POLITECNICO
MILANO 1863

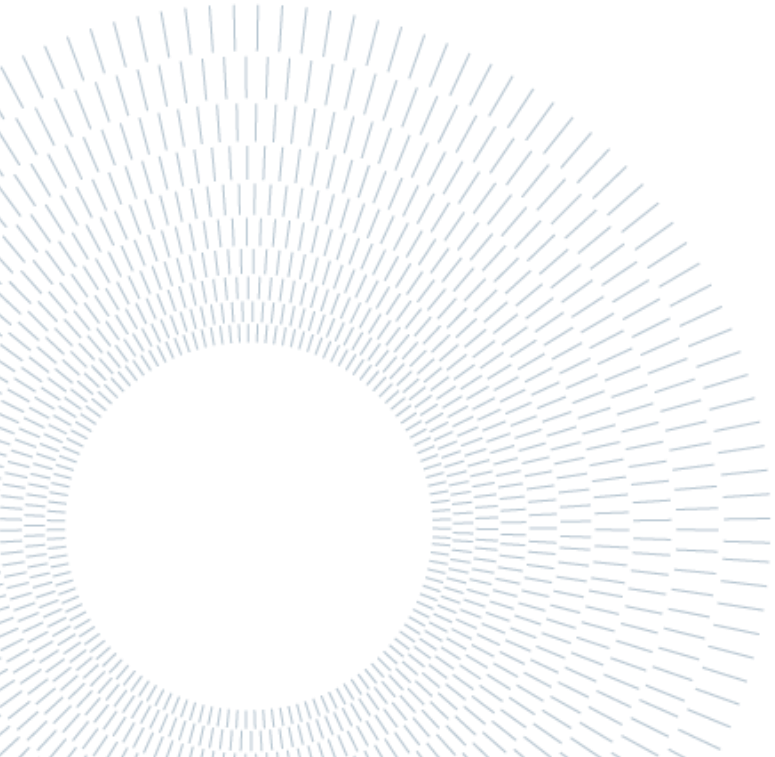
SCUOLA DI INGEGNERIA CIVILE, AMBIENTALE E TERRITORIALE

Modelling the disease ecology of Myxomatosis in European rabbits

TESI DI LAUREA MAGISTRALE IN
ENVIRONMENTAL AND LAND PLANNING ENGINEERING
INGEGNERIA PER L'AMBIENTE E IL TERRITORIO

Autore/Author: **Davide Bogani**

Matricola/Student ID: 996452
Relatore/Advisor: Prof. Marino Gatto
Correlatore/Co-advisor: Prof.ssa Isabella Cattadori
Anno accademico/Academic year: 2022-23



Abstract

The detection and control of emerging infectious diseases represents a major challenge in the modern world, due to multiple factors across pathogens, hosts and environment that generate complex non-linear dynamics. *Myxoma Virus* in the European rabbit (*Oryctolagus cuniculus*) has been and continues to be an excellent system to study the ecology and evolution of infectious diseases. Most of the scientific research has been focusing on the genetics of virus strains and the evolution of their virulence, including the immunology of host-virus interaction and the processes of virus transmission by mosquitoes and fleas. However, despite this richness of information, the modelling of the dynamics of myxoma virus within the host still requires in-depth analyses. This thesis aims to improve the understanding of the within-host dynamics of myxoma virus in the European rabbit.

Using data from laboratory infections, a mathematical model approach was developed to describe the viral load growth during the course of infection and the associated survival probability of the host, for different myxoma virus strains. The development of the models was based on a trade-off between accuracy and complexity, and the most representative model was selected through the application of the Akaike Information Criterion. The virulence of the strains affected the selection of the viral load models. For medium to high virulence strains that overcome the host immune response, growing over the entire course of the infection, the model selection favored Malthusian or logistic growth. On the other hand, for less virulent strains that are controlled by the host, the model selection favored a framework that includes the effect of the immune response on myxoma virus regulation as a dynamic killing rate of the virus. The host survival probability over time was analyzed by developing a model that describes the mortality rate as a function of the viral load. The model selection process identified model structures with threshold as the best option: the mortality rate remains null below a certain load of myxoma virus, after which it increases proportionally to the viral growth for high virulent strains, or it assumes a constant value for mid-virulent strains. Overall, it is possible to observe that the growth rate of a strain is not the only element to characterize its virulence, quantified as host fatality rate. The virulence of a strain is also clearly associated with the severity of the disease effects, expressed by the intensity parameter of the survival model, whereas the deadly load thresholds do not seem to show a pattern correlated with the virulence grade of the strains.

To further explore the interaction between virus and host, this thesis also examined

the dynamics of myxoma virus in rabbit co-infected with *Trichostrongylus retortaeformis* gastrointestinal helminths. Specifically, using data from laboratory experiments, a series of statistical tests were carried out to investigate differences in host survival time and viral load at host death between rabbits co-infected with helminths and rabbits infected only with myxoma virus. For most viral strains there is no significant evidence to reject the null hypothesis of similarity of virus growth and host survival between dual- and single-infected rabbits. Helminths abundance at host death was also considered in the analysis and the inclusion of an interaction between viral load and number of helminths at death provided much more accurate predictions of survival times compared to analyses that considered only an additive effect of these two terms. Similarly, the interaction between survival time and helminth abundance turned out to be a relevant term in predicting viral loads at host death. Nevertheless, the model coefficients of these interactions differ considerably between strains, highlighting a strong variability in the outcomes.

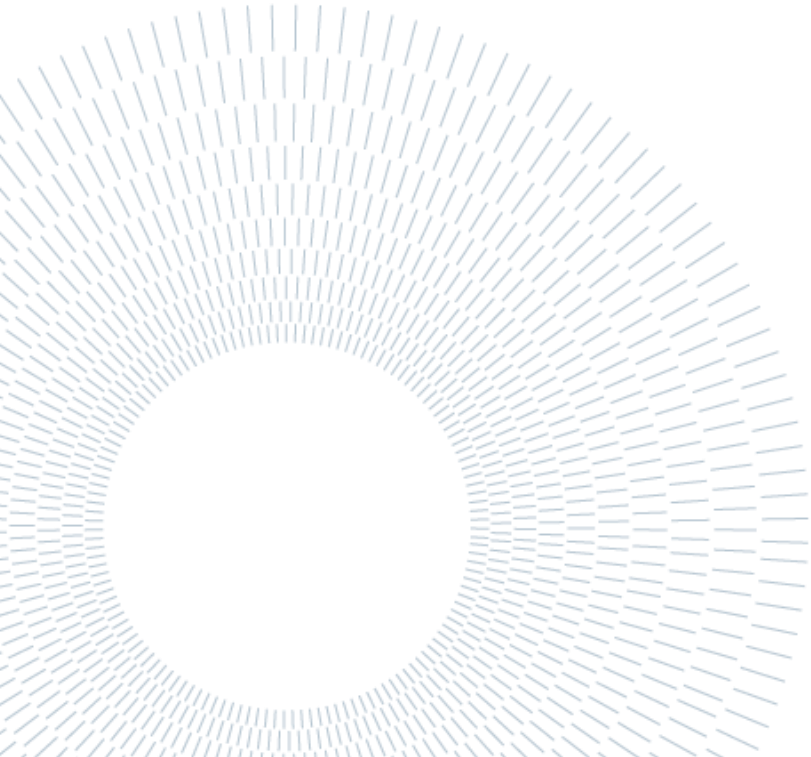
Keywords: Myxoma virus, within-host models of viral dynamics, host survival models, gastrointestinal helminths, consequences of co-infection.

Sommario

La diagnosi e il controllo di malattie infettive emergenti rappresentano una grande sfida nel mondo moderno, a causa di molteplici fattori di interazione fra patogeni, ospiti e ambiente che generano complesse dinamiche non lineari. Il *Myxoma virus* nel coniglio Europeo (*Oryctolagus cuniculus*) è stato e continua ad essere un eccellente sistema per studiare l'ecologia e l'evoluzione di malattie infettive. La maggior parte dello sforzo di ricerca scientifica è stato concentrato sulla genetica dei ceppi virali e sull'evoluzione della virulenza, incluse l'immunologia dell'interazione ospite-virus e la dinamica di trasmissione del virus da parte di zanzare e pulci. Tuttavia, nonostante questa ricchezza di informazioni, la modellazione delle dinamiche del myxoma virus all'interno dell'ospite necessita ancora di analisi approfondite. Questa tesi ambisce a migliorare la comprensione delle dinamiche del virus all'interno del coniglio Europeo. Un approccio modellistico è stato sviluppato, a partire da dati di laboratorio, per descrivere l'evoluzione temporale del carico virale durante il corso dell'infezione e la relativa probabilità di sopravvivenza dell'ospite, per diversi ceppi virali. Lo sviluppo di tali modelli si è basato su un trade-off fra accuratezza e complessità e il modello più rappresentativo è stato selezionato attraverso l'applicazione del Criterio di Informazione di Akaike. La virulenza dei ceppi virali ha influenzato la selezione dei rispettivi modelli di carico virale. Per ceppi di virulenza medio o alta che prevalgono sul sistema immunitario dell'ospite, con una tendenza di crescita continua durante l'infezione, la selezione ha favorito modelli di tipo malthusiano o logistico. Invece, per ceppi virali meno virulenti che l'ospite riesce a controllare, il processo di selezione dei modelli ha favorito uno schema che include l'effetto della risposta immunitaria sotto forma di un tasso dinamico di uccisione del virus. La probabilità di sopravvivenza dell'ospite nel tempo è stata analizzata sviluppando un modello del tasso di mortalità in funzione del carico virale. Il processo di selezione ha identificato modelli con soglia come miglior opzione: il tasso di mortalità è nullo sotto ad una certa quantità di carico virale, oltre il quale cresce proporzionalmente alla crescita virale per ceppi molto virulenti o assume un valore costante per ceppi di media virulenza. Combinando i modelli ottenuti è possibile constatare come la rapidità di crescita di un ceppo non sia l'unico elemento caratterizzante la sua virulenza, quantificata dal tasso di fatalità degli ospiti. La virulenza di un ceppo è anche associata alla gravità delle conseguenze della malattia, espresse dal parametro di intensità del modello di sopravvivenza, mentre i valori di soglia mortale del carico virale non sembrano mostrare un pattern correlato con il grado di virulenza.

Per esplorare ulteriormente l'interazione fra virus e ospite, questo lavoro di tesi esamina anche le dinamiche del virus myxoma in conigli co-infettati con elminti gastrointestinali *Trichostrongylus retortaeformis*. A partire da dati di laboratorio, una serie di test statistici è stata eseguita per cercare differenze dei tempi di sopravvivenza e dei carichi virali alla morte fra conigli co-infettati con elminti e conigli infettati solo con virus myxoma. Per la maggior parte dei ceppi virali non è stato possibile individuare evidenze sufficienti per rifiutare l'ipotesi nulla di similarità di crescita del virus e sopravvivenza degli ospiti fra conigli infettati singolarmente o dualmente. L'abbondanza di elminti alla morte dell'ospite è stata anche considerata nell'analisi e l'inclusione di una interazione fra carico virale alla morte e numero di elminti ha restituito previsioni dei tempi di sopravvivenza molto più accurate di quelle ottenute considerando solo un effetto additivo fra questi due termini. Allo stesso modo, l'interazione fra tempo di sopravvivenza e abbondanza di elminti alla morte è risultata un termine rilevante nella previsione del carico virale alla morte. Ciononostante, i coefficienti di tali interazioni nei modelli differiscono considerevolmente fra diversi ceppi, evidenziando una forte variabilità degli esiti.

Parole chiave: virus Myxoma, modelli di dinamiche virali all'interno dell'ospite, modelli di sopravvivenza dell'ospite, elminti gastrointestinali, conseguenze della co-infezione.



Contents

1 Context

In modern times, a remarkable number of novel infectious agents of humans and animals are regularly discovered. Understanding and predicting the outcomes of emerging pathogens is critical for the control and prevention of epidemics. This task needs knowledge of the ecology of the system, accurate data for the hosts and the pathogens, including environmental components affecting their dynamics, and modelling tools to capture the processes that generate uncertainty and variability in host-pathogen interactions. In this context, the release of *Myxoma virus* in the European rabbit represents an ongoing and classical example of co-evolution of virus virulence and host resistance, through an impressive biological experiment conducted in Australia and Europe [1].

1.1. Myxoma virus

Myxoma virus (MYXV) is a double-stranded DNA virus, genus *Leporipoxvirus* (*Poxviridae* family, *Chordopoxvirinae* subfamily). The virus is predominantly vector-borne and transmitted by mosquitoes or fleas, which probe through the epidermis of an infected host, pick-up virus particles in their mouthparts and inoculate them in the next animal during the feeding process. MYXV does not replicate in mosquitoes and fleas, which simply act as mechanical vectors by moving the virus from one infected host to a naïve animal [2].

MYXV was naturally found in South America circulating as an endemic infection in local populations of Brazilian rabbits (*Sylvilagus brasiliensis*). The virus was generally innocuous, mainly causing skin lesions and localized fibromas that are cleared by the immune system in a matter of a few weeks [3]. In contrast, in the European rabbits (*Oryctolagus cuniculus*) MYXV was found to be extremely lethal and, given its characteristics, it was considered a potential tool for biological control of the rabbit population in Australia and Europe. In fact, the introduction of European rabbits in Australia in 1859 [4] and the subsequent exponential growth through the country caused alarming agricultural lost and damage of arable land, preventing the regeneration of shrub, encouraging soil erosion and consuming/ruining crops. After multiple unsuccessful attempts, in 1950 an isolate strain of MYXV from Brazil (afterwards named *Standard Laboratory Strain*) was released in south-east Australia, where the virus quickly spread killing millions of rabbits in a few months, causing a lethal disease in these new hosts named *myxomatosis* [5].

