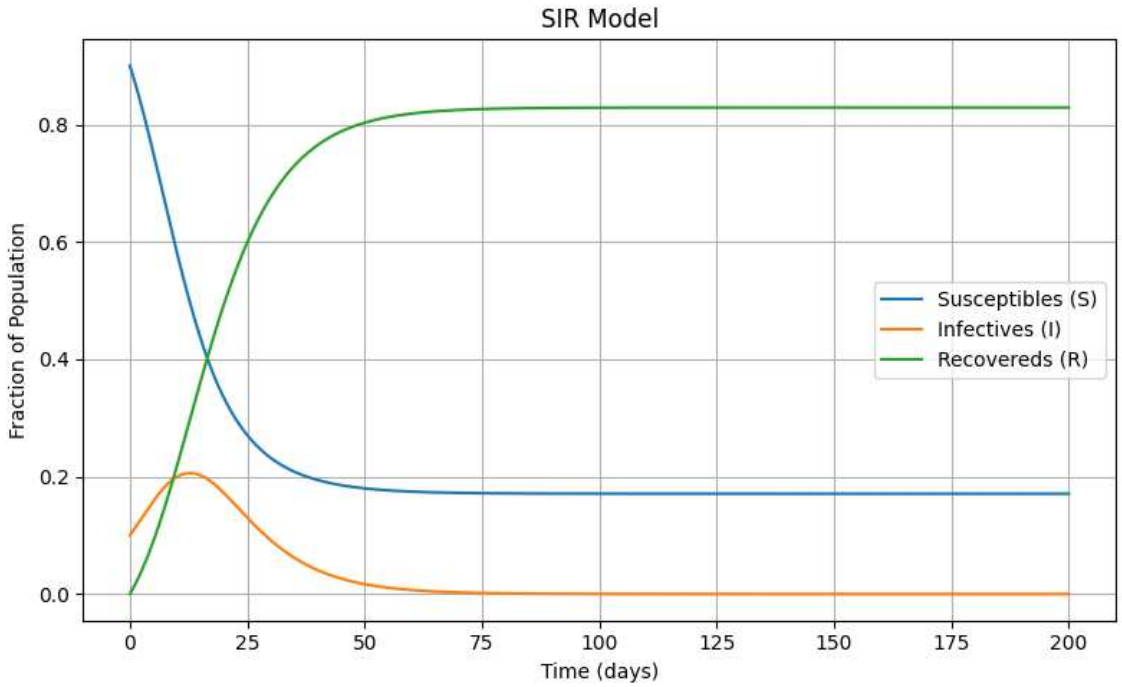
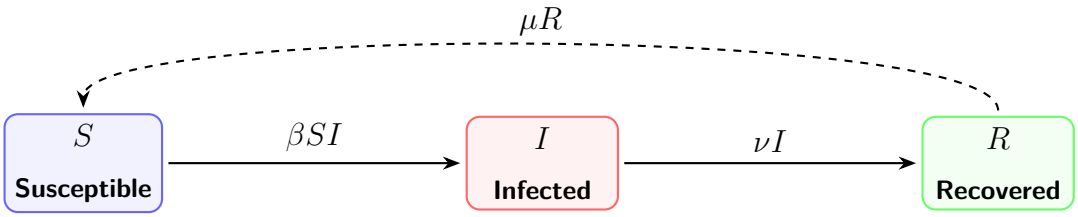


Figure 4.3: Numerical solutions for the *SIR* Model for the following initial conditions and parameters: $S(0) = 0.9,\ I(0) = 0.1,\ \beta = 0.28$ and $\nu = 0.14$. Resulting $R_0 = 2$.

$$SIRS \coloneqq \begin{cases} \dfrac{dS}{dt} = -\beta SI + \mu R, \\ \dfrac{dI}{dt} = \beta SI - \nu I, \\ \dfrac{dR}{dt} = \nu I - \mu R, \end{cases} \tag{16}$$

In this context, the susceptible-infected interaction evolves from a single wave of transmission into a potential "predator-prey" dualism, where the viral predator can persist alongside the prey through continuous cycles of reinfection. This creates a feedback loop where the susceptible population is constantly replenished, enabling the system to bypass the simple transient peak and move toward sustained oscillatory dynamics or a persistent endemic state. The model has the following input-output diagram, with dashed line representing the loss of immunity:



Again, we express the system in two equations form by substituting

$$R = N - S - I$$

so that, unlike the *SIR* model, we now have at most two equilibria: a trivial no disease configuration at $S = N, I = 0$, and a second one at $x^* = (S^*, I^*)$ only present if $N \geq \nu/\beta$. This value, corresponding to the inverse of R_0, is the **threshold level for disease**. A straightforward computation for the disease case shows that both eigenvalues have negative real parts, and so the equilibrium is asymptotically

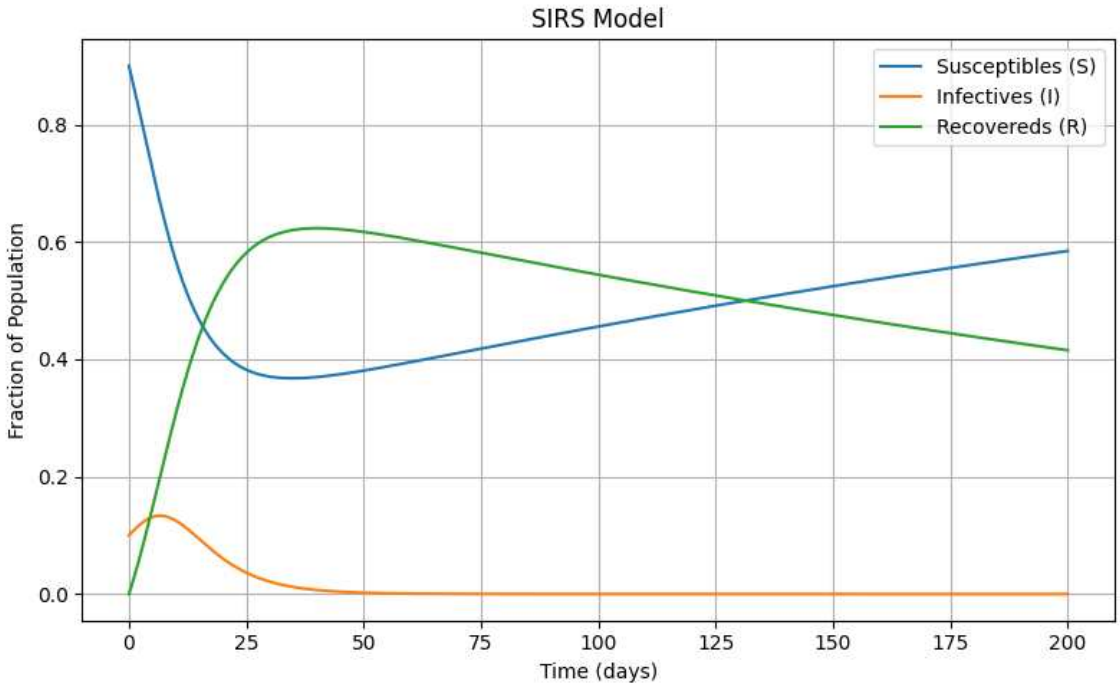


Figure 4.4: Numerical solutions for the *SIRS* model forthe following initial conditions and parameters: $S(0) = 0.9,\ I(0) = 0.1, \beta = 0.375,\ \nu = 0.25$ and $\alpha = 0.0027$. Resulting $R_0 = 1.5$.

stable. Biologically, it means that the epidemic might establishes in the population only for a significant number of total individuals - which, in the initial setting, has to do merely with susceptibility and no congenital immunity.

Let us now make a parallelism for which, within this context, the oncolytic virus acts as the I term that evolves alongside its malignant S "prey", triggering a fight between infection and tumour proliferation. To translate these biological interactions into a quantitative predictive tool, we formalise these transitions through a qualitatively similar system. Therefore, instead of undergoing through detailed analysis of the *SIRS* model, we shift to the OVs setting and present differences and similarities.

4.2 A Minimalist Model for OVT

One could think of this section as more than a mere further step from previous paradigms. Even though models express similar terms and behaviours, the scaling to which dynamic is taking place is sensibly different. For instance, when dealing with an epidemic we consider pathogens and disease development to have predetermined and fixed behaviours and therefore infections follows the same patterns. However, due to the complexity of OVT, several aspects may interfere with malignant cells dynamic. A need emerges for generals yet simple tools for studying interaction between OVs and tumours.

Definition

Literature provides sophisticated models, incorporating a plethora of possible aspects of the in-body response to OVT; however, we are interested to see whether

Table 4.1: Parameters and initial conditions used in OVT models according to literature; please refer to Ref. [19] for a detailed overview.

Parameter estimates		
Symbol	**Definition**	**Value**
r	Tumour growth rate	0.3/day
β	Infectious rate	$\approx 1.5 \times 10^{-9}$mm^3/(viruses×day)
d_I	Infected cells death rate	1/day
d_V	Virus death rate	4/day
α	Viral burst size	3500 viruses/cell
u_0	Initial density of uninfected cells	10^9 cells/mm^3
i_0	Initial density of infected cells	10^7 cells/mm^3
v_0	Initial density of virus particles	1.9×10^{10} viruses/mm^3

a minimalist set of ODEs is sufficient for a qualitatively understanding of such dynamic. We target at describing a three actors regime: free *Virus* particles (v) attack the cancer tissue, while malignancy is hereby sub-classified either as *Uninfected* (u) or *Infected* (i) tumour cells. Each one is a function of time τ. This shift allows for better dealing with the OV's action as a comprehensive measure that is, at least ideally, perpetually available for clinicians to be restored. The reasoning behind this setting is that, due to the cyclical waves of tumour replications, patients should be able to manifest stronger and not compromised immune responses by being provided with safe but consistent viral loads. The system is

$$\begin{cases} \dfrac{du}{d\tau} = ru - \beta uv, \\ \dfrac{di}{d\tau} = \beta uv - d_I i, \\ \dfrac{dv}{d\tau} = d_V v - \alpha d_I i, \end{cases} \tag{17}$$

where τ represents the number of days. We assume that uninfected tumour cells replicate by a factor r, proportional to their amount. As we have previously explored, exponential growth is not a feasible system description because of the competition effect. However, due to the short exposure between TME and OVs, it is suitable for this situation. There is, in particular, a term in common with *SIR(S)* system and that is the interaction term deducted via the same law of mass action, occurring again with constant rate β. Parameters are presented in table 4.1.

A fundamental assumption in this basic framework is that once inoculated, host cells are incapable of division, as virions takes control. Finally, the lysis induced burst at rate d_I cause the release of α new viral particles. Moreover, tumour cell death unrelated to treatment are neglected. A scheme for this interaction can be found in figure 4.5.