A short time later, in 1952, a similar scenario happened in Europe. Two wild rabbits were intentionally inoculated in France with another highly virulent MYXV strain (*Lausanne*). The virus rapidly spread across Europe decimating the rabbit populations. Surprisingly, following the pathogen spread in both continents, attenuated strains (case fatality rate = 60% – 95%, i.e. the percentage of rabbits diagnosed with MYXV that ended up dying for it) rapidly evolved from the initially progenitors (case fatality rate ≥ 99.8%) [6]. The attenuated strains lead to longer surviving rabbits, which prolonged the infectious window, and thus the probability of transmission by mosquitoes or fleas. For instance, a viral strain that initially killed 90% of rabbits caught at Lake Urana was killing only 26% of rabbits in the same location 7 years later [7]. In parallel, it was also observed that rabbits developed an increasing resistance against the virus, which forced a selection for stronger virulence, as too attenuated viruses were controlled by the immune response and transmitted poorly [8]. This unstable balance of ongoing counter adaptations between MXXV virulence and rabbit resistance has been the force driving the co-evolutionary pathway of the MYXV-rabbit system, commonly known as a biological host-pathogen "arms race" [9].
A grading system quantifying the virulence of MYXV strains was proposed by Fenner [10] and it was based on case fatality rate (CFR) and average survival time (AST) of the infected rabbits. Five classes were identified, as reported in Table 1-1 [9].

Table 1-1: Virulence grading system for myxoma virus strains.

Virulence grade	AST [days]	CFR [%]
1	≤ 13	100
2	14 - 16	95-99
3	17 - 28	70-95
4	29 - 50	50-70
5	51 - indeterminate	≤ 50

The virus never became benign or completely attenuated and many of the MYXV lineages evolved from the initial progenitors are currently circulating in rabbit populations [10]. Over the years, the trade-off between transmission and resistance has favoured the circulation of MYXV strains of grade 3 virulence [11]. However, emerging pathogens can show remarkably high peak of virulence during the initial epidemic phase before the beginning of transient selections [12]. After 60 years of virus-rabbit co-evolution, a final equilibrium of this arms race has not yet been foreseen [13].

MYXV has not changed just in terms of transmissibility and fatality rate. In the 1950s the most evident clinical manifestation of *Myxomatosis* consisted of nodular skin

tumours, caused by virus replication in cutaneous and mucocutaneous sites such as anogenital region, eyelids, nasal cavities, and the base of the ears. Associated with these symptoms, the hosts showed swollen head, anogenital region and eyelids, with mucoid discharge from the nostrils, drooping ears, upper respiratory tract occlusion and strong inflammatory cell responses [6]. Despite this symptomatology, the definitive cause of death was partially obscure; only a limited load of MYXV was recorded in key organs of dead rabbits, although there was clear evidence that the highly virulent strains caused massive immunosuppression, by which rabbits were incapable to clear the infection and fully recover [14].
In the following decades, some of the circulating MYXV strains disappeared and new emerged, with new disease phenotypes raising: pulmonary oedemas and hemorrhages, inhibition of the immune response, depletion or necrosis of lymphocytes, loss of neutrophils, opportunistic bacterial infections, sudden reactions similar to sceptic shocks and acute cytokines storms non-associated with inflammatory cell response [11].

Laboratory studies have been very successful in disentangling the genes potentially involved in virus virulence and the processes that allowed the virus to manipulate and evade the host immune response during the initial release and subsequent spread into rabbit populations. A large number of MYXV strains were examined and great advancements were made in the comprehension of how this virus interacts with European rabbits, including the role of vector transmission. Ultimately, the study of MYXV in rabbits offers a rare opportunity to bring together molecular, ecological and genetic studies to understand the past, present and future of this pathogen-host interaction, as a unique case-study of infectious diseases, a topic that in the very recent years has drastically impacted the lives of all with the case of SARS-CoV-2 [15].

1.2. The role of co-infection

Co-infection, i.e. the simultaneous infection of a host with multiple pathogens species (*sensu lato*), is anything but an exception in natural systems, including human populations [16]. Indeed, it is very common for a free-living organism to be infected by multiple pathogens and a balanced parasite community is in some cases a sign of a healthy population, as could be the community of bacteria in the gut on an individual. In the hosts, co-infecting pathogens can exhibit positive or negative interactions, or show no association [17]. In the latter case, the infections and the immune responses occur independently from each other. In the case of interactions, instead, two non-mutually exclusive patterns can be observed. Pathogens can interact directly, for example competing for resources, which should be expected when two taxonomically similar pathogens select the same niche or food. Alternatively, pathogens can interact indirectly, for example when they infect different organs or tissues and the interaction is, for example, via the effects on the host immune response, through cross-immunity or antagonist immunity [18]. In presence of a cross-immunity situation, the host is

simultaneously co-infected with two different pathogens that force a shared immune response, while during antagonistic immunity the host is simultaneously co-infected with two different pathogens, but each parasite initiates a different response of the immune system.
The outcomes of co-infections are strongly determined by the pathogens involved, the relative order in which they infected the host, the dynamics of their interaction and the activity of the immune system. Co-infection outcomes can be neutral, detrimental, or beneficial, with consequences that can strongly influence the host's wellbeing, including diagnostic tests, and drug treatments [19]. The presence of detrimental co-infection means that there is an aggravation of the hosts' condition compared to the conditions caused by the pathogens if occurring alone. These conditions involve mortality, morbidity, intensity of the immune response, reactions to therapies and treatments, and transmissibility. To provide a well-known example, HIV immunodeficiency increases the risks when interacting with tuberculosis or hepatitis [19]. By contrast, beneficial outcomes concern a situation in which co-infection can reduce some of the aspects of severity of the single infections. As an example, the nematode *Ascaris lumbricoides* resulted in many cases protective from malaria and its severe manifestations [20]. A neutral co-infection indicates that there are no substantial diverse effects compared to those caused by the pathogens in single-infection processes.

Because helminths are among the most prevalent parasites on the planet and affect almost a quarter of the world human population [21], the study of the dynamics of co-infection with helminths can provide important insights into the mechanisms of pathogen regulation. This information can also be used to develop intervention strategies targeted to the control of multiple infections in humans and livestock [22] and to the assessment of the zoonotic risk. In general, the hosts become infected with helminths by exposure to media contaminated with free-living parasite stages like food, water bodies and soils or by interaction with vectors carrying helminths. Helminths can have a direct or complex life cycle, in the first case infection occurs by direct exposure to infective stages, while in the second case parasite maturation must go through an intermediate host that releases infective stages ready to infect the definitive host [23]. In the host larvae develop into adults, reproduce sexually, and females shed eggs or larvae that are expelled in the environment through the host's feces.
The interaction of helminths with other pathogens like viruses or bacteria can be indirect and characterized by the stimulation of two broad branches of the immune response: type 1 helper T cells biased immune response (Th1) is developed against intracellular pathogens (like viruses) and type 2 helper T cells biased immune response (Th2) is developed against the helminths [24]. These antagonistic forces cause a split of the response, often reducing its effectiveness against one of them. While co-infection modelling cannot rely only on this Th1-Th2 paradigm to explain complex interactions

[25], this is a general useful starting approach.
Scientific literature already provides many examples of helminth interactions with other pathogens, with various outcomes: African buffaloes with gastrointestinal helminths and *Mycobacterium bovis* bacterium suffer an accelerated death, while gray treefrogs coinfected with the helminth *Echinoparyphium* and the *Ranavirus* virus have lower viral load than individuals only infected by the virus [26]. The comprehension of these interaction dynamics can be an essential step to evaluate how to treat similar cases, for example to understand whether deworming could be a reliable strategy or even a pejorative approach [27]. In this context, the co-infection of *Myxoma virus* and gastrointestinal helminths *Trichostrongylus retortaeformis* in European rabbits (*Oryctolagus cuniculus*) can be considered a useful case-study of multi-parasite infections.
Using mathematical models, the impact of MYXV on the susceptibility of *Trichostrongylus retortaeformis* was investigated in natural populations of rabbits [24], suggesting that co-infected rabbits face higher helminth intensity compared with virus-free hosts. In separate research [28], it was also observed that the grade of MYXV affects some relevant helminths aspects such as difference in male and female distribution in the host, female fecundity and number of eggs in utero. The effects of helminths on MYXV dynamics remains instead unclear. On the side of transmissibility, it was studied whether gastrointestinal helminths played a role in the MYXV-hosts arms race that followed the initial release in rabbit populations [29]. On the side of the infection development, moving from between-hosts to within-host point of view, the effects of helminth co-infection on the MYXV dynamics have yet to be understood. Detailed analyses on this front should be done to achieve a more exhaustive understanding of the dynamics of this co-infection system.

1.3. Modelling infectious diseases

Mathematical models represent an invaluable approach to study the dynamics of infectious diseases, both within-host and between-hosts. Indeed, the development of models was proved to be essential to analyse the mechanisms behind the processes of infection and transmission and to identify how complex processes of host-virus interaction emerge [30]. Modelling can be applied to examine laboratory and field data, to simulate possible scenarios and interactions, and to study the relative contribution of multiple factors that can restrict or expand the conditions for which a pathogen can prosper [31].
The within-host interactions between pathogens and the immune system are often characterized by large variability, since the replication power of the pathogens strongly depends on the host individual characteristics and on load and distribution of the virus in the body. Therefore, the development of mathematical models must find an adequate trade-off between complexity and accuracy. On one side, unnecessary complexity can obscure fundamental results and may require large

computational and data collection efforts; on the other side over-simplification can miss the detection of mechanisms fundamental to understand the processes of infection and transmission. If properly obtained, model outcomes can play a critical role in public health control and prevention [32], by testing fundamental scientific hypotheses, informing data-collection strategies and assisting the design of disease-control policies.

2 Scope and approach

Myxoma virus (MYXV) has been and continues to be a fascinating pathogen to study. The within-host kinetics of infection and the epidemiology of transmission at the population level provide opportunities to examine patterns and mechanisms that generate variability in host-virus interaction. The work behind this thesis aims to contribute to the understanding of MYXV within-host dynamics by addressing two specific goals:

- To develop solid and simple mathematical models to describe the growth of different strains of MYXV in the European rabbit, until eventual death, using laboratory data.
- To investigate how the co-infection with the gastrointestinal helminth *Trichostrongylus retortaeformis* affects MYXV viral load at death and host survival time.

The first part of this research compares different model frameworks to describe the infection with MYXV strains of different virulence in the rabbits. The performed work examines the viral growth post infection (longitudinal modelling) as well as the survival statistics at death time (cross-sectional modelling). The complex interaction between MYXV strains and rabbit responses makes modelling the biological processes extremely challenging. The aim of this work is to develop a system of conceptual models with the tools of population ecology, based on a trade-off between accuracy and complexity, able to provide a parsimonious explanation of the patterns exhibited by different strains.

The second part of this thesis moves from the general modelling of within-host MYXV kinetics to focus on the co-infection between *Myxoma virus* and *Trichostrongylus retortaeformis* in the European rabbit. In this perspective, this work aims to investigate the effects of helminths on MYXV growth post infection and on the survival time of the host, including the comparison of these effects between MYXV strains of different virulence. This research was conducted considering two approaches: i- the binary presence or absence of helminths, namely the comparison between rabbits infected with MYXV with and without helminths, and ii- the analysis of the estimated helminth abundance in the hosts at death time. A variety of parametric and non-parametric statistical tests were performed to seek significant differences between single-infected and co-infected rabbits. The estimated number of helminths at death

was introduced to test different hypotheses of interaction, using the selection of model parameters to identify the role of helminth co-infection on the prediction of survival time and viral load at death.

3 Available data

The work presented in this thesis is based on laboratory data provided by Cattadori's laboratory at the Center for Infectious Diseases (CIDD) of Pennsylvania State University (PSU) and by available published literature. Data from Cattadori's laboratory were collected from previous experiments as part of projects on the evolution of *Myxoma virus* (MYXV) virulence and *Myxoma virus-Trichostrongylus retortaeformis* co-infection. This chapter summarizes the structure of the datasets and the related methodological differences of the associated experiments.

3.1. Datasets

Four different datasets were subjected to analyses.

3.1.1. Longitudinal modelling: recent data (Kerr dataset)

This dataset contains the results of experiments conducted at PSU in 2012. The dataset monitors the longitudinal growth of MYXV load in the primary lesion of 6 laboratory rabbits. At fixed days post infection, the amount of virus in three of the six rabbits is measured with biopsies in the primary lesion, alternating group of three animals. The dataset structure is as follows:

- Strain: the administered strain of MYXV, namely: *Swh 8/2/93*, *Ws6 346*, *Bd23*, *Swh 1209*, *Brk 12/2/93*, *Uriarra* and *Km13*.
- ID: identification number assigned to each rabbit, unique inside each strain.
- Survival time: death time in days considering the infection with MYXV as time=0. All the alive animals were euthanized at the end of the experiments.
- Virus amount: quantity of MYXV measured in the primary lesion on fixed days post infection [PCR_{copies}/mg].

The inoculation of 100 plaque farming units (PFU) of MYXV in 0.1 ml of phosphate buffered saline (PBS) solution was realized by one single intradermal administration in the right side of the individual's rump. Tissue at the primary lesion was subjected to disposable dermal punch biopsies and the amount of virus was measured at fixed days by quantitative polymerase chain reaction (qPCR). Further indications of the experimental methods can be found in [9]. The strains involved in this source come from Australia in the 1950s and the 1990s.

3.1.2. Longitudinal modelling: historical data (Fenner dataset)

This dataset contains the results of experiments conducted at the John Curtin School of Medical Research (JCSMR) of the Australian National University (ANU), before 1956. The dataset describes the longitudinal growth of MYXV load of different strains in 2 laboratory rabbits. The dataset structure is as follows:

- Strain: the administered strain of MYXV virus, namely: *Sls*, *Neuromyxoma*, *Km13 II*, *Uriarra III*, *Loiret 55* and *Lausanne*.
- ID: identification number assigned to each rabbit, unique inside each strain.
- Virus amount: load of MYXV measured in the rabbit's primary lesion on fixed days post infection [RID/g].

The quantification of MYXV load follows a different procedure compared to the dataset in section 3.1.1. The rabbits were infected intradermally in ten marked skin sites with 0.1 ml of a solution containing 10E+4.3 rabbit infectious doses (RID). At intervals of 1 or 2 days, slices of skin from the occurring lesions were ground and titrated on the chorioallantois of developing chick embryos. Further indications of the experimental methods can be found in [2]. The strains involved in this source come from Australia and Europe in the 1950s and before.