Through an appropriate change of variables, that is described in Ref. [20], equa-

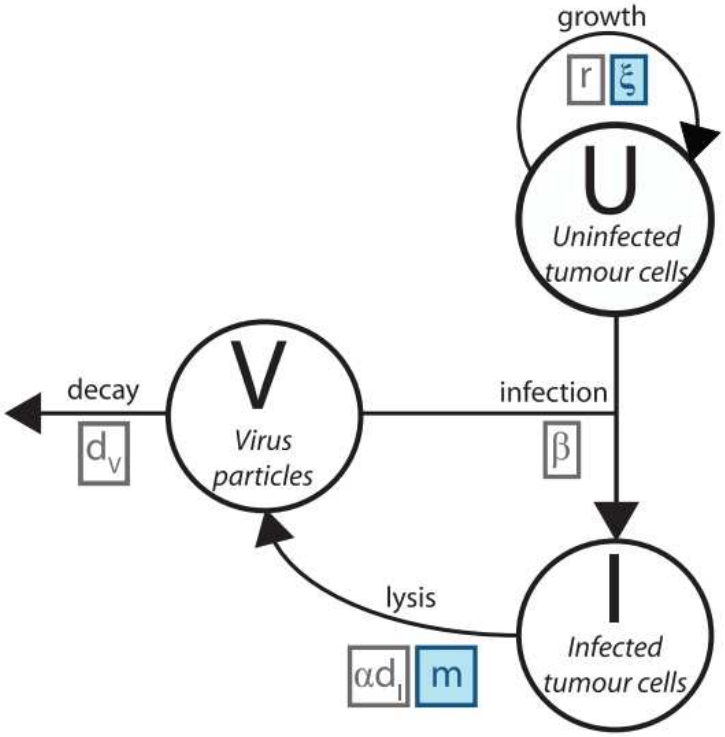


Figure 4.5: Flow diagram from the research article [20] that sums up interaction between described populations. Parameters listed relate to the original model in grey boxes and to the non-dimensionalised form of the model in blue boxes.

tions can be simplified into their *dimensionless form*:

$$\begin{cases} \dfrac{dU}{dt} = \xi U - UV, \\ \dfrac{dI}{dt} = UV - I, \\ \dfrac{dV}{dt} = -mV + I. \end{cases} \tag{18}$$

Explicitly, $\xi = r/d_I,\ m = d_V/d_I$. The parameter ξ represents the **dimensionless growth rate** of the uninfected tumour population, typically ranging from 0.0001 to 1.0 to describe various levels of cellular aggressiveness, whereas m defines the **viral clearance rate**, with values between 0.1 and 1.0 that facilitate oscillatory limit cycles and values above 3.0 correspond to rapid viral depletion. Finally, U, I, V and t represent the corresponding populations and time.

This formulation preserves the properties of the original model, while carrying the advantage of dealing with only two parameters, which is less computationally demanding. We remark that limitations that could potentially inhibit OVT efficacy are simply ignored here. The model dynamics rely heavily on initial conditions and choosing realistic values for each parameter and population density is crucial for obtaining meaningful fits. Since we considered an exponential growth term, if the ratio U_0/V_0 is large enough, the tumour will outperform over the viral action, producing an unbounded growth for the uninfected population.

Table 4.2: Parameters and initial conditions used in the non-dimensionalised system in 18.

Parameter estimates			
Symbol	Name	Definition	Value
ξ	Dimensionless growth rate	r/d_I	0.3
m	Viral clearance rate	d_V/d_I	0.25
U_0	Initial proportion of uninfected cells	$U(0)$	0.4
I_0	Initial proportion of infected cells	$I(0)$	0.01
V_0	Initial proportion of virus particles	$V(0)$	0.1

Stability and long term behaviour

We refer to results from Ref. [20], where computations are presented in details. The first equilibrium is located at the origin

$$U = 0, \quad I = 0, \quad V = 0, \tag{19}$$

while a non-trivial one is to be found at $x^* = (U^*, I^*, V^*)$ such that:

$$U^* = m, \quad I^* = m\xi, \quad V^* = \xi. \tag{20}$$

This is an equilibrium where the tumour coexists with the virus. Values of $m = 0$ and $\xi = 0$ yield to one more equilibrium each; the former implies that $d_V = 0$ (no viral decay is occurring by the immune system), resulting in $U = 0, I = 0, V \in \mathbb{R}$. The latter instead requires $r = 0$ (no measured tumour replication) and the consequent equilibrium locates at $I = 0, V = 0, U \in \mathbb{R}$. These are all trivial.

At the origin, by varying ξ and m, eigenvalues lead to the non-trivial equilibrium being a saddle node, except for $m \in (0, 2) \wedge \xi \in (-2, 0)$ where they all take a negative value determining a stable node. At x^* the stability is reported by figure 4.6. Note that negative parameters do not have biological significance

By studying stability of the system we know that, in absence of tumour growth and viral decay, model predicts that OVT is able to completely eradicate cancer cell population; however, in more realistic proliferation settings, no stable equilibria is present. Nevertheless, oscillations are present in this case's dynamic. OVs can reduce a remitting tumour to a stable oscillatory state, without achieving complete eradication. To date, the possibility for a virus to be concealed from the immune system and thus not sustaining any significant decay, remains just an hypothesis. In contrast, a deep insight on **amplitude** and **period** of oscillations resulted that the most effective strategy for improving OVT is lowering both the ratio of viral particles decay to cell burst rate, i.e. m and the ratio of tumour replication rate to cell burst rate, i.e. ξ. The consequent long-term evolution is interested by a more delayed tumour cycles and smaller uninfected cells population. This accounts for either tumour remission or, perhaps, a chance for complete eradication if assessed by the inclusion of an further therapeutic element.

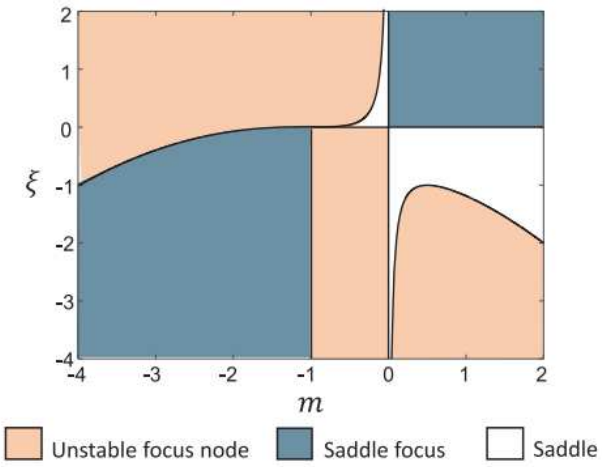


Figure 4.6: Stability regions for varying m, ξ values at the x^* equilibrium. Please refer to Ref. [20] for in depth stability analysis.

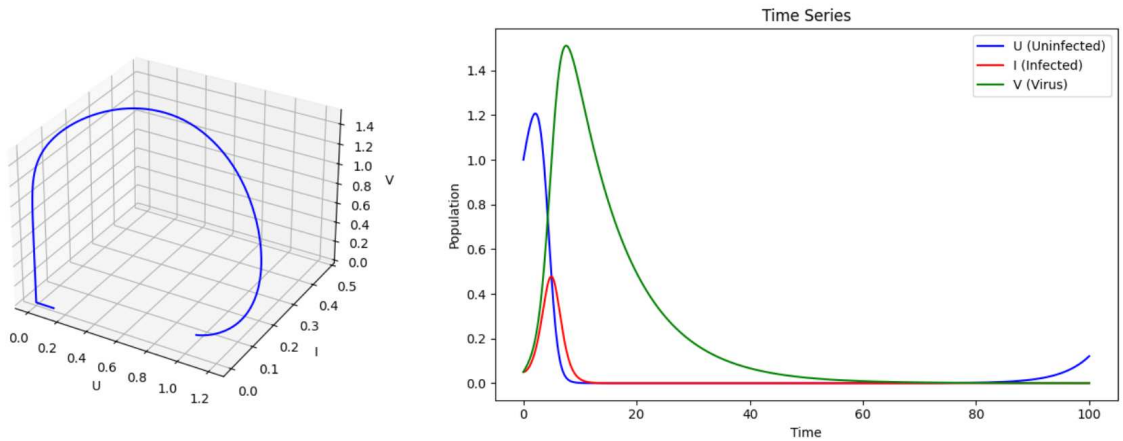


Figure 4.7: 3D plot and numerical solutions time series of the dimensionless model for: $U_0 = 0.95, I_0 = 0, V_0 = 0.05$ and $\xi = 0.2, m = 0.1$.

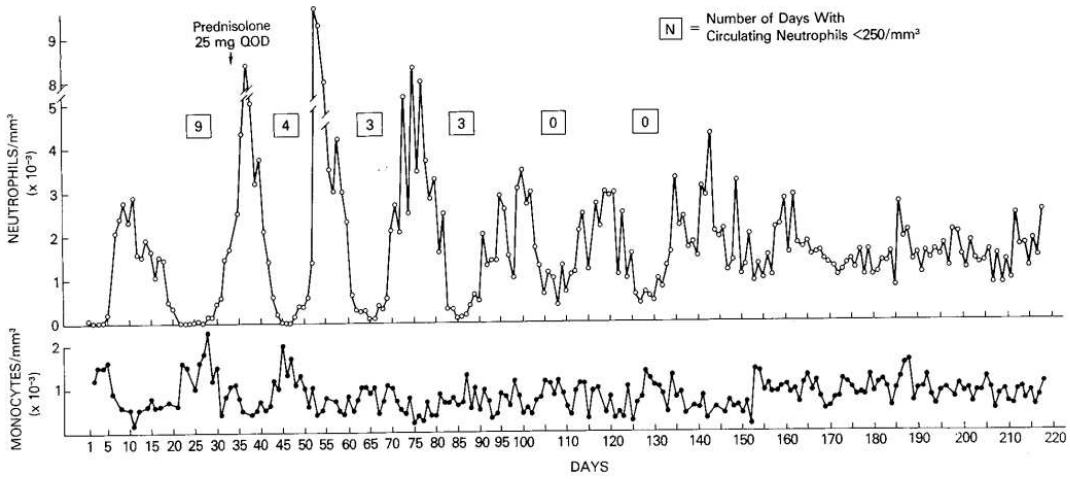


Figure 4.8: Daily blood neutrophil and monocyte counts (main WBC components) during treatment of a patient with CyN. Data are taken from Ref. [21]. Note that these are measurements on actual patients, not mice or guinea pigs.