3.1.3. Cross-sectional modelling: survival titers (survival dataset)

This dataset contains the results of experiments conducted at PSU in 2013. The survival time and the amount of virus in 6 rabbits was collected at death for several different strains. The dataset structure is as follows:

- Strain: the administered strain of MYXV, namely: *Perthshire 1537*, *Perthshire 1792*, *Perthshire 2082*, *Perthshire 2282*, *Perthshire 1527*, *Yorkshire 127*, *Yorkshire 135* and *Yorkshire Col*.
- ID: identification number assigned to each rabbit, unique inside each strain.
- Survival time: death time in days considering the infection with MYXV as time=0. All the alive animals were euthanized at the end of the experiments.
- Virus amount: quantity of MYXV contained in the lungs and in the primary site of infection at death time [PFU/g].

The virus load was inoculated in the amount of 100 plaque farming units (PFU) in 0.1 ml of phosphate buffered saline (PBS) solution by intradermal administration in the right side of the rabbit's rump and the load was measured by cell culture. The unit of measure is conceptually similar to the one used in the 'Fenner dataset' (rabbit infectious doses, section 3.1.2), with the difference that in the first case MYXV was grown on chorioallantois chicken membranes rather than on other specific cells on *Petri* dishes. Further indications of the experimental methods can be found in [3]. In contrast, the experiments that provided the 'Kerr dataset' (section 3.1.1) used another technique (qPCR), which makes the comparison between dataset less straightforward.

The strains involved in this source are the most recent, coming from Europe after year 2000.

3.1.4. Cross-sectional modelling: co-infection titers (co-infection dataset)

This dataset contains the results of experiments conducted at PSU in 2014, 2015 and 2017. The dataset includes data at death time of MYXV single-infected and MYXV-*Trichostrongylus retortaeformis* co-infected rabbits. The dataset structure is as follows:

- Strain: the administered strain of MYXV, namely: *Coomandook* (32 rabbits), *Perthshire 2082* (16 rabbits), *Ws 61071* (16 rabbits), *Lausanne* (20 rabbits) and *Sls* (20 rabbits).
- ID: identification number assigned to each rabbit, unique inside each strain.
- Survival time: death time in days considering the infection with MYXV as time=0. All the alive animals were euthanized at the end of the experiments.
- Virus amount: MYXV load in the eyelid and in the lungs at death time [PFU/g].
- Presence of helminths: binary indication of whether the rabbit was co-infected with *Trichostrongylus retortaeformis*.
- Helminths amount: the abundance of *Trichostrongylus retortaeformis* in the small intestine of the co-infected host at death time.

Briefly, for each MYXV strain, half of the available rabbits were periodically infected (gavage) with 800 third-stage larvae of gastrointestinal helminths *Trichostrongylus retortaeformis*, to simulate the natural ingestion with contaminated herbage. The helminth infections of experiments with strains *Ws 61071* and *Perthshire 2082* occurred on days -17, -10, -3, +4, +11 and +18 in relation to MYXV infection, while for strains *Coomandook*, *Lausanne* and *Sls* only on days -17, -10, -3, +4 and +11. Helminth infection was started before the inoculation of the virus to allow the development of an immune response against the helminths (developed as a function of cumulated exposure to the parasite). Further details regarding the within-host growth of helminths in presence of MYXV can be found in [28]. At death time the total number of helminths in the small intestine was estimated using 2.5ml aliquots. At the same time, the other animals were infected only with MYXV, to provide a baseline for comparison. The data of this dataset are unpublished. The strains involved in this source come from Australia and Europe in the 1950s and the 1990s.

3.2. Summary of available strains

A summary of all the available strains is reported in Table 3-1. This collection represents a wide range of virulence grades, but many strains are available in only one of the datasets. For this reason, some longitudinal virus data lack information of host survival times and vice versa. Moreover, different units of measurement and different sites of collection warrant attention in the interpretation of the research outcomes.

Table 3-1: Collection of the MYXV strains under research. Where the degree of virulence was found to be variable in different experimental tests, two values are reported.

Strain	Virulence grade	Dataset
Lausanne	1	Fenner dataset; co-infection dataset
Sls	1	Fenner dataset; co-infection dataset
Yorkshire 135	1	Survival dataset
Brk 12/2/93	1-2	Kerr dataset
Bd23	1-2	Kerr dataset
Ws6 346	2	Kerr dataset
Swh 8/2/93	2	Kerr dataset
Yorkshire Col	2	Survival dataset
Perthshire 1792	2	Survival dataset
Yorkshire 127	2.3-3	Survival dataset
Swh 1209	3	Kerr dataset
Km 13 II	3	Fenner dataset
Perthshire 2282	3	Survival dataset
Perthshire 2082	3-3.1	Co-infection dataset; survival dataset
Ws 61071	3.1	Co-infection dataset
Coomandook	3.2	Co-infection dataset
Perthshire 1537	3-3.4	Survival dataset
Uriarra III	3-4	Fenner dataset
Loiret 55	4	Fenner dataset
Uriarra	5	Kerr dataset
Neuromyxoma	5	Fenner dataset
Km13	5	Kerr dataset
Perthshire 1527	5	Survival dataset

4 Modelling of myxoma virus dynamics

From its discovery up to these days, *Myxoma virus* (MYXV) evolved into many strains with different grades of virulence. The infection consequences on the host have also changed, spreading from fibromas to immune collapses [33]. The purpose of this chapter is to develop models which can provide a simple but accurate way to describe the within-host growth of a broad range of strains and the survival of the host. The optimal models among the developed options were identified by using a selection process based on a trade-off between accuracy and complexity through the Akaike Information Criterion [34]. First, a set of models for the within-host viral growth was selected. Then, the virus growth was linked with the host survival by modelling the mortality rate as a function of the viral load, with a structure selected by fitting the consequent survival probabilities in time on the experimental survival curves. The uncertainty of the optimized models was inspected by developing confidence intervals of the simulations.
After infection, MYXV quickly grows in the host. Due to this fast replication process and the high biological variability, the virus load (V_t) is measured in logarithmic units, as usually performed in relevant literature. The datasets involved in the following analyses are the 'Kerr dataset' (section 3.1.1), which provides hosts' survival times and measurements of the viral load along time for a series of strains, the 'Fenner dataset' (section 3.1.2), which provide measurements of the viral load along time for a series of strains and the 'Survival dataset' (section 3.1.3), that provides hosts' survival times for a series of strains. All the analysis of this chapter were performed using MATLAB tools (*ode45*, *fminsearch*, *fminsearchbnd* [35]).

4.1. Modelling of virus growth in time

As far as the 'Kerr dataset' is concerned, in some fixed days post infection (DPI) the amount of virus was measured in 3 of the 6 rabbits infected with each strain, alternating the rabbits. Instead, the 'Fenner dataset' monitors the viral load of the same two rabbits per strain, in fixed DPI. Even if different from each other, both datasets are suited to perform a longitudinal analysis with the same method. The strains in 'Kerr dataset' and 'Fenner dataset' have different units of measures, so a generic unit *[V]* is used later in this chapter to express suitable units.

The viral load dynamics of the strains are showed in Figure 4-1 ('Kerr dataset') and Figure 4-2 ('Fenner dataset'). For each strain, the rabbits are marked by colors.

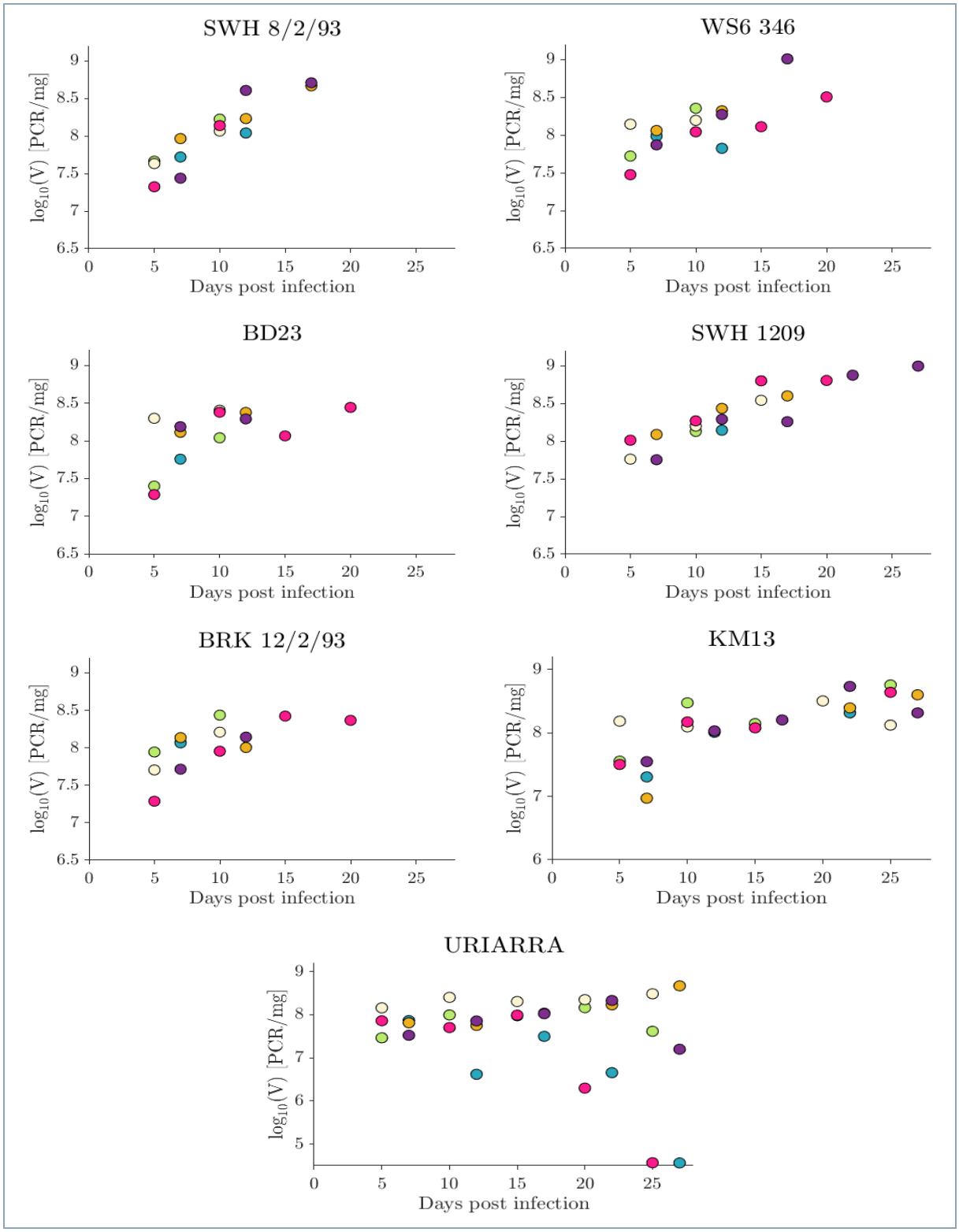


Figure 4-1: Representation of viral load in time for different strains from 'Kerr dataset'. Inside each strain group, the rabbits are distinguished by colors.

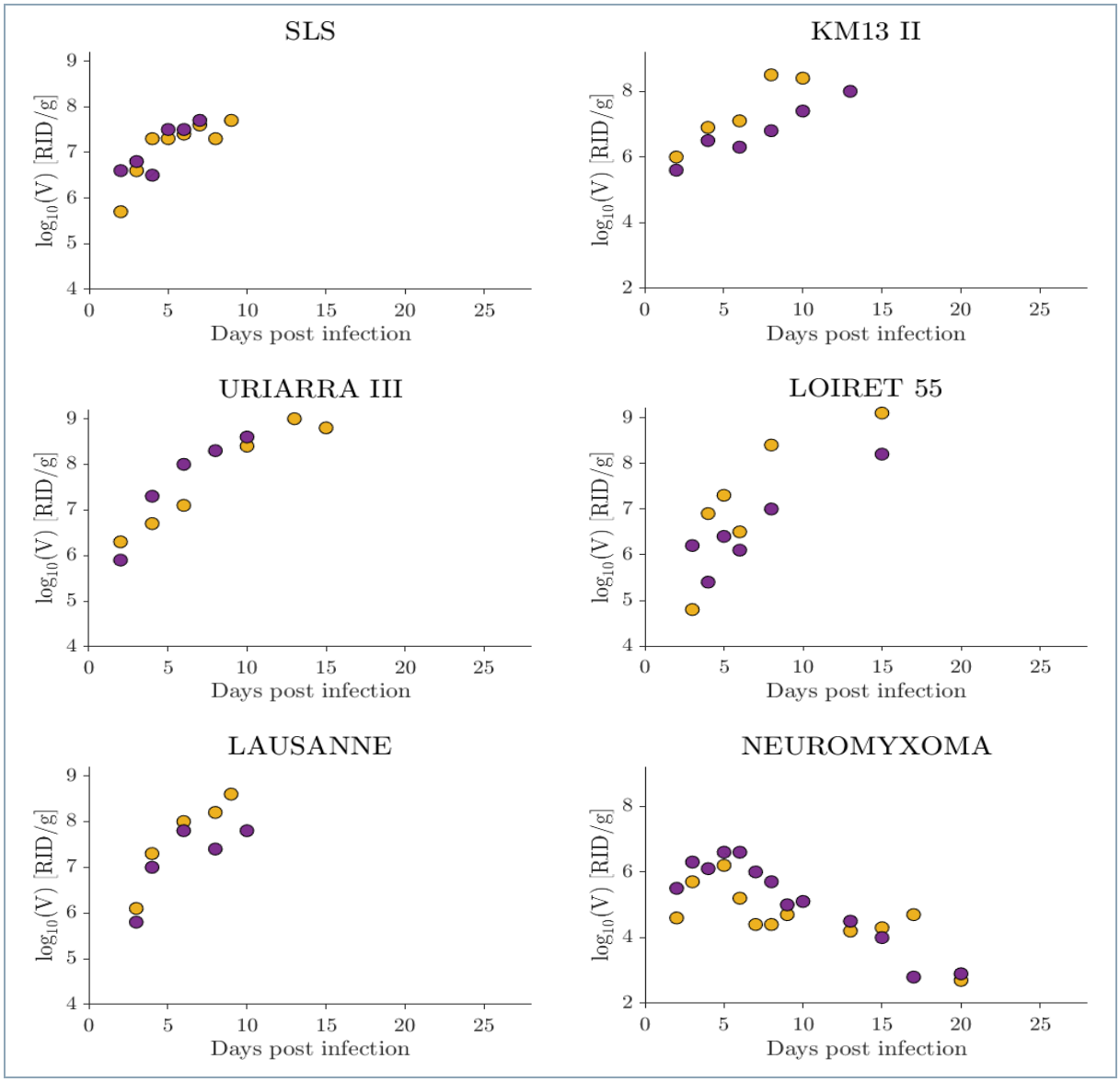


Figure 4-2: Representation of viral load in time for different strains from 'Fenner dataset'. Inside each strain group, the rabbits are distinguished by colors.