4.3 Numerical Solution and Fitting

Cyclic Neutropenia (CyN) is a rare haematologic disorder that tends to occur approximately every three weeks and lasting for few days at a time. It causes a temporary condition with a low absolute neutrophil count. Since the neutrophils make up the majority of circulating **White Blood Cell** (WBC), it places the body at severe risk of inflammation and infection. A review of such conditions is presented in Ref. [21] and figure 4.8 shows a patient's immune profile during a therapy. It is worth remarking that leukemias and a number of blood tumours have similar dynamics and the effect of some of the therapies for CyN is very similar to cases of oncolytic virotherapy in blood tumour. The mathematical advantage of CyN is that it offers time-series that are very long and accurate. We will use CyN as a proxy for OVT for blood tumour.

Data

To rigorously test the oscillatory potential of our ODE model, we utilised datasets from CyN conditions as the interest dynamic: as the WBC count drops, we would expect the Immune System to be unable to react to tumour proliferation. The following rise from the lower peak would bring back an anti-cancer potential, that would force a tumour remission. This mechanism is hereby triggering a dual growth phase, opposite to the WBC oscillation frequency. A time-series showing this cycle can be found in figure 4.8. Interestingly, the number of days for which WBC is low, but non-zero, explains the long period orbits behaviour, rescaled to a few days interval.

While the biological drivers differ, the mathematical structures of WBC oscillations and the hypothesised tumour-virus interactions share a common feature: they both represent feedback-regulated systems capable of oscillations occurring with a relatively stable period. By fitting our model to these high-frequency datasets (e.g., 4-5 full cycles over 300 days), we demonstrate the model's ability to capture complex periodic behaviours that simpler sigmoids are not able to achieve.

We remark that we aim at capturing the oscillatory dynamics over the absolute

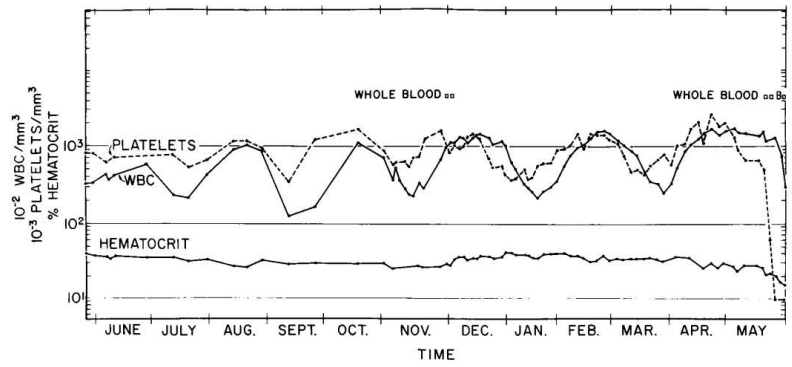


Figure 4.9: Datasets from Ref. [22]. Note the length and the spacing between some measurements.

magnitude. Indeed, efforts in a quantitative "exact" fit means nothing but a chance for over-fitting, where the model usually tends to average between cycles by producing a line. Hence, we now consider as a metric the number of full cycles that the model might be able to fit and predict.

Numerical Optimisation and Parameter Estimation Framework

The transition from a theoretical biological model to a data-driven predictive tool requires rigorous estimation of unknown parameters. In our context, the interaction between tumour cells and oncolytic viruses is governed by the tumour growth rate ξ and the viral clearance m, which must be inferred from experimental observations.

A significant challenge in fitting nonlinear ODEs is the high sensitivity to initial conditions (U_0, I_0, V_0). To ensure the model captures the correct dynamical regime, we implement a specialised two-stage optimisation process. First, a **Grid Search** is performed over a range of the biological rates ξ and m to identify a *good* guess to start with. This is followed by the optimisation process that includes those rates and also the initial conditions. This approach allows the model to adjust the starting state, which is fundamental to get meaningful trajectories. Since experimental measurements report the total mass of the tumour, the model is calibrated against the composite variable $U + I$. This step allows to "internalise" the predation term. We then plot the evolution of U, I, and V, while adjusting data to the sum variable as the total tumour volume.

To rigorously evaluate the model, data is partitioned into **training** and **testing** subsets. This split approach ensures that the model is not merely replicating known points but is instead identifying the underlying biological rhythm. By doing so we aim at performing qualitative forecasting over oscillatory information. Metrics measuring adherence to data would appear unreliable at this stage, but by focusing on frequency and amplitude we can consider it an engaging premise for next analysis.

Model Fit

In the optimisation of non-linear ODE systems, the choice of the objective function significantly dictates the qualitative behaviour of the fitted model. While the standard SSR is computationally efficient, it often fails to capture the oscillatory nature

of biological data, such as those observed in cyclical haematological disorders. Our intent is to prioritise dynamical consistency over point-wise accuracy, hence we focus on cyclical characteristics such as amplitude and frequency.

We now consider data from Ref. [22], which we report in figure 4.9. *WBC* dataset presents an irregular density of points: for the third cycle the frequency of measurements is irregular. This aspect may results in odd fitting patterns in certain situations where the train set separates clusters of data. However, intervals of high density, yet coherent, data appear to be of advantage for the optimiser. Examples of fitting for various combination of training set are presented in figure 4.10.

Further experimenting with interval selection lead us to believe that the term *forecast* itself is improper: the ODEs adimensionalised system here is not just achieving precise fits and consequent predictions, but rather capturing the whole dynamic. In fact, by looking at cycles preceding the fit, it appears that frequency and amplitude are coherently described. This is particularly interesting since it would attribute a deep meaning to a short window of data. We could express this effect as if the model were to **sync** with the oscillations frequency, instead of simulating them. This comes at the expense of a risk for "asynchronism" with the dynamic. It is the case for some fit attempts to fail noticeably; the good news is that by looking at such an explicit oscillatory dataset, it is rather obvious when the model is not performing *well*.

We are now interested to see to which extent this synchronisation property holds, so we reference to figure 4.11. Our guess is that every time the model is correctly capturing a cycle worth of frequency and amplitude, it is most likely adapted to a regular experimental oscillatory effect. It seems that this system is able to "predict the past".

Moreover, in some cases, the model is strongly adherent to the dynamic with just few points. The key element here is probably the regularity to which the experimental data occurs. Figure 4.12 offers a clear overview on the model's potential. However, clinical data are not always as ideal as this dataset; we present other examples to show model's limitation.

We remark the main, counterintuitive, fitting accuracy limitation: it appears that forcing the model to spread its optimisations resources over a large interval causes a loss for specific region's reliability. It is the case for the values after the *Fit End* vertical line. One would expect that the model performs better if trained for longer but, as we previously mentioned and experienced with sigmoids, a better fit not always translate to a more accurate forecasting. We hence focus on different interval elections to estimate which periods of oscillations result in better future predictions. For example, cases of figure 4.13.

Other examples

Cyclical Neutropenia We extracted data from Figure 4.14 (Ref. [23]). A first remark is about the model ability to fit cycles. Whilst there is no phenomenological difference between the real and the simulated data, the scale of data has several order of magnitudes of distance. However, since the model is dimensionless by definition, we operate a scaling factor to data, therefore obtaining a precise fit over 4 complete cycles in both datasets. We point out that, due to the naive data generation technique, one dataset's maximum is just the other's minimum. We remark that this result is already an impressive demonstration of fitting and adaptability potential for a two parameters, three dimensional model. Curves solutions are plotted in figure

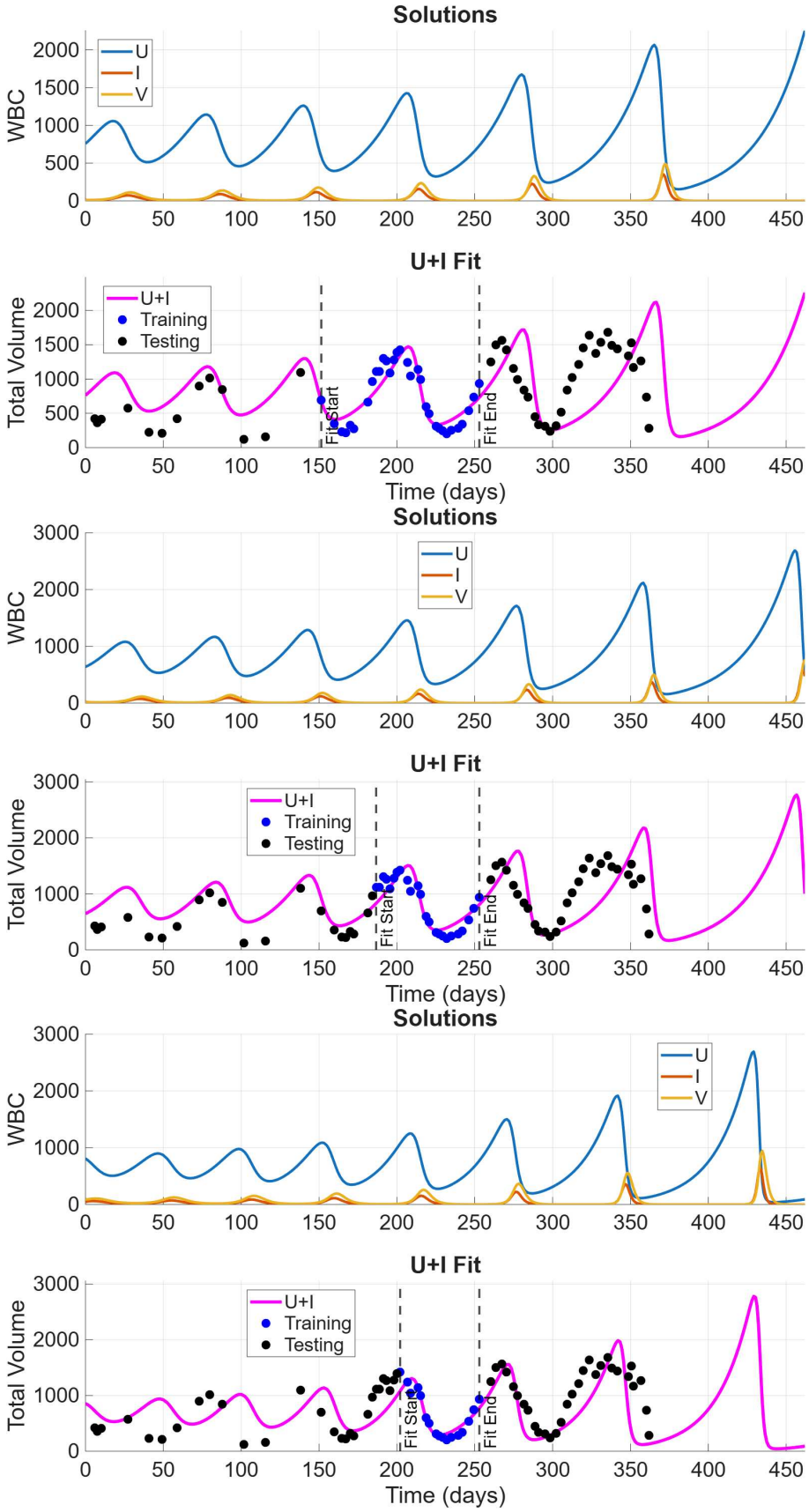
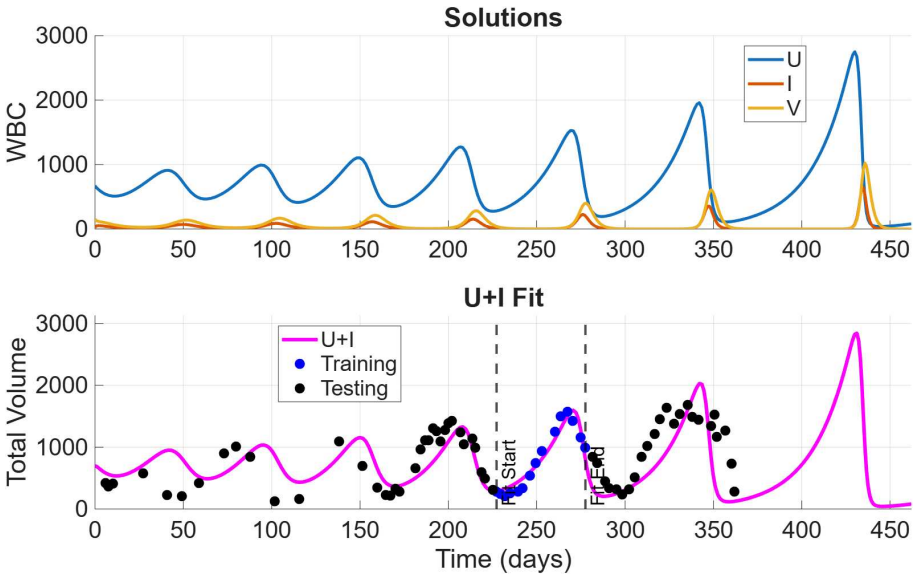
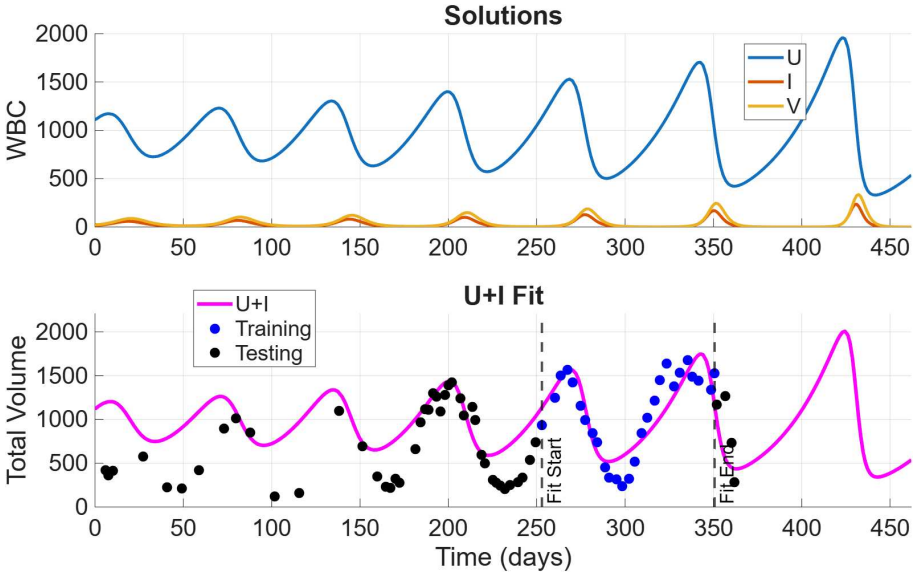


Figure 4.10: Data from figure 4.9. We observe an increase of fit and forecast accuracy by reducing past values. It is hereby remarkable that less than a cycle is enough to forecast the frequency and amplitude of the following two.



(a) $\xi = 0.0471,\ m = 0.4960.$



(b) $\xi = 0.0284,\ m = 0.6536.$

Figure 4.11: Data from figure 4.9. We performed fit on two late intervals.

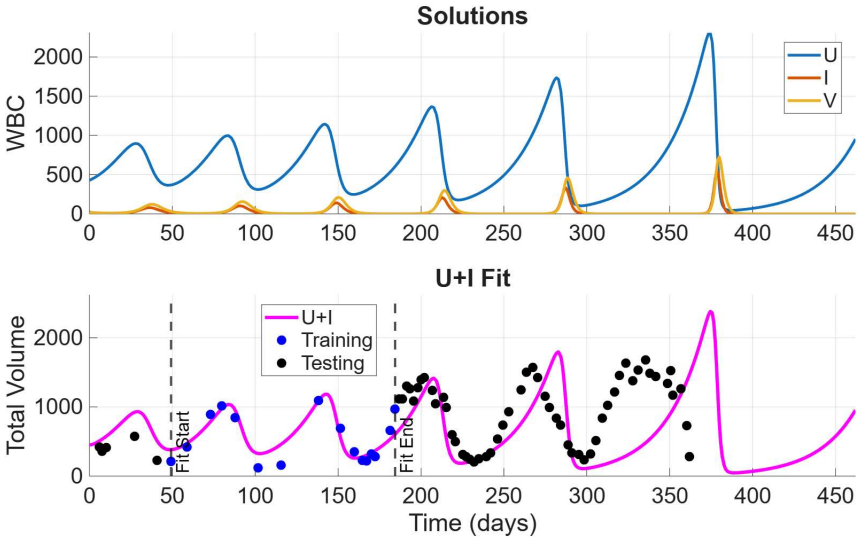


Figure 4.12: Data from figure 4.9, with a different training and test set choice. $\xi = 0.0422$, $m = 0.6203$.

Table 4.3: Parameters for figure 4.15.

Dataset	**Type**	ξ	m	U_0	I_0	V_0
Tumour	Grid search guess	0.0474	0.4526	–	–	–
	Parameters	0.0352	0.5713	0.473	0.001	0.055
WBC	Grid search guess	0.0369	0.4526	–	–	–
	Parameters	0.0344	0.5700	0.670	0.016	0.030

4.15.

Moreover, data present a form of varying amplitude and median values: the periodic frame is shifting over time in an asymmetrical and growing fashion. Thus model is not always able to follow the increasing (or decreasing) trend in the same way. It seems that data here follows a sort of hysteresis pattern: oscillations revolves around a "support" axis that is not horizontal, causing each cycle to affect the next. This effect yields to a data memory; even though amplitude remains to similar order of magnitudes. Since the ODEs system considers only short to long term orbits, it cannot exhibit the same trend of consecutive increase in baseline values.

If we admit to this limitation, it is still remarkable for such a minimalist system to perform such good fits and forecasting. Moreover, the aforementioned sync property holds here. Thus, fitting accuracy is highly sensible to the interval selection. By trying several combinations, it does not appear to have a trend of reliable fitting.

Longer orbits We consider Ref. [21] data plotted in figure 4.17. These cohorts manifest a longer steady period where values are close to zero. However, due to biological differences between the observed phenomenon and the tumour-virus interaction, in terms of rate at which tumour remission may take place.

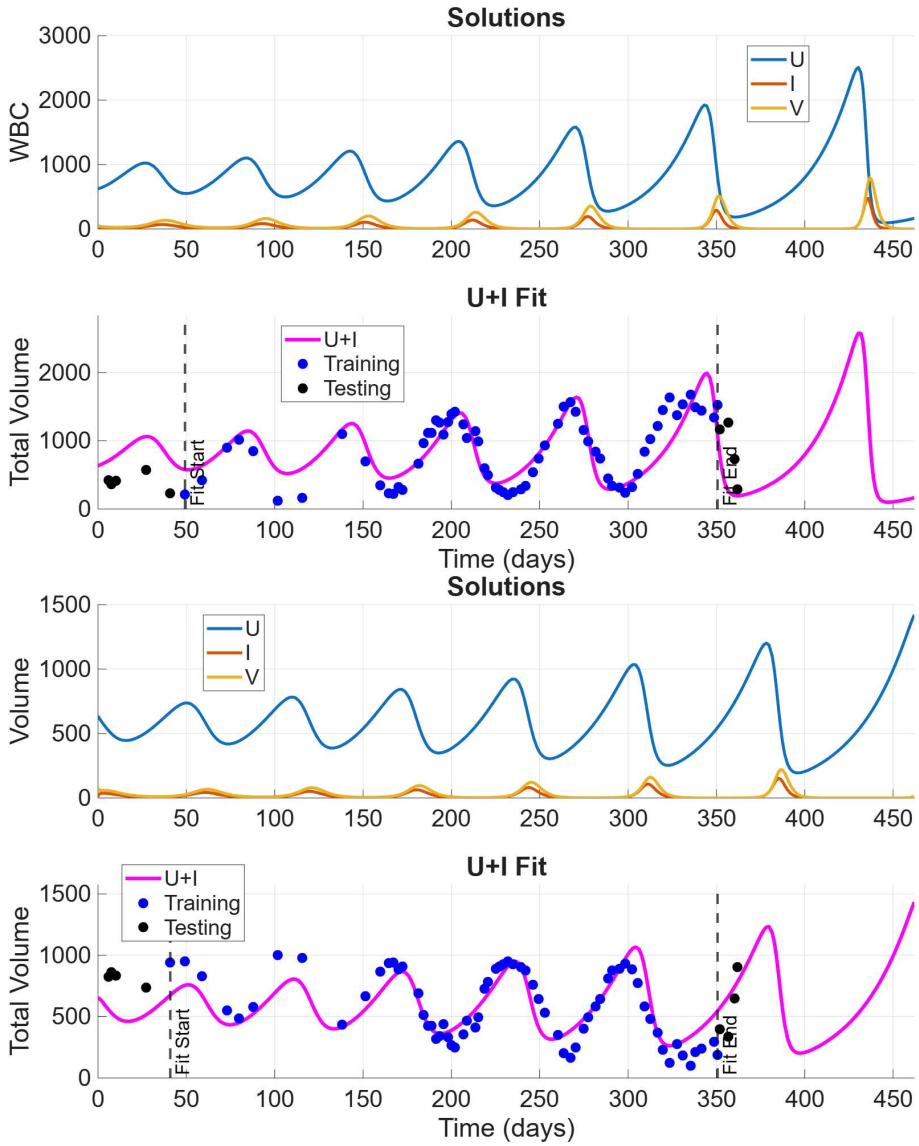


Figure 4.13: Data from figure 4.9 and synthetic generated ones. We performed fit on an interval covering 85% of the dataset.