It is possible to observe that *Neuromyxoma* ('Fenner dataset') and *Uriarra* ('Kerr dataset') act differently from all the other strains. In fact, these two strains are the only cases where the host immune system is able to efficiently control the virus growth, that decreases after an initial increase. The effect is rather clear for *Neuromyxoma* and less evident for *Uriarra*, since two of the six rabbits are still in the increasing phase in the last days of the experiment.

The data size is not suitable to run a model rabbit by rabbit, hence the fit was realized using the data of all the rabbits infected by the strain together. Since more than one rabbit is sampled for each DPI, the residuals represent not only the temporal variability, but also the variability of the rabbits themselves. First, the within-host viral

load was modelled as a growing viral population with no predators and no competitors. The effect of the hosts' immune response was then introduced, described as a killing rate of the virus. Once the best model for each strain was identified, the confidence intervals of simulations and the parameters were obtained by applying a bootstrap approach [36].

4.1.1. Models in literature

The main model that can be found in literature to describe the within-host viral load is a simple function of time developed by Dwyer et al. [13] as an input for a study of MYXV in rabbits' populations. The following equation was applied (see Figure 4-3) to represent the virus load in time for various strains, collected by Fenner et al. [2] and grouped by degree of virulence, as part of an analysis of populations.

$$log_{10}(V_t) = log_{10}(K_1 te^{-K_2 t}) \quad (4.1)$$

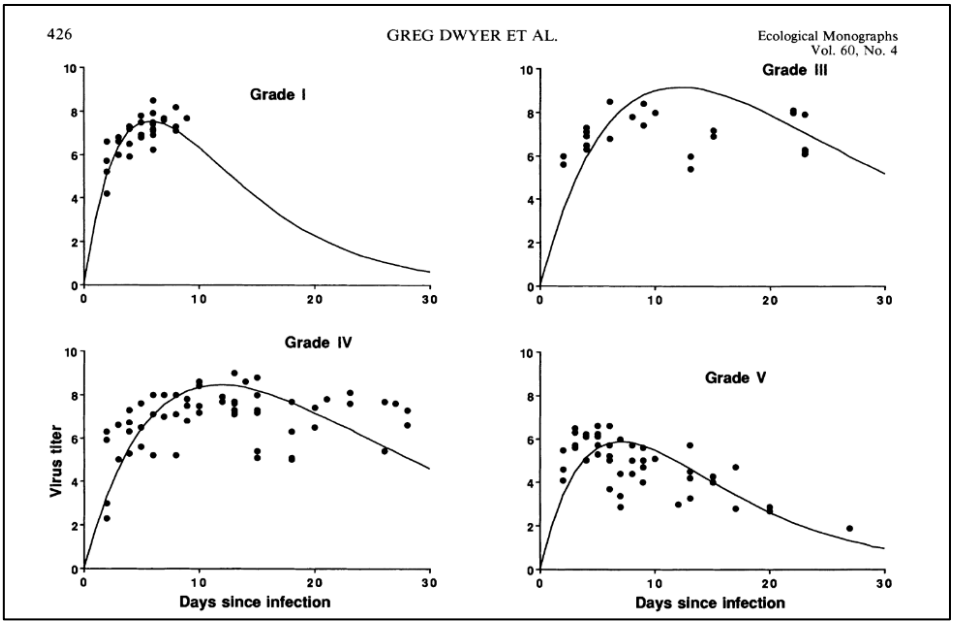


Figure 4-3: Applications of model (4. 1) for different virus strains: $\log_{10}$(V) [RID/g] in time post infection [days], by Dwyer et al. [13].

The choice of this function to describe the within-host viral growth leads to the identification of a decrease of viral load in time. However, the less virulent strain (grade 5) is the only one that seems to effectively control MYXV and generate such behavior. For the other cases, there is no general evidence of a decreasing turning point during the infection, characterized instead by a virus load that initially grows fast until it slows down towards a plateau until the death of the hosts. The decline of virus load

takes place for the small percentage of long-term survivors. Function (4. 1) is therefore not completely suitable for the description of the within-host MYXV growth dynamics.

4.1.2. Model candidates

To grasp the mechanisms of the MYXV growth dynamics in time (t [*days*]), four models were considered as candidates to fit the within-host dynamics of the virus, expressed as V_t and measured in suitable units.

4.1.2.1. Malthusian model

The first evaluated option is a Malthusian model. For many of the analyzed strains the growth of the pathogen is extremely rapid and uncontrolled by the host, therefore a model without a limiting capacity can be considered to fit the evolution of virus load in the examined interval of the disease [37]:

$$\frac{dV}{dt} = rV \qquad (4.2)$$

with the log-form explicit solution as below. The initial condition V_0 and the growth rate r are the parameters to be optimized (V_0>0 [[V]], r>0 [$days^{-1}$]).

$$log_{10}(V_t) = log_{10}(V_0 e^{rt}) \qquad (4.3)$$

Although at the start of the experiment the body of the host is still clean of the pathogen agent, considering DPI=0 at the administration of the doses, it's possible to identify the initial condition V_0 as the equivalent assimilation of those doses by the body. The capability of the model to weigh the impact of the initial condition is very relevant to analyze the 'Kerr dataset', which has no data before the fifth day after infection.

4.1.2.2. Gompertz model

The second considered option is a Gompertz growth model. This function is suited to catch the fast initial growth that proceeds towards a capacity [38]:

$$\frac{dV}{dt} = -rVln\left(\frac{V}{K}\right) \qquad (4.4)$$

with the log-form explicit solution as below. The initial condition V_0 (i.e. the resulting assimilation of the infection doses by the body), the initial growth rate r and the

population capacity K and are the parameters to be optimized (V_0>0 [[V]], r>0 [$days^{-1}$], K>0 [[V]]).

$$log_{10}(V_t) = (1 - e^{-rt})log_{10}(K) + e^{-rt}log_{10}(V_0) \quad (4.5)$$

4.1.2.3. Logistic model

The third evaluated option is a logistic growth model. This option limits again the growth of the pathogen considering a carrying capacity due to factors like intraspecific competition [39]:

$$\frac{dV}{dt} = rV\left(1 - \frac{V}{K}\right) \quad (4.6)$$

with the log-form explicit solution as below. The initial condition V_0 (i.e. the resulting assimilation of the infection doses by the body), the initial growth rate r and the carrying capacity K are the parameters to be optimized (V_0>0 [[V]], r>0 [$days^{-1}$], K>0 [[V]]).

$$log_{10}(V_t) = log_{10}\left(\frac{KV_0e^{rt}}{K - V_0 + V_0e^{rt}}\right) \quad (4.7)$$

4.1.2.4. Model with immune response

To include the effect of the immune response, the fourth model option was conceptualized as a differential system with one equation for the virus and one for the effect of the immune system. In this system, the growth of the virus is counteracted by a viral decrease induced by a predation effect of the immune response [18]. No data about the immune system is available in the datasets. In absence of data, the intervention of the immune response was introduced in the model through the variable I_t [$days^{-1}$], which represents a killing rate of the virus. This killing rate varies over time as a response to a dynamic MYXV load.
The virus load V_t grows logistically (with parameters V_0, r, d considered equal to $\frac{r}{K}$) and is contrasted by the killing rate through a type-II functional response [40] modulated by parameter b. The dynamic virus killing rate I_t ([$days^{-1}$]) grows with a hyperbolic response to the viral load with parameters a and c, possibly autocatalytic (according to the shape parameter e, which takes on the values 0 or 1) and contrasted by a natural decrease of the immune system mediated by a decay parameter u; the initial condition I_0 was set to 0 [$days^{-1}$]. The model structure results therefore as follows.

$$\begin{cases} \frac{dV}{dt} = rV - dV^2 - \frac{VI}{1+bV} \\ \frac{dI}{dt} = \frac{aV}{1+cV} I^e - uI \end{cases} \quad (4.8)$$

This complete model structure was investigated to identify the combination of parameters to maintain or remove to find the best trade-off between complexity and accuracy. The values that the parameters can assume are reported in Table 4-1.

Table 4-1: Parameter ranges of the within-host MYXV growth model with immune response, for the AIC model selection of strains Uriarra and Neuromyxoma.

Parameter			Unit
$r > 0$			$[days^{-1}]$
$V_0 > 0$			$[\,[V]\,]$
$d = 0$	or	$d > 0$	$[(days[V])^{-1}]$
$b = 0$	or	$b > 0$	$[[V]^{-1}]$
$a = 0$	or	$a > 0$	$\begin{cases} [1/days^2[V] \ if \ e = 0 \\ [\,(days[V])^{-1} \ if \ e = 1 \end{cases}$
$c = 0$	or	$c > 0$	$[[V]^{-1}]$
$e = 0$	or	$e = 1$	$[\,/\,]$
$u = 0$	or	$u > 0$	$[days^{-1}]$

4.1.3. Methods

The selection and optimization of the best model among the candidates was based on a trade-off between complexity and accuracy through the Akaike Information Criterion (AIC) [34]. It was assumed that the virus load data V_i is affected by uncorrelated lognormal noise ε_i, i.e. $V_i = \tilde{V}_i \, \varepsilon_i$ with ln (ε_i) being normally distributed with 0 mean and unknown variance σ^2, or in other words that $log_{10}(V_i)$ is affected by normal noise. Let $\tilde{V}_i$ be a function of time and be defined by a vector θ of h unknown parameters, therefore ln $(V_i/f(i;\theta)) =$ ln (ε_i). The likelihood function is structured as follows, where n is the number of data,

$$L(V;\theta,\sigma) = \prod_{i=1}^{n}\left[(2\pi\sigma^2)^{-\frac{1}{2}} V_i^{-1} \exp\left(-\frac{\left(\ln\left(\frac{V_i}{f(i;\theta)}\right)\right)^2}{2\sigma^2}\right)\right] \quad (4.9)$$

and the loglikelihood is thus expressed as

$$\mathcal{L}(V;\theta,\sigma) = -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n}\ln(V_i) - \frac{\sum_{i=1}^{n}\left(\ln\left(\frac{V_i}{f(i;\theta)}\right)\right)^2}{2\sigma^2} \quad (4.10)$$

This loglikelihood function is maximized by finding the estimated value $\hat{\sigma}^2$ and $\hat{\theta}$ that maximize the loglikelihood. For any given $\hat{\theta}$ the optimal value of the variance σ^2 is provided by the following formula.

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(\ln\left(\frac{V_i}{f(i;\hat{\theta})}\right)\right)^2}{n} \quad (4.11)$$

Therefore, the resulting loglikelihood is structured as

$$\mathcal{L}\left(V;\hat{\theta},\hat{\sigma}\right) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\hat{\sigma}^2) - \sum_{i=1}^{n}\ln(V_i) - \frac{n\hat{\sigma}^2}{2\hat{\sigma}^2} = const - \frac{n}{2}\ln(\hat{\sigma}^2) \quad (4.12)$$

The optimal $\hat{\theta}$ is thus the one that minimizes $\hat{\sigma}^2$, namely the squared error of the fit of the model to the data. By discarding the constant and multiplying by 2, the Akaike Information Criterion score is obtained as follows [41], where h is the number of parameters.

$$AIC = 2h + n\ln\left(\frac{\sum_{i=1}^{n}\left(ln\left(\frac{V_i}{f(i;\hat{\theta})}\right)\right)^2}{n}\right) \quad (4.13)$$

The model is now ready for AIC inspection as summarized in Table 4-2.

Table 4-2: Model selection structure for MYXV within-host dynamics.

Variable $x_i = V_i$
Parameters $\Theta = (V_0\, r, K, d, b, a, c, e, u,);\ \ h = count(\Theta)$
Objective function $J(\Theta) = AIC = 2h + n\ln(\frac{\sum_{i=1}^{n}\left(\ln(\frac{x_{i,obs}}{x_{i,sim}})\right)^2}{n})$
Optimization φ^* such that $J(\varphi^*) = J^* = \min_{\varphi}(J(\Theta))\ ;\ \varphi \in \Theta$

The results also report the outcomes of the selection considering an additional 'parsimony criterion', for which models whose AIC scores do not differ more than 2 are considered equivalent and between them the model with fewer parameters is suggested as optimal option. The results identified with this approach describe the system with further reduced complexity.
The uncertainty of the calibration process was investigated by applying a bootstrap approach on the selected model for each strain. The data of each strain was randomly resampled 100 times and the calibration was repeated for each obtained series of data. With the obtained parameters the MYXV growth was simulated 100 times and the 5% - 95% confidence interval was identified.

4.1.4. Results

The results of the MYXV growth model selection are reported considering two cases: the strains with an increasing trend in the entire experiment interval and the strains that show a viral decrease.

For all the strains with an increasing trend, the outcomes of the selection are summarized in Table 4-3. As regards the model with immune response, just the case [V_0>0, r>0, d=0, b=0, a>0, c=0, e=0, u=0] reported below is made explicit, while all the other combinations of parameters are discussed later.

$$\begin{cases} \frac{dV}{dt} = rV - VI \\ \frac{dI}{dt} = aV \end{cases} \qquad (4.14)$$

The AIC scores are shown together with the coefficients of determination R^2 of the optimized models, formulated as follows, to provide an idea of the goodness of the fit.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(log_{10}(x_{i,obs}) - log_{10}(x_{i,sim}))^2}{\sum_{i=1}^{n}(log_{10}(x_{i,obs}) - log_{10}(x_{mean}))^2} \qquad (4.15)$$

Table 4-3: Results of MYXV growth model selection of all the strains, apart from Uriarra and Neuromyxoma. The model selected by the lowest AIC score is marked in bold while the model selected by the parsimony criterion is underlined.

Strain	Dataset	Virulence grade	AIC Malthus model	R^2 Malthus model	AIC Gompertz model	R^2 Gompertz model	AIC logistic model	R^2 logistic model	AIC immune model (4.14)	R^2 immune model (4.14)
Swh 8/2/93	Kerr	2	**-21.300**	0.829	-20.138	0.839	-20.248	0.840	-20.251	0.840
Ws6 346	Kerr	2	**-13.854**	0.525	-11.907	0.527	-11.918	0.528	-11.930	0.528
Bd23	Kerr	1-2	-5.698	0.324	-7.123	0.480	**-7.181**	0.483	-5.481	0.411
Swh 1209	Kerr	3	**-33.937**	0.824	-32.602	0.826	-32.761	0.831	-32.784	0.832
Brk 12/2/93	Kerr	1-2	-14.754	0.452	**-16.195**	0.565	-15.897	0.562	-14.865	0.529
Km13	Kerr	5	-14.235	0.554	-14.026	0.586	-14.301	0.591	**-14.380**	0.592
Sls	Fenner	1	-4.170	0.666	-8.020	0.780	**-8.420**	0.786	-7.284	0.768
Km13 II	Fenner	3	**8.527**	0.656	9.502	0.687	9.298	0.693	9.231	0.694
Uriarra III	Fenner	3-4	2.039	0.836	-6.168	0.930	**-6.383**	0.931	-5.289	0.925
Loiret 55	Fenner	4	**15.024**	0.669	15.506	0.708	15.429	0.710	15.413	0711
Lausanne	Fenner	1	7.046	0.644	1.155	0.838	**-1.163**	0.872	4.348	0.777

Looking at the coefficients of determination, all the available models fit the data of most strains quite well and they are confirmed as valid model options to describe the MYXV dynamics. In particular, the logistic models return very good performances in terms of R^2, slightly better than the ones of the Gompertz solutions. Since the Gompertz model shows the lowest AIC score just in one case (strain *Brk 12/2/93*), this model option was discarded. However, considering the balance between accuracy and complexity provided by the Akaike Information Criterion, the Malthusian model returns the lowest score for many strains (*Swh 8/2/93*, *Ws6 346*, *Swh 1209*, *Km13 II* and *Loiret 55*). In these cases, the MYXV viral load grows in the host in a fast and uncontrolled way and the carrying capacity is not clearly perceived before the fatalities. In other cases, instead, a carrying capacity of the viral load is more appreciable, and the logistic model turns out to be the best option also according to the Akaike Information Criterion (strains *Bd23*, *Sls*, *Uriarra III*, *Lausanne*, and *Brk 12/2/93* no longer considering the Gompertz option). For the strains *Bd23* and *Brk 12/2/93*, however, the parsimony criterion identifies that the Malthusian model can be considered as equivalent ($\Delta\text{AIC} < 2$) and selected to reduce the complexity.

As far as the model with immune response is concerned, option (4. 14) is the only combination of parameters that results significant for these strains. In fact, the optimization process set the other parameters as substantially irrelevant, meaning that the AIC score of other combinations of parameters deviates from the option [$V_0>0$, $r>0$, $d=0$, $b=0$, $a>0$, $c=0$, $e=0$, $u=0$] or from a logistic model (coinciding with [$V_0>0$, $r>0$, $d>0$, $b=0$, $a=0$, $c=0$, $e=0$, $u=0$]) by approximately 2 times the extra number of parameters ($+2\Delta h$). This model option returns the lowest AIC score only for one strain, *Km13*, and can be considered equivalent as the simpler Malthusian option ($\Delta\text{AIC} < 2$) applying the parsimony criterion. Therefore, without data to represent the immune system and in absence of a viral decrease the model selection does not identify the inclusion of I_t as essential. It is however possible, in some cases, that the effect of the immune response may really be almost negligible, due to the identified immunosuppressive characteristics of MYXV.

The optimized parameters of the evaluated models for all the considered strains are reported in Table 4-4.

Table 4-4: Optimized parameters of the selected models for all the strains, apart from Uriarra and Neuromyxoma. The model selected by the lowest AIC score is marked in bold, while the model selected by the parsimony criterion is underlined.

Strain	Dataset	Virulence grade	Malthus		logistic			Model with immune response (4. 14) [V₀>0, r>0, d=0, b=0, a>0, c=0, e=0, u=0]		
			V_0 [IV]	r $[\frac{1}{days}]$	V_0 [IV]	r $[\frac{1}{days}]$	K [IV]	V_0 [IV]	r $[\frac{1}{days}]$	a $[\frac{1}{days^2[V]}]$
Swh 8/2/93	Kerr	2	**1.154E+07**	**0.230**	-	-	-	-	-	-
Ws6 346	Kerr	2	**3.428E+07**	**0.133**	-	-	-	-	-	--
Bd23	Kerr	1-2	4.084E+07	0.113	**1.949E+06**	**0.679**	**2.035E+08**	-	-	-
Swh 1209	Kerr	3	**4.445E+07**	**0.125**	-	-	-	-	-	-
Brk 12/2/93	Kerr	1-2	3.742E+07	0.112	**3.266E+06**	**0.582**	**1.827E+08**	-	-	-
Km13	Kerr	5	2.767E+07	0.101	-	-	-	**1.523E+07**	**0.175**	**5.178E-11**
Sls	Fenner	1	-	-	**1.633E+05**	**1.118**	**3.924E+07**	-	-	-
Km13 II	Fenner	3	**3.906E+05**	**0.504**	-	-	-	-	-	-
Uriarra III	Fenner	3-4	-	-	**2.781E+05**	**0.836**	**6.854E+08**	-	-	-
Loiret 55	Fenner	4	**1.510E+05**	**0.566**	-	-	-	-	-	-
Lausanne	Fenner	1	-	-	**149.480**	**2.903**	**9.325E+07**	-	-	-

Generally, the values of the growth rate r of the strains provided by the 'Fenner dataset' result higher than the ones of the strains provided by the 'Kerr dataset'. This effect is mainly due to two aspects: the different viral unit of measure adopted in the two datasets and the temporal location of the first available data. The 'Fenner dataset' viral data are collected from the second day post infection, while the 'Kerr dataset' starts from DPI=5. This lack of viral data in the first few days tends to increase the initial value of assimilation of the doses V_0 and reduce the growth rate r. For the strains contained in the 'Fenner dataset', the availability of the entire temporal spectrum of data results in a larger performance gap between the Malthusian models and the logistic models, which better describe the curve and partially show a pattern of parameter values that allows to distinguish the virulence of the strains (high virulent strains are generally associated to high r and low K).
It is possible to observe that the values of r and V_0 of *Lausanne* are quite different from those of the other strains, mainly due to the role of the data at DPI=2. A test was performed to verify that the behavior of this strain does not differ much from the others, by requiring the initial condition of the strain to remain aligned to the general optimized values (for example V_0>1E+05 [V]). As expected, the new optimization still returned a good fit (V_0=1E+05 [V], r=1.119 [$days^{-1}$], K=1.240E+08 [V], R^2=0.779).

A different situation regards the rabbits infected with *Uriarra* and *Neuromyxoma*, for which the immune system controls the growth of the virus generating a viral decrease. The results of the AIC selection for these strains are listed in Table 4-5, highlighting only some relevant cases of the model with immune response. The selection of the immune model option revealed that all the combinations of parameters including e=1 led to unacceptable shapes of the curves. In the table all the combinations of d, b and c are reported with e=0 and u>0, while the options with u=0 are reported just the relevant combinations.

Table 4-5: Results of MYXV growth model selection for strains Uriarra and Neuromyxoma. The model selected by the lowest AIC score is marked in bold while the model selected by the parsimony criterion is underlined.

Model option	Number of parameters	Uriarra Kerr dataset (grade 5)		Neuromyxoma Fenner dataset (grade 5)	
		AIC	R^2	AIC	R^2
Malthusian model	2	52.436	0.000	51.007	0.000
Logistic model	3	52.564	0.058	44.045	0.301
Gompertz model	3	51.844	0.083	25.479	0.668
Model with immune response [V_0>0, r>0, d=0, b=0, a>0, c=0, e=0, u>0]	4	52.414	0.126	22.379	0.729
Model with immune response [V_0>0, r>0, d=0, b=0, a>0, c>0, e=0, u>0]	5	54.175	0.132	20.811	0.765
Model with immune response [V_0>0, r>0, d=0, b>0, a>0, c=0, e=0, u>0]	5	53.948	0.139	24.379	0.729
Model with immune response [V_0>0, r>0, d=0, b>0, a>0, c>0, e=0, u>0]	6	55.908	0.140	22.460	0.768
Model with immune response [V_0>0, r>0, d>0, b=0, a>0, c=0, e=0, u>0]	5	54.414	0.126	19.963	0.773
Model with immune response [V_0>0, r>0, d>0, b=0, a>0, c>0, e=0, u>0]	6	56.175	0.132	21.898	0.773
Model with immune response [V_0>0, r>0, d>0, b>0, a>0, c=0, e=0, u>0]	6	55.948	0.139	21.963	0.773
Model with immune response [V_0>0, r>0, d>0, b>0, a>0, c>0, e=0, u>0]	7	57.908	0.140	23.963	0.773
Model with immune response [V_0>0, r>0, d>0, b=0, a>0, c=0, e=0, u=0]	4	51.948	0.139	**17.989**	0.773
Model with immune response [V_0>0, r>0, d=0, b=0, a>0, c=0, e=0, u=0]	3	**50.414**	0.126	21.367	0.718

The within-host growth of these strains is best described by one of the combinations of the model with immune response. As shown, the AIC selection pushes u to be

turned off to zero. Without data to fit I_t, the decay mechanism of the virus killing rate cannot be identified and it is considered negligible. It is however plausible that the immune activity starts to fade away only after the virus has substantially disappeared from the body. The selection of a simple model structure with e=0 and u=0 implies therefore that the killing rate I_t is proportional to the integral of the viral load and it is not appreciably inactivated, or modulated (b=0, c=0). The optimization of *Uriarra* identifies the option with d=0 as the best, producing the only difference from *Neuromyxoma*. The two optimized solutions show therefore a Malthusian and a logistic growth (considering $d = \frac{r}{K}$) as the other strains, with the immune component modelled with only one extra parameter. The obtained models result as follows, with the optimized parameters shown in Table 4-6.

$$\begin{cases} \frac{dV}{dt} = rV - VI \\ \frac{dI}{dt} = aV \end{cases} \quad (4.16)$$

$$\begin{cases} \frac{dV}{dt} = rV - dV^2 - VI \\ \frac{dI}{dt} = aV \end{cases} = \begin{cases} \frac{dV}{dt} = rV(1 - \frac{V}{K}) - VI \\ \frac{dI}{dt} = aV \end{cases} \quad (4.17)$$

Table 4-6: Optimized parameters of the viral growth model of Uriarra and Neuromyxoma strains. The selections by lowest AIC score and by parsimony criterion coincide.

Strain	Model (4. 16) [V_0>0, r>0, d=0, b=0, a>0, c=0, e=0, u=0]			Model (4. 17) [V_0>0, r>0, d>0, b=0, a>0, c=0, e=0, u=0]			
	V_0 [[V]]	r [$\frac{1}{days}$]	a [$\frac{1}{days^2[V]}$]	V_0 [[V]]	r [$\frac{1}{days}$]	d [$\frac{1}{days[V]}$]	a [$\frac{1}{days^2[V]}$]
Uriarra Kerr dataset (grade 5)	**1.665E+07**	**0.222**	**3.468E-10**	-	-	-	-
Neuromyxoma Fenner dataset (grade 5)	-	-	-	**1.565E+02**	**3.308**	**1.708E-06**	**7.858E-0.7**

The optimization of *Neuromyxoma* identifies a low value of V_0 and a high value of r. As it was done with strain *Lausanne*, a test was performed by requiring the initial condition of the strain to remain aligned to the general optimized values (for example V_0>1E+05). Also in this case, the new optimization still returned a good fit (V_0=1E+05 [V], r=0.947 [$days^{-1}$], d=4.888E-07 [$([V]days)^{-1}$], a=3.212E-07 [$([V]days^2)^{-1}$], R^2=0.734).
The immune response results more effective on the strain *Neuromyxoma* than on

Uriarra, despite both being grade 5 virulence. Grade 5 includes in fact a wide range of MYXV strains, from medium-low to null fatality rate. The strain *Uriarra* shows a great variability rabbit by rabbit, especially in the last days of the experiment, with two rabbits not yet in a clear decreasing viral phase, two with already low viral loads, and two in an intermediate situation. This variability is reflected in a low coefficient of determination R^2 as the model describes an average behavior between specific cases. In addition, even among the models with an increasing viral growth there was a strain with grade 5 virulence, namely *Km13*. Although characterized by a low rate of fatality (all the animals were still alive at the end of the experiment), a clear decrease in viral loads was not yet observed until DPI=27. Interestingly though, despite this difference, *Km13* is the only other strain for which the model with immune response returned the lowest AIC score, relating this case with *Uriarra* and *Neuromyxoma*.

The fit and confidence interval of the MYXV growth model with the lowest AIC score for each strain of the 'Kerr dataset' and the 'Fenner dataset' are respectively represented in Figure 4-4 and Figure 4-5. Each rabbit is marked with a different color.