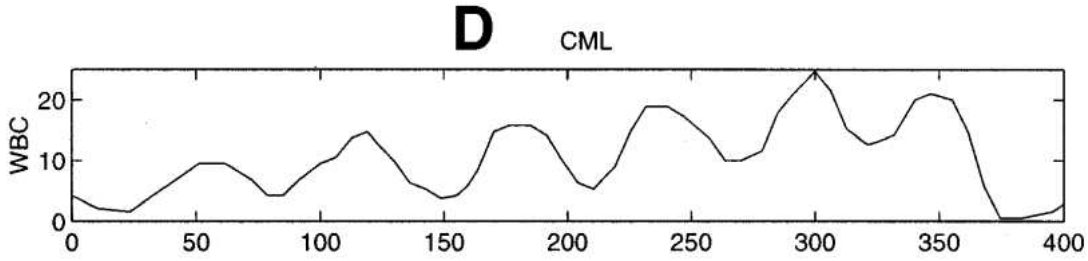
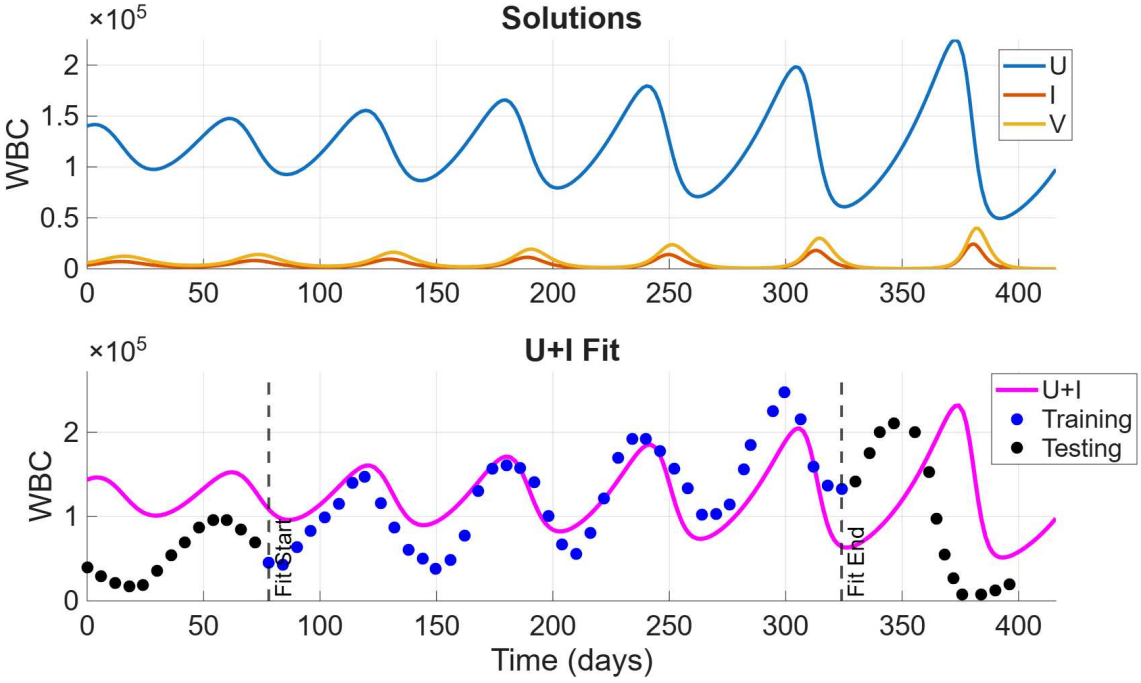
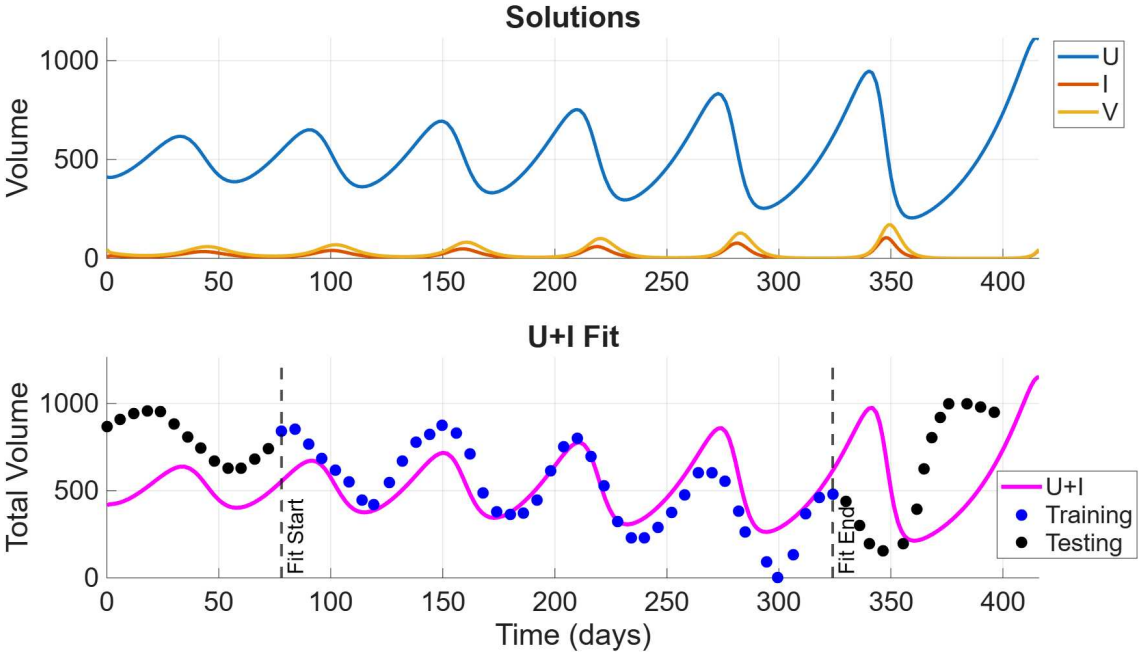


Figure 4.14: Representative patterns of WBC levels in a periodic haematological disorder. Scale corresponds to 10^4 cells; Ref. [23].



(a) Experimental CyN data from figure 4.14 WBC dataset.



(b) Synthetic data, simulated by reversing the frequency of (a).

Figure 4.15: ODEs system 18 fit to data, devised from figure 4.14. Parameters are reported in Table 4.3.

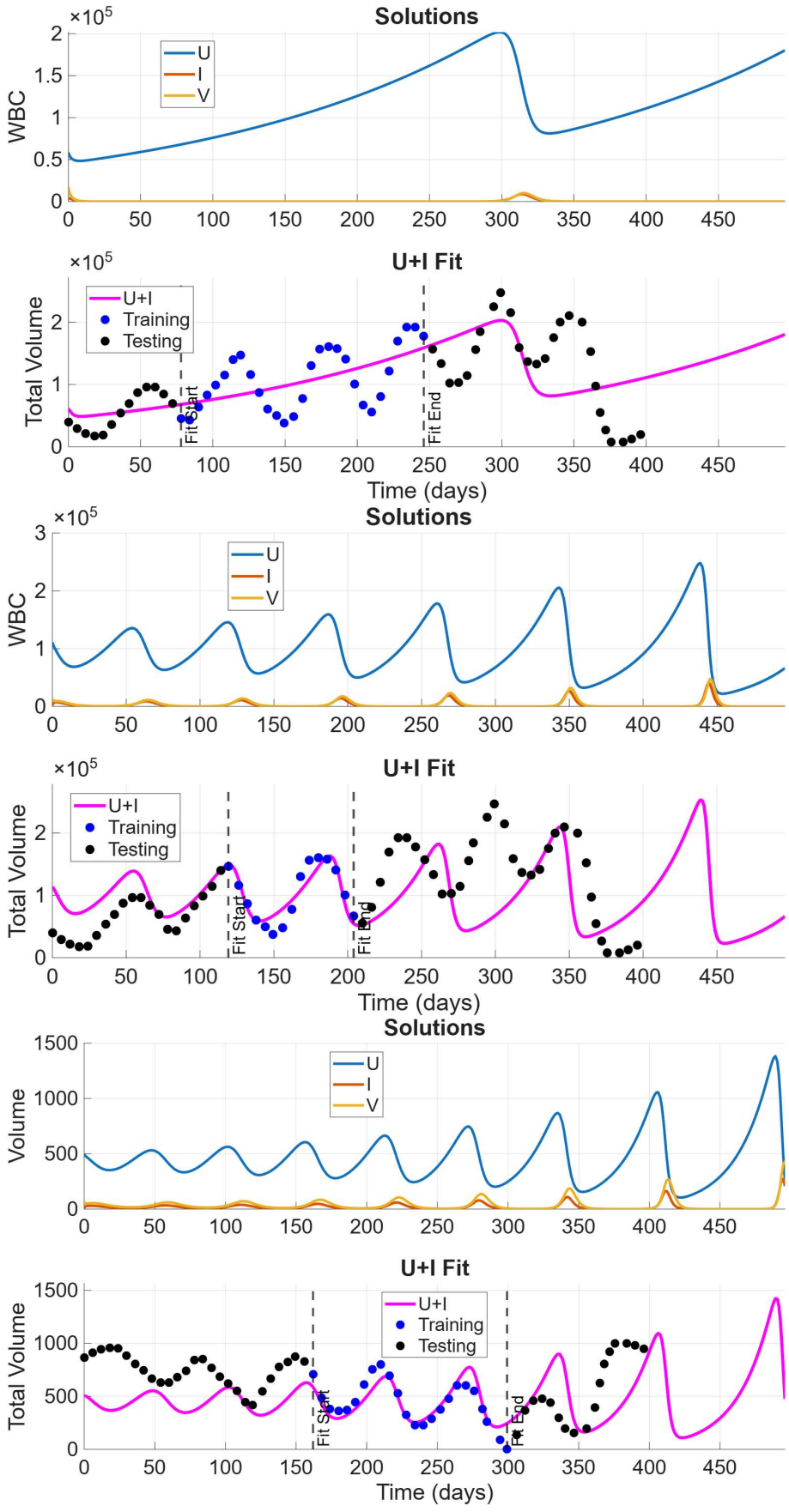


Figure 4.16: Data devised from figure 4.14. Several fits performance over varying interval selections, showing the model sensitivity to virus-cancer dynamics.