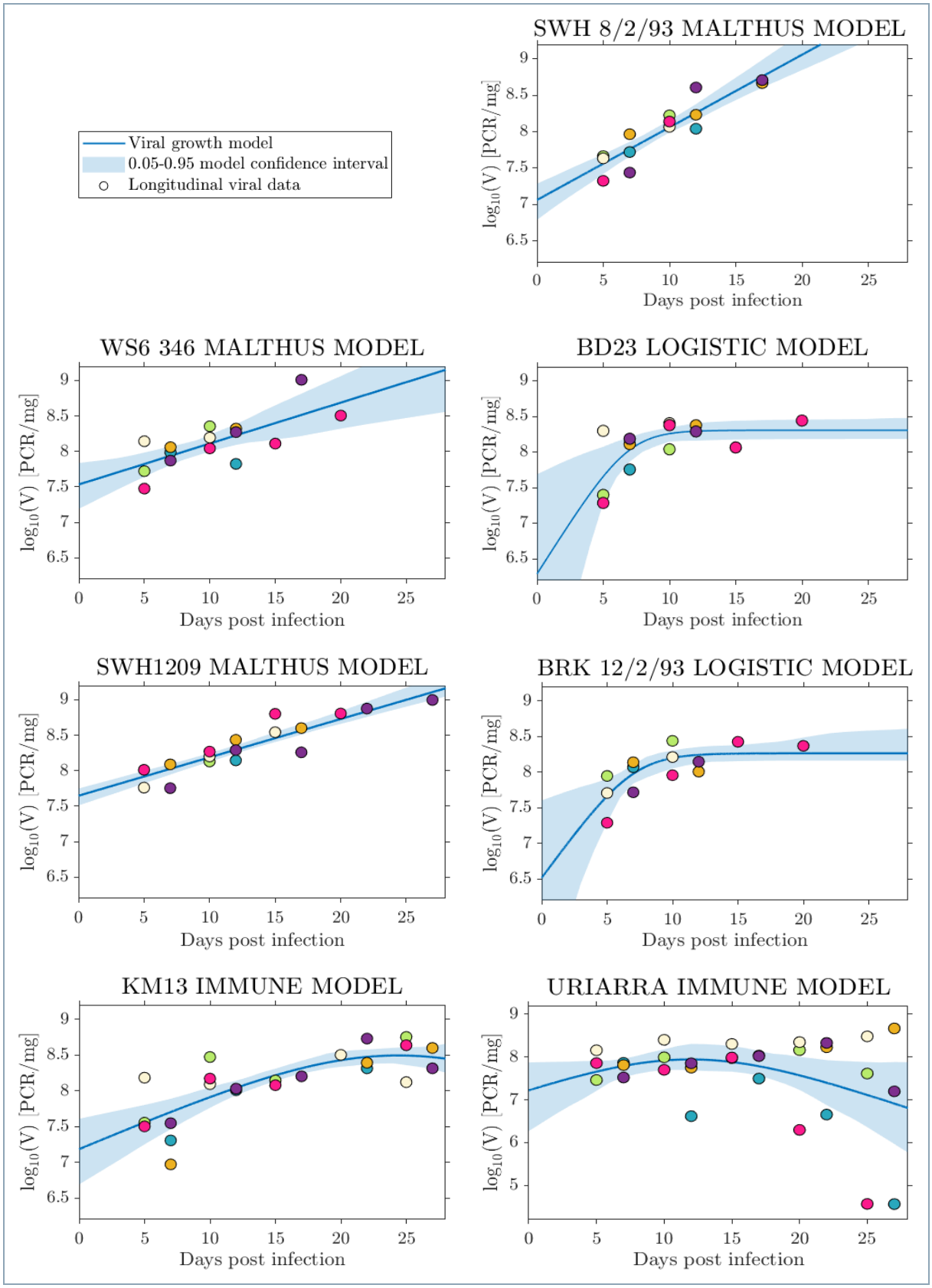


Figure 4-4: Fit and confidence interval of the MYXV growth model with the lowest AIC score for each strain of the 'Kerr dataset'. Each rabbit is marked with a different color.

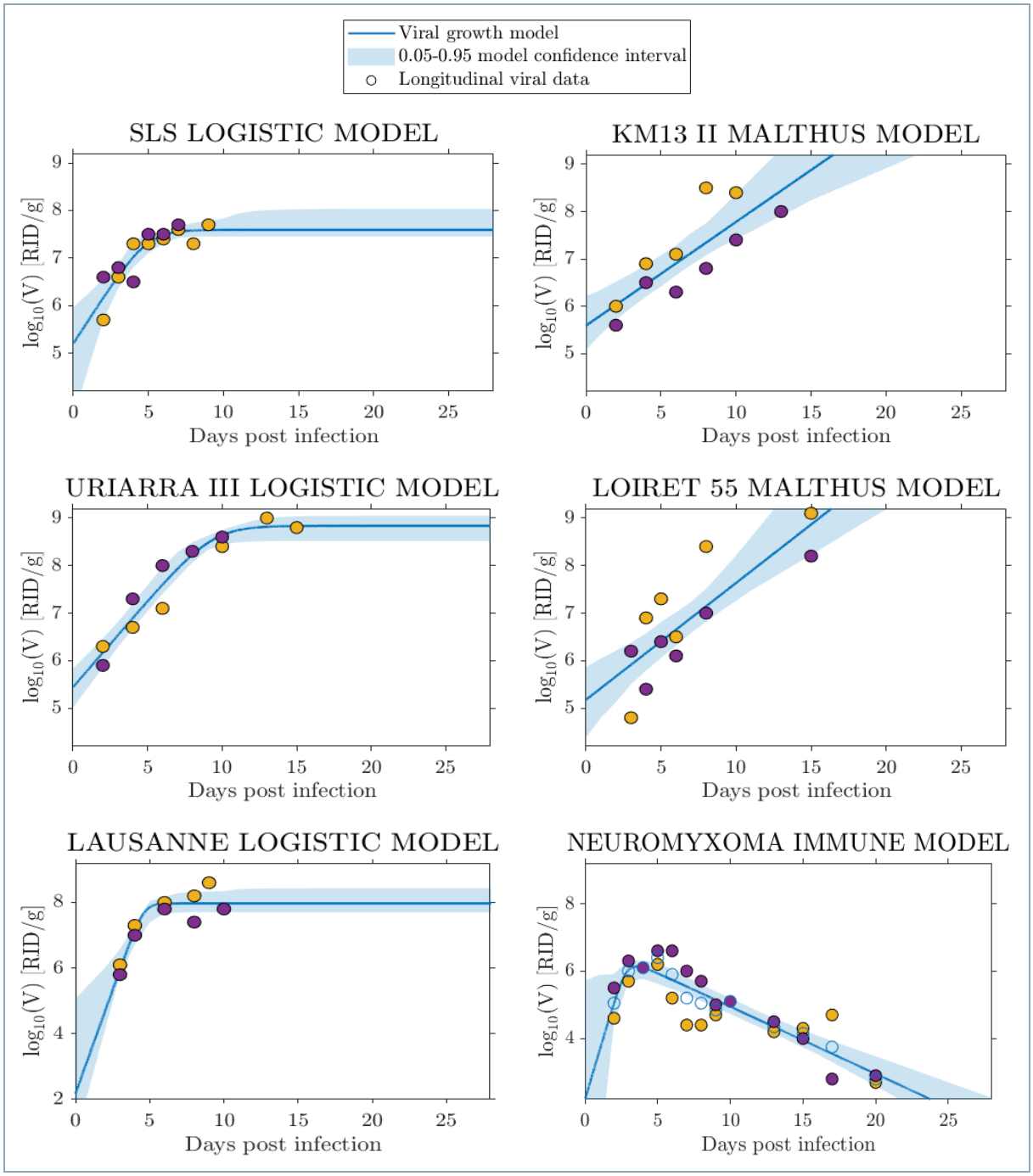


Figure 4-5: Fit and confidence interval of the MYXV growth model with the lowest AIC score for each strain of the 'Fenner dataset'. Each rabbit is marked with a different color.

Some of the models, especially for the strains provided by the 'Kerr dataset', show great uncertainty in the first DPI, where calibration data are missing. The variability expressed with the bootstrap process then narrows at intermediate DPI and does not include the entire cloud of data in the confidence interval. The intervals tend to widen

again and show greater uncertainty after the last DPI with data, especially for the simple Malthusian cases. The confidence intervals of the parameters are reported in Appendix A.

4.1.5. Comments

Summarizing, the model selection identifies different solutions depending on the capability of the immune response to control and overcome virus growth. The selected models provide a sufficiently complete framework to describe the dynamics of MYXV inside the host after infection. The Malthusian model can describe the initial growth phase of almost all the strains and fits well those cases where the viral growth continues to increase with exponential rapidity until the death of the host. The logistic model, at the cost of one extra parameter, returns very good performances and is suitable for those cases where there's sufficient time to observe the viral load slowing down towards carrying capacity before the death of the host. Finally, the model with immune response makes the role of the immune system explicit by adding the contribution of a virus killing rate to the Malthusian or logistic growth. This option is useful to describe all the cases where the host response manages to control the infection and eradicate the virus after an initial growth.
These models allow to predict the viral load over time and at the time of death for a wide range of possible strains. The optimized values of the growth parameters cannot be used alone to distinguish the virulence of the strains, but it is necessary to model also the severity of the disease to understand the effects that cause different average survival times.

4.2. Modelling of mortality rate and survival probability

In subchapter 4.1 the within-host MYXV growth was analyzed by developing a model to describe its evolution in the days post infection (DPI). The goal of this subchapter is to complete the characterization of the infection dynamics of different MYXV strains, modelling the survival probability of the hosts. It is therefore necessary to link the information provided by the within-host viral growth and by the survival times of the hosts to develop a complete survival model. The analysis focuses on strains from grade 1 to 3. Not enough survival data were available for grade 4, while all the animals infected with strains of grade 5 remained alive for the entire period of the experiments. The survival probability in time p_t of a host infected with MYXV can be derived from the mortality rate $\mu(V_t)$, function of the viral load V_t, as expressed as follows; it was assumed that the host mortality rate increases with the viral load, as suggested by data.

$$p_t = e^{-\int_0^t \mu(V(z))dz} \quad (4.18)$$

Four model options were compared for each strain, to describe the mortality rate as a function of the viral load estimated by the viral growth model of subchapter 4.1 that showed the lowest AIC score for the strain. The selection and optimization of the best survival model was based on the accuracy of the fit of the estimated survival probabilities on the curves built with host survival time data.
As regards the two datasets used in the previous subchapter, the 'Kerr dataset' contains viral growth data and host survival times for each strain (except *Brk 12/2/93*, for which the survival times are missing), while the viral growth data of the strains in the 'Fenner dataset' lack the information of the host survival times. Instead, the 'survival dataset' provides only host survival times for a series of extra strains. The survival analysis can therefore be directly performed for the strains of the 'Kerr dataset', while the strains of the 'Fenner dataset' need a preliminary step. The 'survival dataset' and the 'Kerr dataset' were rearranged by mixing all the data of strains with the same virulence and building overall survival curves for grades 1, 2 and 3. The survival probabilities of *Brk 12/2/93* and of the strains provided by the 'Fenner dataset' were fitted on those combined curves, made with data of three or more strains each. All the available survival curves are reported in Figure 4-6.

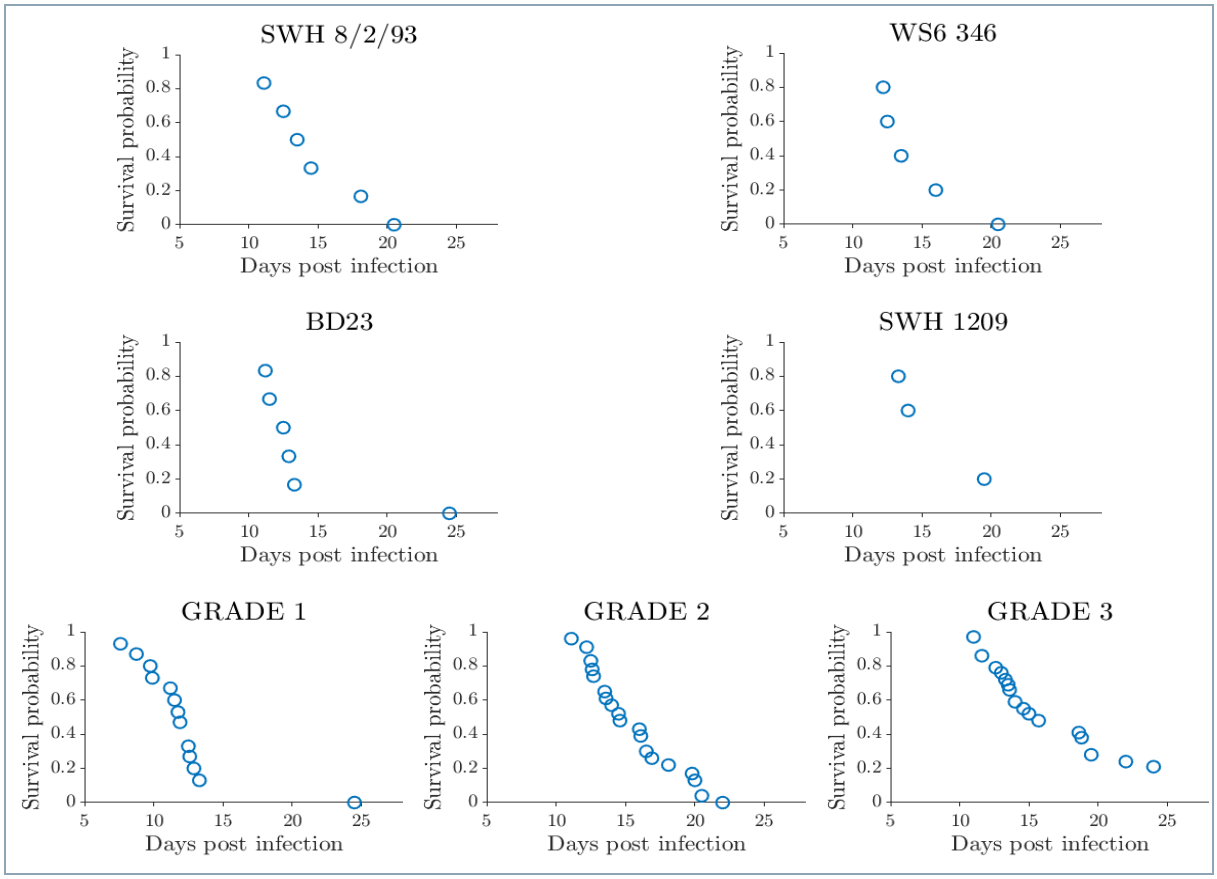


Figure 4-6: Survival curves built with survival times of hosts infected by strains from 'Kerr dataset' or from 'Kerr dataset' and 'survival dataset' aggregated by virulence grade.