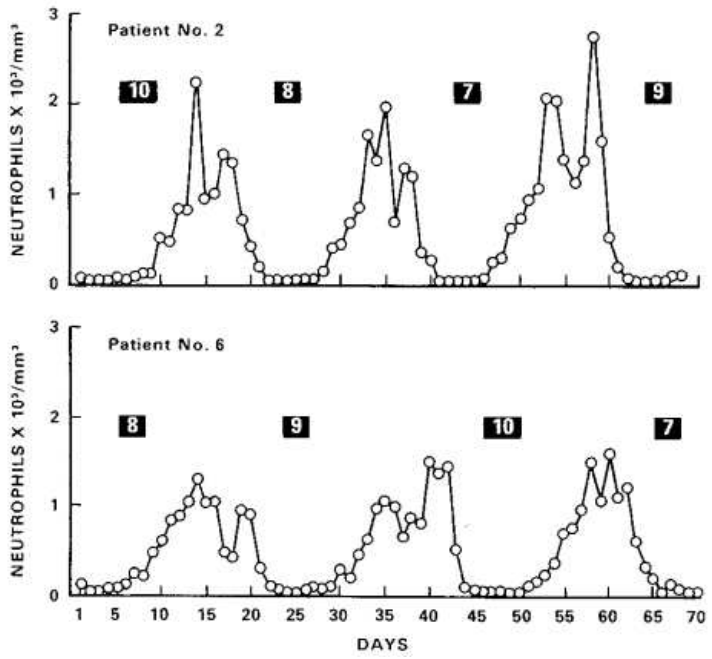


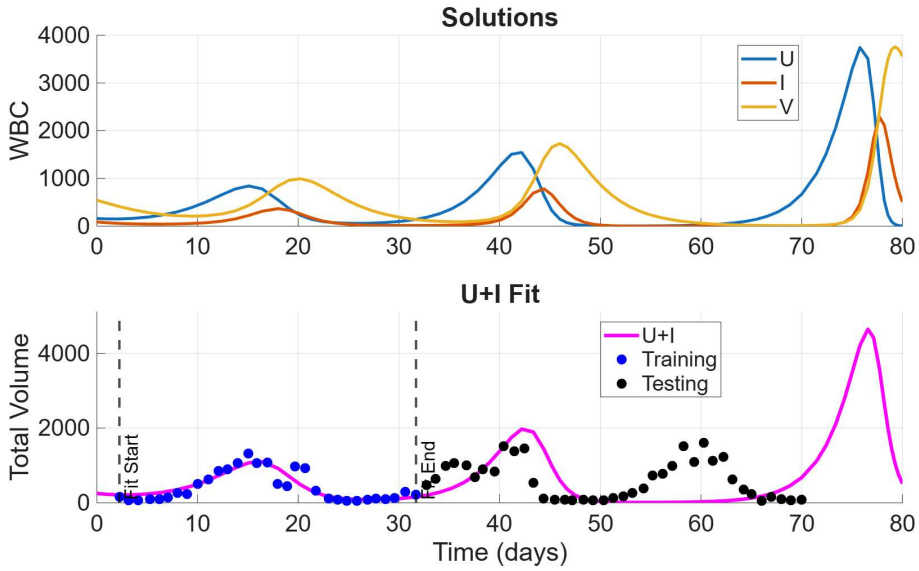
Figure 4.17: Data from Ref. [21] showing long time periods of extremely low neutrophils count in a regular periodic fashion. Note that these measured oscillations occur in human patients - No.2 showing shorter orbits than No.8.

Minimal Data requirements

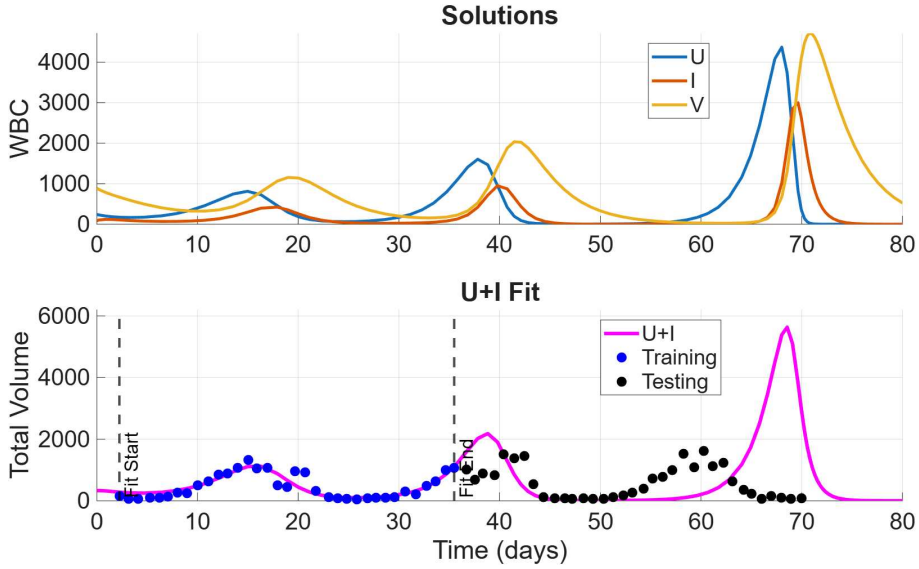
Through our analysis, we noted that without reaching a threshold of minimum number of points, it is impossible to fit any cycle. Take for instance data from Ref. [24], reported in figure 4.22. The *WBC* dataset here is composed of roughly just one point for each peak; the algorithm treats it as a random wave. The resulting fit, for various train set choice, shows similar profiles, such as the one of figure 4.23. We suspect that a focused data collection accounting for frequent and periodic evaluations would allow the optimisation process to adapt. The same pattern repeats for figure 4.8 dataset, for which the narrow and "edgy" cycles make it impossible for the optimiser to fit the curve properly through the points.

Adaptation to sigmoids

In order to prove the system 18 versatility, we arranged the grid search to include parameters ranges that admits growth. The system of ODEs is able to achieve a versatile steepness, even though no wavy neither oscillatory effect is contemplated in such fits. While the main growth trend gets captured quite easily, even by approximating local exponential patterns, it is only to later manifest a sudden drop due to the guessed effect of the OVT. However, not withstanding this aspect for the model, all this analysis points to encouraging results and may end up with a new mathematical tool for forecasting. From a macroscopic point of view, the model has the ability to fit the slope of tumour growth and tends to provide a forecasting horizon that, at least in the first days, is relatively close to the test sample, such as cases of figure 4.24.

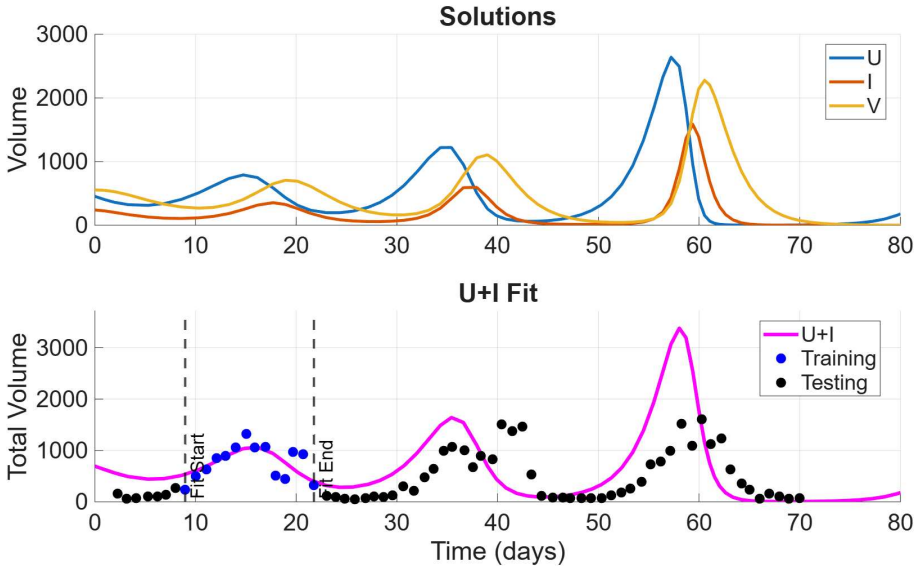


(a) Model here is able to successfully fit the train interval and achieves a delayed forecast for the next cycle, while failing at the following one.

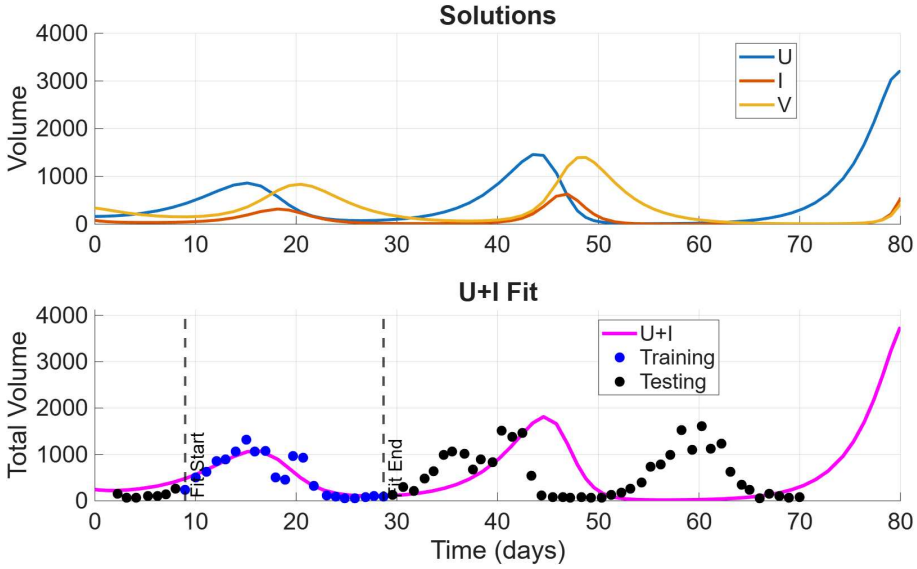


(b) A later *fit_end* value yields to achieving a good fit for the first oscillation occurring in test interval.

Figure 4.18: Data from figure 4.17, patient No.8.



(a) $\xi = 0.3867,\ m = 0.4364.$



(b) $\xi = 0.0019,\ m = 0.7953.$

Figure 4.19: Data from figure 4.17, patient No.8. Note how different solutions to the ODE can approximate well (or not so well) the clinical data; in some circumstances, the fit is highly sensible to minor changes in data selection.

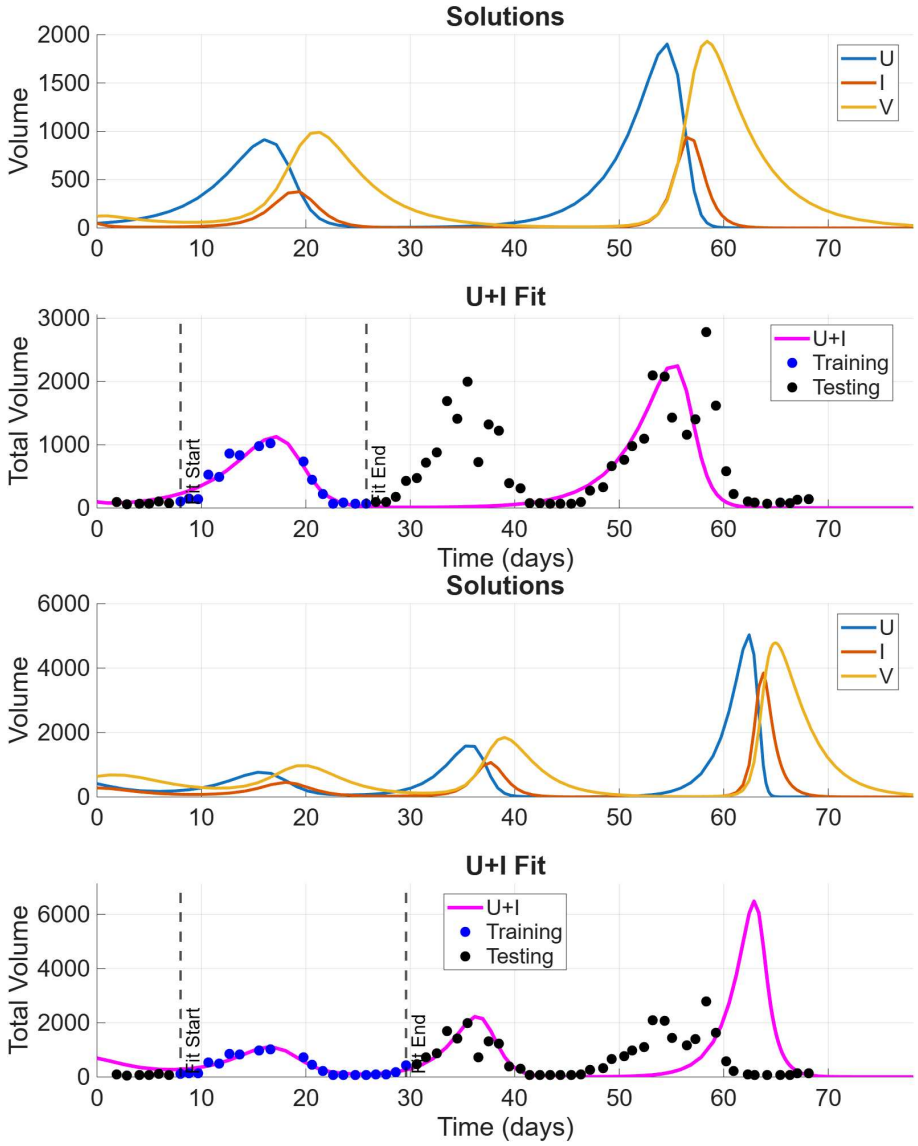


Figure 4.20: Data from figure 4.17, patient No.2. Note how accurate training can be a good indication of future prediction, but not always.

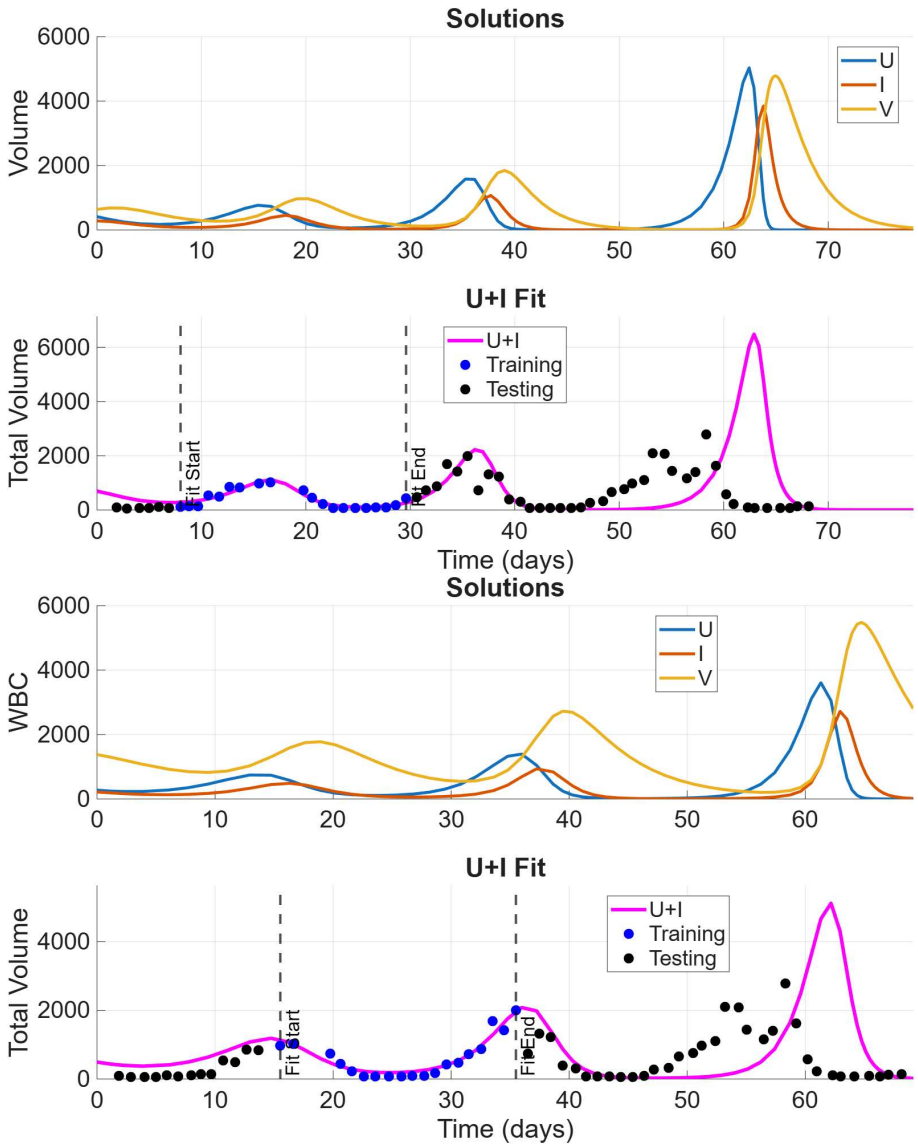


Figure 4.21: Data from figure 4.17, patient No.2 It appears that there is no clue on whether it is more efficient to study an interval in between two minima or two maxima.

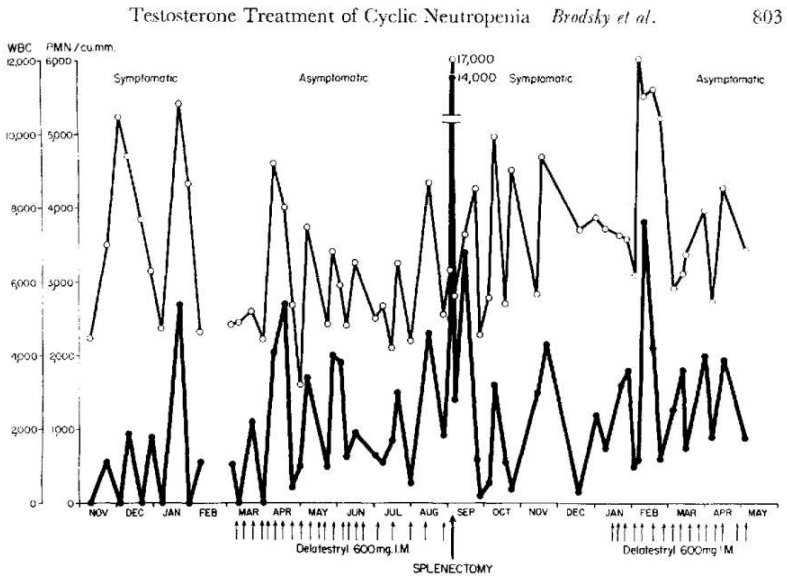


Figure 4.22: Data taken from Ref. [24], from a real patient undergoing testosterone injections. Arrows indicate when such injections have been administered.

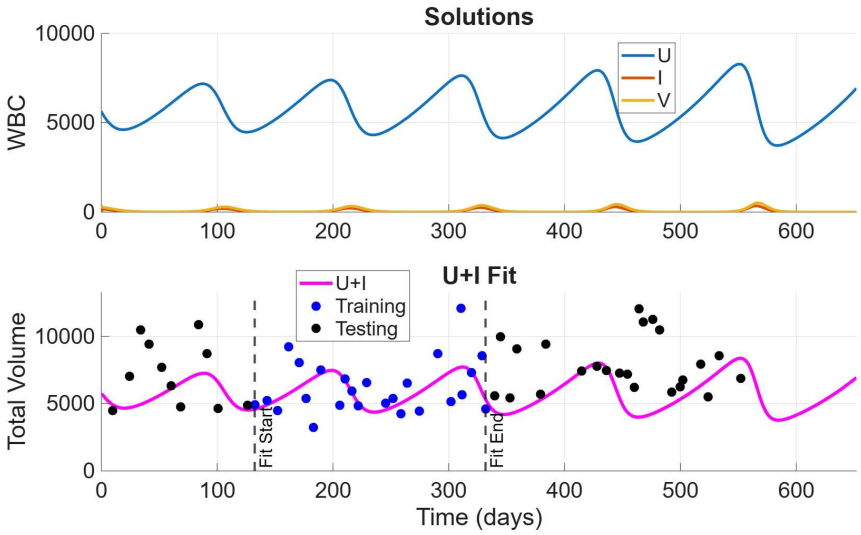
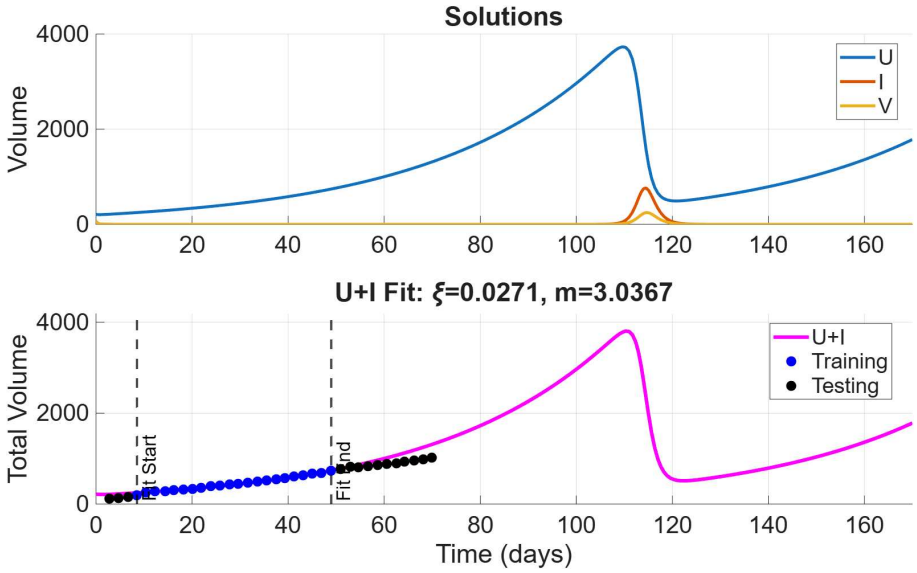
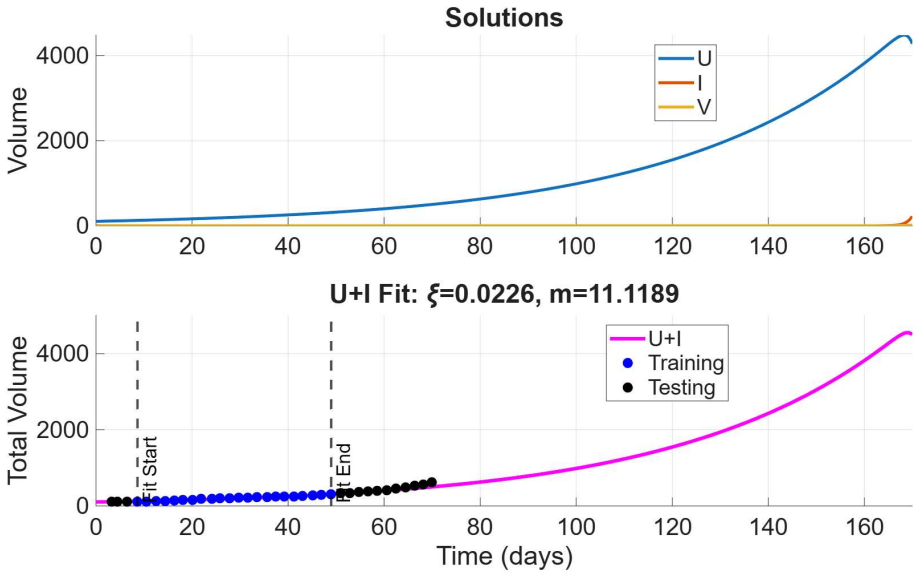


Figure 4.23: The system here is not able to distinguish a clear trend in data from figure 4.22, even though several oscillations occur. An emerging necessity for dense dataset rises here.