The number of survival time data of some strains was not suitable to efficiently inspect the variability of the process with traditional methods as bootstrap [36] or jackknife [42]. Simulations of 50 infected rabbits were therefore repeated 100 times to build confidence intervals of the models, assigning to each rabbit a DDPI (death day post infection) coinciding with the temporal step when the modelled probability of survival falls below a randomly generated number. Thanks to the newly obtained artificial data, it was possible to build 5%-95% confidence intervals of the survival probability to complete the characterization of the strains.

4.2.1. Model candidates

Four options were considered to describe the mortality rate in time ($\mu_t[days^{-1}]$), two based on a sigmoidal solution and two based on a viral threshold. These model structures are designed to mimic the decreasing trend of the survival probabilities.

4.2.1.1. Sigmoidal model with fixed exponent

The first option is a model based on a sigmoidal growth of the mortality rate, able to describe an inverse S-shape of the related survival probability. The model has two parameters, δ and γ, while the exponent is quadratic and fixed ($\delta[days^{-1}]>0$, $\gamma[[log_{10}(V)]]>0$):

$$\mu(V_t) = \frac{\delta(\log(V_t))^2}{\gamma^2+(\log(V_t))^2} \quad (4.19)$$

4.2.1.2. Sigmoidal model with optimizable exponent

The second option improves the first sigmoidal-based mortality rate model by optimizing the exponent τ in addition to δ and γ, as in the following equation, trying to adhere more to the shape of the curve at the cost of one extra parameter ($\delta[days^{-1}]>0$, $\gamma[[log_{10}(V)]^{\tau/2}]>0$, $\tau[/]>0$).

$$\mu(V_t) = \frac{\delta(\log(V_t))^\tau}{\gamma^2+(\log(V_t))^\tau} \quad (4.20)$$

4.2.1.3. Threshold model with proportional mortality (delta)

The third model candidate is based on a viral threshold, as shown below. In the experiments rabbit didn't die in the first week after infection, suggesting that MYXV could start to be fatal over a certain viral load. For this reason, the two-parameters model assumes a viral threshold *thr* below which the mortality is null; after that, the

death rate increases proportionally (through β) to the delta from that value (thr[[$log_{10}(V)$]]>0, β[$([V]days)^{-1}$]>0).

$$\begin{cases} if\ log_{10}(V_t) < thr: \ \mu(V_t) = 0 \\ otherwise: \ \mu(V_t) = \beta(log_{10}(V_t) - thr) \end{cases} \quad (4.21)$$

4.2.1.4. Threshold model with on-off mortality (switch)

The fourth model option, as reported below, in based again on a viral threshold thr, after which the mortality rate switches from 0 to a constant value ω. The sudden and fast decrease of the survival probabilities suggests that the option of a simple switch function could be suitable to describe the mortality rate (thr[[$log_{10}(V)$]]>0, ω[$days^{-1}$]>0)

$$\begin{cases} if\ log_{10}(V_t) < thr: \ \mu(V_t) = 0 \\ otherwise: \ \mu(V_t) = \omega \end{cases} \quad (4.22)$$

4.2.2. Methods

The optimization of the parameters was realized fitting p_t on the available survival curves. The optimized models were then compared by applying the Akaike Information Criterion model selection (Table 4-7).

Table 4-7: Model selection structure for survival probabilities.

Variable X= $p(\mu(V))$
Parameters $\Theta = (\delta, \gamma, \tau, thr, \beta, \omega)$; $h = count(\Theta)$
Objective function $J(\Theta) = AIC = 2h + n\ln(\frac{\sum_{i=1}^{n}(X_{i,obs} - X_{i,sim})^2}{n})$
Optimization φ^* such that $J(\varphi^*) = J^* = \min_{\varphi}(J(\Theta))$; $\varphi \in \Theta$

Once the structure was selected and optimized, the model was used to perform simulations to build survival confidence intervals. For 100 repetitions, 50 uniformly distributed random real numbers in the interval [0 1] were generated, each representing one rabbit. A death time was associated to each of these simulated hosts, coinciding with the time when the modelled survival probability $p(t)$ becomes lower than the random extracted value. Consequently, a viral load at death was also assigned to each rabbit, equal to the MYXV amount provided by the corresponding viral growth model. The distribution of the survival probability in time was calculated from the

cumulated temporal distribution of dead hosts. From the results of the 100 repetitions of this process it was possible to obtain the 5%-95% confidence intervals of survival probabilities.

4.2.3. Results

The results of the survival model selection are reported in Table 4-8 and displayed in Figure 4-7 for the 'Kerr dataset' and in Figure 4-8 for the 'Fenner dataset'.

Table 4-8: AIC results for survival model selection, with lowest AIC score marked in bold.

Strain	Virulence grade	Viral growth model	AIC surv. model (4. 19)	AIC surv. model (4. 20)	AIC surv. model (4. 21)	AIC surv. model (4. 22)
Swh 8/2/93	2	Malthusian	-15.566	-26.711	**-34.621**	-32.661
Ws6 346	2	Malthusian	-11.559	-19.351	-23.567	**-27.009**
Bd23	1-2	Logistic	-14.825	-21.150	**-31.453**	-30.492
Swh 1209	4	Malthusian	-6.060	-10.629	-13.550	**-15.792**
Brk 12/2/93	1-2	Logistic	-55.138	-97.961	**-120.218**	-119.712
Sls	1	Logistic	-37.719	-49.231	**-59.544**	-58.200
Km13 II	3	Malthusian	-64.343	-77.735	-87.375	**-105.283**
Uriarra III	3-4	Logistic	-60.679	-98.245	-102.170	**-105.283**
Lausanne	1	Logistic	-38.818	-42.214	**-58.360**	-58.200

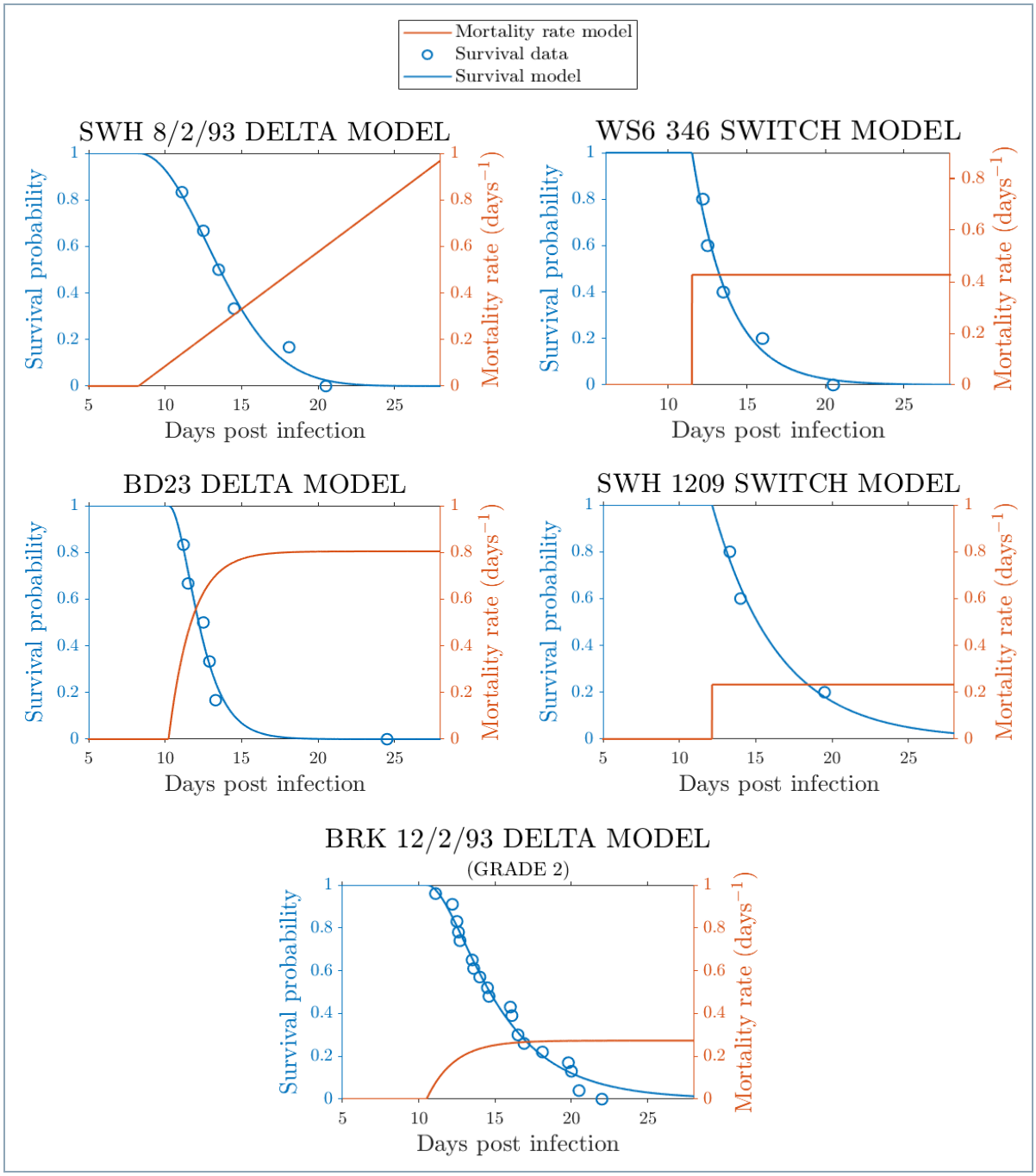


Figure 4-7: Survival fit of the optimized selected models for the strain of Kerr dataset. Where the grade is specified, the survival curve is obtained from the aggregation of other strains.

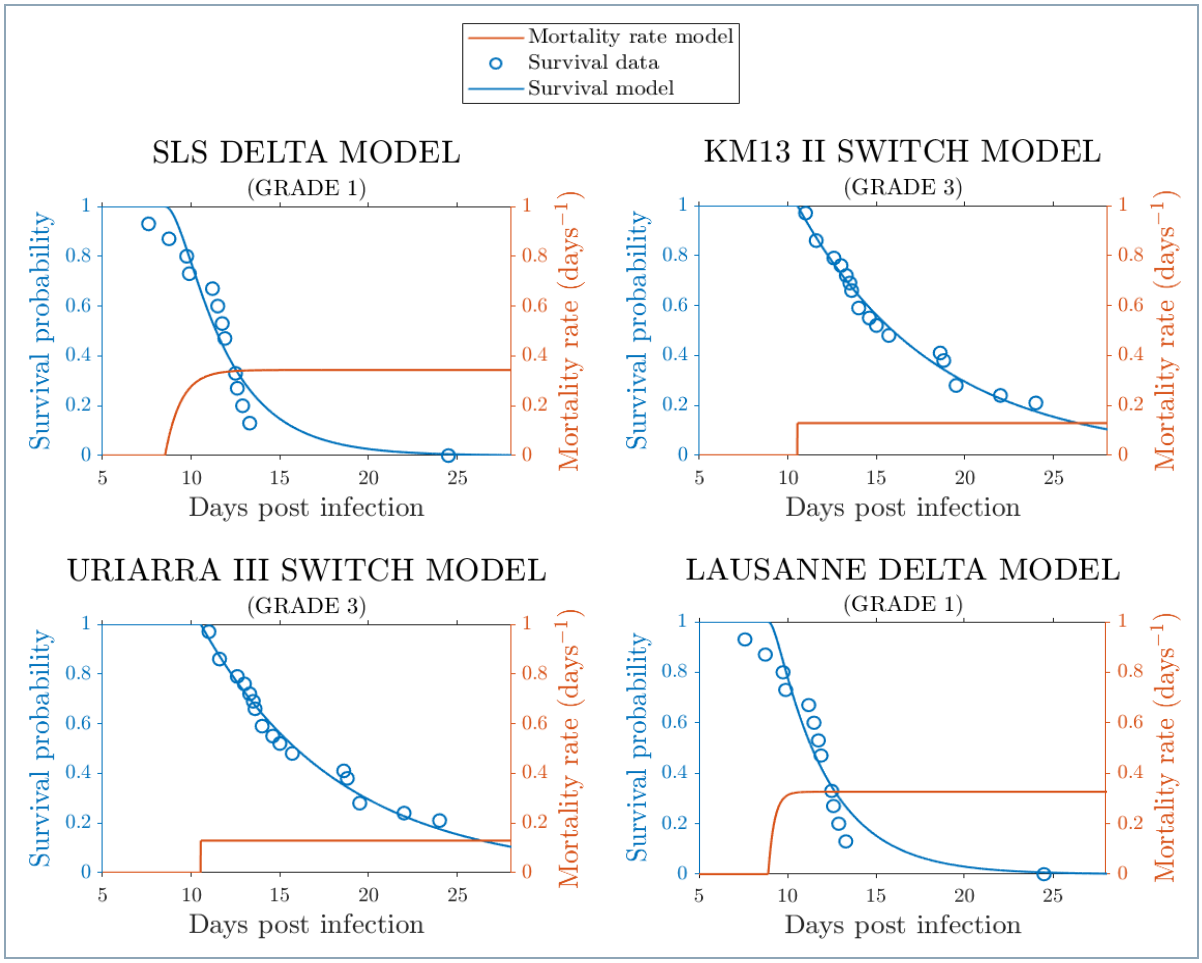


Figure 4-8: Survival fit of the optimized selected models for the strain of Fenner dataset. The survival curve are obtained from the aggregation of other strains.

The sigmoidal model option (4. 19) provides low-level performances in describing the evolution of the mortality rate. The fixed quadratic exponent considered by the model is not sufficient to follow the sudden and rapid decrease of the survival probabilities. In fact, the optimization process of the sigmoidal model option (4. 20), which also consider the exponent as a parameter (τ), tends to assign extremely high values to τ to portray the characteristics of the curves. However, even leaving the parameter τ free to assume very high values, the sigmoidal-based models perform worse than the ones based on a threshold. Model (4. 21), based on a delta from the threshold, resulted the best option for strains with virulence of grade 1 or grade 2 (*Swh 8/2/93*, *Bd23*, *Brk 12/2/93*, *Sls* and *Lausanne*), whereas model (4. 22), based on a switch from the threshold, performed the best for strains of grade 2 or grade 3 (*Ws6 346*, *Swh1209*, *Km13 II* and *Uriarra III*). For the most virulent strain, all the rabbits die within a few days and in this interval the progression of the viral growth from the threshold value results in a factor that quickly pushes the survival probability to zero. Instead, the survival times of the hosts infected with less virulent MYXV strains are distributed

over a wider fatality interval and a constant mortality rate can better describe the spread of these values.
The optimized parameters of the best model of each strain are listed in Table 4-9.