(a) *PBS* dataset showing a forecasting trend of underestimation.



(b) V dataset showing a forecasting trend of overestimation. We remark that for the first 15 days of prediction, residuals are less than 50.

Figure 4.24: Data from figure 1.4 fitted with the ODEs system. We remark that the optimiser here forces the model to delay its oscillatory nature.

4.4 Conclusions

The core finding of this analysis is that the ODE model captures the system's intrinsic frequency. The model's ability to replicate the rhythmic nature of the data confirms the ODE's structural validity, proving that the underlying biological mechanism is solid, regardless of point-wise residual errors. In fact, measuring the physical distance of fits (or predictions) from data would lead at missing the qualitative result achieved. As mentioned, more complicated models exists in the literature, but a strong meaning is present in such a minimalist sight on the field. Indeed, results suggest that the tumour-virus interaction can be captured by a predator-prey feedback loop, governed by the law of mass action.

We remark that a similar computational, data-driven approach, has not been tested yet, and the accomplished steps are a promising direction for this field of research. By implementing and enhancing these techniques, new insights on OVT can hopefully be achieved.

Chapter 5

Conclusions and Future Ideas

Throughout this work, we have navigated the delicate balance between biological complexity and mathematical identifiability. We started from the simplest Sigmoidal growth models, questioning the very nature of the noise we assume in our measurements, and moved towards complex three-dimensional systems capable of describing oscillations.

The main experience of this journey is capture by the fact that a "good fit" is never enough. As we have seen through the use of Profile Likelihood and AIC_c, the risk of over-parametrisation is always around the corner. We have learned that "visual perfection" of the curves is not always a positive signal for a deeper understanding. What we can actually say about the parameters is usually more powerful at explaining the underlying dynamic.

Even with a robust MLE framework, we noticed that the optimiser often "gets lost" in the parameter space, especially when the model is flexible (like the Richards model) but the data is sparse, limited or ambiguous. This often results in "cut-offs" or non-identifiable parameters where the likelihood surface becomes flat. To overcome this, future developments may shift focus on more advanced global optimisation strategies. However, we also remark that the real solution might not be in the code, but in the lab, since the first element to start with is always the time series. By reversing the framework, we might give a significant contribution to this research field: instead of asking what to do with data, we investigated crucial identifiability thresholds, suggesting that improving in measurement collection frequency would yield to better fit results. This would bridge the gap between the blackboard and the Petri dish.

A crucial point arose during the analysis of the ODEs system in 4. When we looked at the oscillations, we noticed a fundamental discrepancy between the model's geometry and the biological reality. Our adimensionalised system works on what we can call a "constant support". The orbits, while capturing the frequency and amplitude, are essentially "memoryless" in their centering. Each oscillation is a new loop that doesn't account for the long-term drift of the system. In biology, however, *hysteresis* may occur. The memory of the previous cycle affects the next one, creating a trend our model cannot fully capture.

To fix this, we could investigate ways of adding a lag: a time delay could naturally represent the time it takes for the virus to replicate, creating the skewed geometry we observed, for example using delayed differential equation systems

In this thesis, we predominantly investigated *Category I* and oscillations. A natural next step is to apply the full multi-step pipeline to *Category II* datasets using

different models. It would be interesting to see if the superiority of the log-Normal approach holds across different treatment types and different tumours. Given the large noise and the small number of points, we suspect that advanced statistical methods may result in stronger and more reliable fit.

It is also important to remark that it is relevant to apply presented methods to other biological processes. The framework developed here is universal and can be easily exported to study of interactions between other populations or different tumour-immune interactions.

This thesis was never about finding a single "magic equation". It was an exercise in **Mathematical Honesty and Parsimony**: we have shown that while models are powerful tools, they are only as good as our ability to interrogate them. By acknowledging the limits of our functions and optimisers, the sparsity of our data, the difficulty of accurately capturing oscillations methods for improvement. By refining our statistical and mathematical tools, we are getting closer to understanding the "rules of the game".

Bibliography

[1] Hiroki Nagai and Young Hak Kim. Cancer prevention from the perspective of global cancer burden patterns. *Journal of Thoracic Disease*, 9(3):448–451, March 2017.

[2] Elizabeth Kelly and Stephen J Russell. History of oncolytic viruses: genesis to genetic engineering. *Molecular Therapy*, 15(4):651–659, February 2007.

[3] Nicole M Anderson and M Celeste Simon. The tumor microenvironment. *Current Biology*, 30(16):R921–R925, August 2020.

[4] Douglas Hanahan and Robert A Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, 2000.

[5] Palaniyandi Muthukutty and So Young Yoo. Oncolytic virus engineering and utilizations: Cancer immunotherapy perspective. *Viruses*, 15(8), July 2023.

[6] Kristin R Swanson, Carly Bridge, J D Murray, and Ellsworth C Alvord, Jr. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *J Neurol Sci*, 216(1):1–10, December 2003.

[7] WebPlotDigitizer. automeris.io. https://automeris.io/.

[8] A-Rum Yoon, Jinwoo Hong, and Chae-Ok Yun. A vesicular stomatitis virus glycoprotein epitope-incorporated oncolytic adenovirus overcomes CAR-dependency and shows markedly enhanced cancer cell killing and suppression of tumor growth. *Oncotarget*, 6(33):34875–34891, October 2015.

[9] Janina Marek, Lorenz Hanesch, Teresa Krabbe, Nadia El Khawanky, Simon Heidegger, and Jennifer Altomonte. Oncolytic virotherapy with chimeric VSV-NDV synergistically supports RIG-I-dependent checkpoint inhibitor immunotherapy. *Molecular Therapy - Oncolytics*, 30:117–131, August 2023.

[10] Joo-Hang Kim, Young-Sook Lee, Hoguen Kim, Jing-Hua Huang, A-Rum Yoon, and Chae-Ok Yun. Relaxin expression from tumor-targeting adenoviruses and its intratumoral spread, apoptosis induction, and efficacy. *Journal of the National Cancer Institute*, 98(20):1482–1493, October 2006.

[11] Jana de Sostoa, Carlos Alberto Fajardo, Rafael Moreno, Maria D Ramos, Martí Farrera-Sal, and Ramon Alemany. Targeting the tumor stroma with an oncolytic adenovirus secreting a fibroblast activation protein-targeted bispecific t-cell engager. *ournal for ImmunoTherapy of Cancer*, 7(1):19, January 2019.

[12] Ela Guo and Hana M Dobrovolny. Mathematical modeling of oncolytic virus therapy reveals role of the immune response. *Viruses*, 15(9), August 2023.

[13] Pierre-François Verhulst. Recherches mathématiques sur la loi d'accroissement de la population. *Mémoires de l'Académie royale de Belgique*, 1845.

[14] A K Laird. Dynamics of tumour growth: comparison of growth rates and extrapolation of growth curve to one cell. *British Journal of Cancer*, 19(2):278–291, June 1965.

[15] Matthew J. Simpson, Alexander P. Browning, David J. Warne, Oliver J. Maclaren, and Ruth E. Baker. Parameter identifiability and model selection for sigmoid population growth models. *Journal of Theoretical Biology*, 535:110998, 2022.

[16] Wuyue Yang, Dongyan Zhang, Liangrong Peng, Changjing Zhuge, and Liu Hong. Rational evaluation of various epidemic models based on the covid-19 data of china. *Epidemics*, 37:100501, 2021.

[17] Eckhard Limpert, Werner A. Stahel, and Markus Abbt. Log-normal distributions across the sciences: Keys and clues: On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability-normal or log-normal: That is the question. *BioScience*, 51(5):341–352, 05 2001.

[18] Dimiter Prodanov. Analytical solutions and parameter estimation of the SIR epidemic model. In *Mathematical Analysis of Infectious Diseases*, pages 163–189. Elsevier, 2022.

[19] Pantea Pooladvand, Chae-Ok Yun, A-Rum Yoon, Peter S Kim, and Federico Frascoli. The role of viral infectivity in oncolytic virotherapy outcomes: A mathematical study. *Mathematical Biosciences*, 334:108520, December 2020.

[20] Adrianne L. Jenner, Adelle C. F. Coster, Peter S. Kim, and Federico Frascoli. Treating cancerous cells with viruses: insights from a minimal model for oncolytic virotherapy. *Letters in Biomathematics*, 5(sup1):S117–S136, 2018.

[21] D G Wright, D C Dale, A S Fauci, and S M Wolff. Human cyclic neutropenia: clinical review and long-term follow-up of patients. *Medicine (Baltimore)*, 60(1):1–13, January 1981.

[22] H Vodopick, E M Rupp, C L Edwards, F A Goswitz, and J J Beauchamp. Spontaneous cyclic leukocytosis and thrombocytosis in chronic granulocytic leukemia. *New England Journal of Medicine*, 286(6):284–290, February 1972.

[23] C Haurie, D C Dale, and M C Mackey. Cyclical neutropenia and other periodic hematological disorders: a review of mechanisms and mathematical models. *Blood*, 92(8):2629–2640, October 1998.

[24] Isadore Brodsky, Hobart A. Reimann, and Lewis H. Dennis. Treatment of cyclic neutropenia with testosterone. *The American Journal of Medicine*, 38(5):802–806, 1965. Symposium on Viruses.