Table 4-9: Optimized parameters of the AIC best model of each strain.

Strain (virulence grade)	Dataset	Survival model	Threshold $[[log_{10}(V)]]$	$\beta\ [\frac{1}{[V]days}]$	$\omega\ [\frac{1}{days}]$
Swh 8/2/93 (2)	Kerr	Delta (4. 21)	7.886	0.492	-
Ws6 346 (2)	Kerr	Switch (4. 22)	8.198	-	0.427
Bd23 (1-2)	Kerr	Delta (4. 21)	8.267	19.427	-
Swh 1209 (4)	Kerr	Switch (4. 22)	8.305	-	0.233
Brk 12/2/93 (1-2)	Kerr	Delta (4. 21)	8.212	5.477	-
Sls (1)	Fenner	Delta (4. 21)	7.586	45.867	-
Km13 II (3)	Fenner	Switch (4. 22)	7.900	-	0.129
Uriarra III (3-4)	Fenner	Switch (4. 22)	8.700	-	0.129
Lausanne (1)	Fenner	Delta (4. 21)	7.970	2.069E+05	-

The fatality mechanisms of a MYXV strain are described by the intensity of the disease (β or ω depending on the model) or by a different amount of virus after which the consequences of the infection lead to the death of the hosts (*thr*). Strains with a higher grade (hence a lower virulence) are characterized by weaker effects and therefore lower β for model (4. 21) or lower ω for model (4. 22). Instead, the resulting values of *thr* do not appear to show a pattern correlated with the grade of the strains. The β value found for strain Lausanne results very different from the others because of the difference in the parameters of the growth model discussed in section 4.1.4.

The results of the 5%-95% confidence intervals of survival and death probabilities are reported in Figure 4-9 for the strains of the 'Kerr dataset' and in Figure 4-10 for the strains of the 'Fenner dataset'. The intervals are obtained from 100 repetitions of the process in which the death times of 50 individuals were randomly generated. The cumulated death probability π is simply obtained from the survival probability p as follows.

$$\pi = 1 - p \qquad (4.23)$$

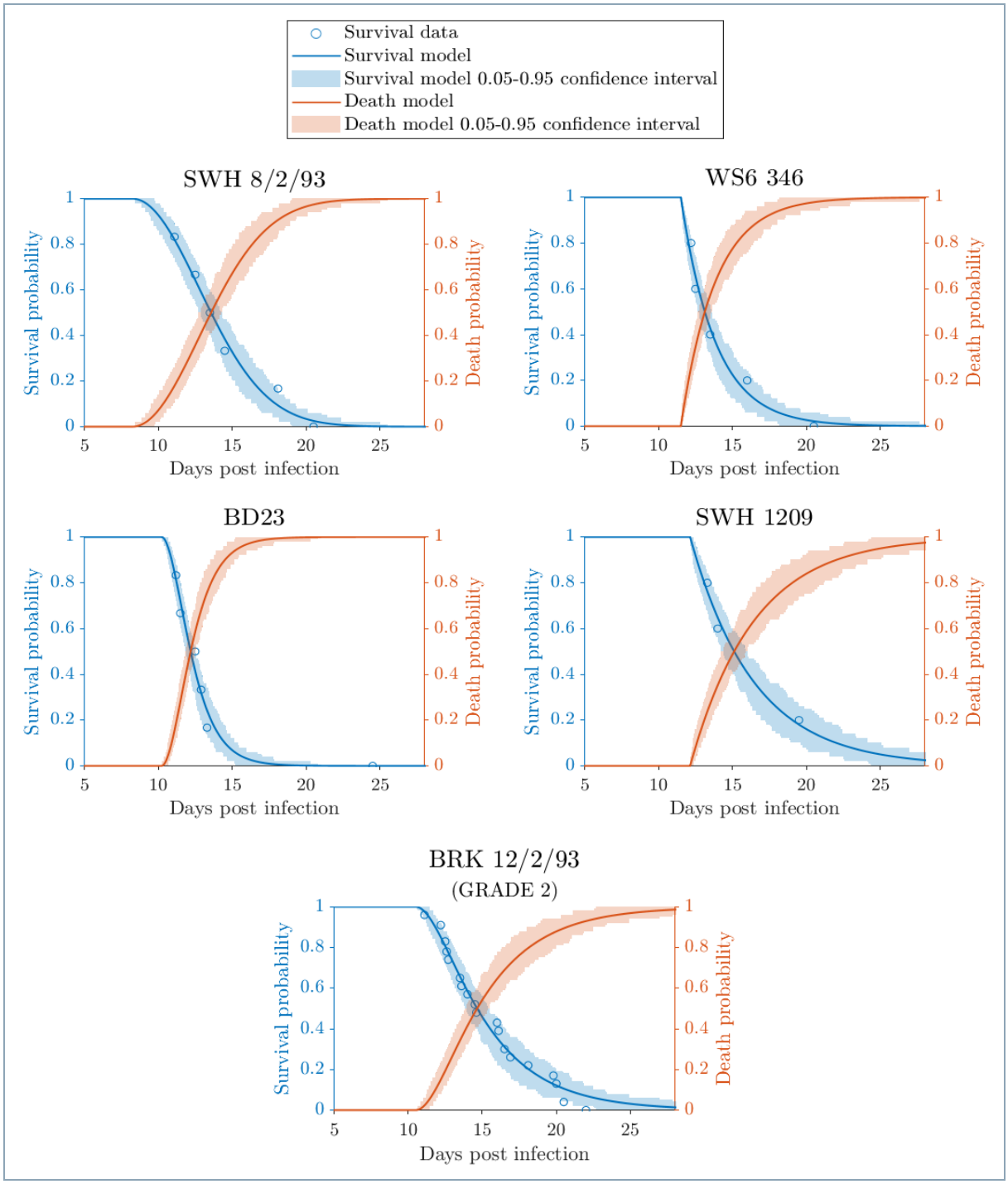


Figure 4-9: 5%-95% confidence interval of simulated survival and death probabilities. Where the grade is specified, the survival data are obtained from the aggregation of other strains.

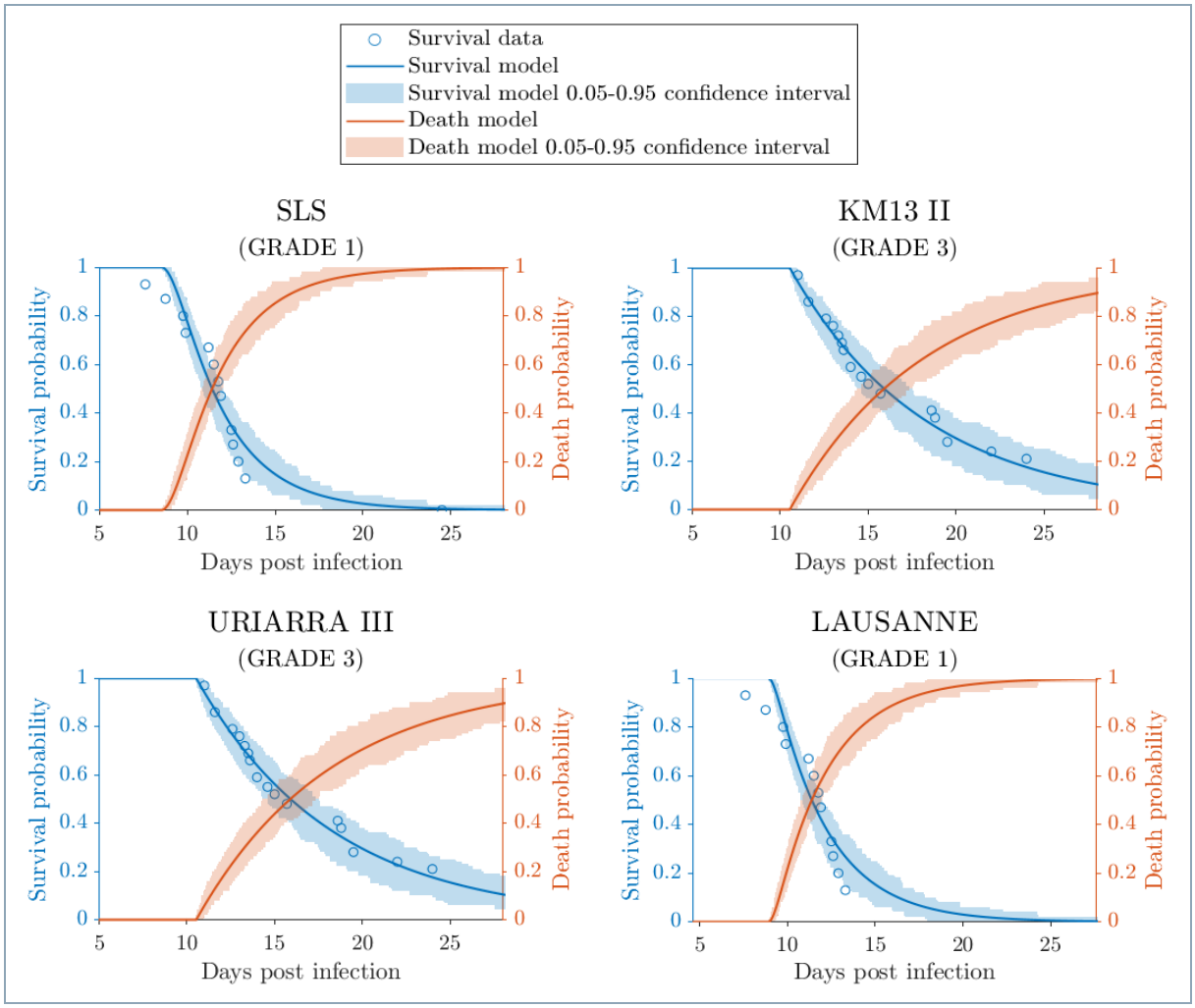


Figure 4-10: 5%-95% confidence interval of simulated survival and death probabilities. Where the grade is specified, the survival data are obtained from the aggregation of other strains.

4.2.4. Comments

Summarizing, the survival probability of the hosts was analyzed with a threshold-based mortality rate model correlated to the results of longitudinal models of MYXV within-host growth. With the combination of the two models, survival simulations of a fictious large number of rabbits can be performed to build confidence interval and better assess the biological variability of the system. Using the longitudinal and the survival models together, a complete characterization of the infection process of a strain was performed.

5 Effects of helminth co-infection on myxoma virus dynamics

This chapter of the thesis will cover the analysis of co-infection between *Myxoma virus* (MYXV) and the gastrointestinal helminth *Trichostrongylus retortaeformis* in the European rabbit. The overarching goal is to contribute to clarify how co-infections can generate heterogeneities in infection and transmission processes, by examining the effect of helminths on MYXV dynamics and host survival.
Previous studies have showed that the co-infection of helminths and viruses often triggers a trade-off in the host immune response of the host, where a Th1 reaction is generated against the virus while an antagonistic Th2 reaction is produced against the helminths [24]. However, this antagonist immune mechanism is affected by the intensities, the sites, and the relative timings of the two infections, including the role of host age, sex and breeding conditions. Recent studies have found that MYXV does affect the infection dynamics of *Trichostrongylus retortaeformis*, which has been showed to be immune-regulated, and less clearly of *Graphidum strigosum*, that appears to be weakly or no regulated by host immunity [43]. In contrast, the impact of helminths on MYXV within-host dynamics is fundamentally unknown. Here, the hypothesis that gastrointestinal helminths influence the survival rate and the viral load of rabbits co-infected with MYXV is examined. More specifically, by altering the host immune response and the ability to cope with multiple infections, helminths may facilitate MYXV infection, which in turn could lead to higher host mortality.

The 'co-infection dataset' (section 3.1.4) used in this chapter provides data of both co-infected and single-infected rabbits. Two approaches were explored:

- A comparison of host survival and MYXV viral load at death between rabbits co-infected with MYXV and helminths and rabbits only single-infected with MYXV. A series of statistical tests were performed, to seek significant differences between these two groups.
- The evaluation of the impact of helminths abundance of on host survival and MYXV load at host's death. Linear models were applied to understand the role of the estimated number of helminths as a predictor term for survival time and viral load at death of co-infected hosts.

5.1. Statistical tests on survival time and viral load at death of co-infected and single-infected rabbits

Tests to compare the survival curves and the averages, medians, and variances of survival time (*ST*) and viral load at death ($\log_{10}(VLD)$) of rabbits co-infected with *Myxoma virus* and *Trichostrongylus retortaeformis* and rabbits only single-infected with MYXV were carried out, to investigate whether the presence of helminths generates significant impacts. Each test was performed twice, first considering just the hosts deceased in the 30-days experiment and second also considering the surviving rabbits. In this chapter the results are reported only for the analysis without survivors, whereas the analysis with survivors can be found in Appendix B. For the MYXV strains *Perthshire 2082* and *Lausanne*, with no survivors, the two situations coincide.

5.1.1. Tests on survival time: methods

To look for significant differences between the survival times of single and co-infected rabbits, the following tests were performed using a threshold value α of 0.05 as level of significancy for the null hypotheses H_0.

- Log-rank test [44] on survival curves Ψ_{ST} (H_0: $\Psi_{ST,a} = \Psi_{ST,b}$, α=0.05). Censoring was applied on the data of rabbits surviving the experiment period, whose fate is unknown.
- t-test [45] on averages μ_{ST} (H_0: $\mu_{ST,a} = \mu_{ST,b}$, α=0.05). t-tests were carried out considering whether the variances were similar or not, in relation to the F-test results.
- Wilcoxon test [46] on indistinguishability of populations, through true location shift ξ_{ST} (H_0: $\xi_{ST,a\text{-}b}$=0, α=0.05). Non-parametric Wilcoxon tests were performed to provide a different statistical view from the t-test in the comparison of the two groups, covering the case of non-normal distributions and trying to catch possible bias due to sample size. Assuming the distributions of the two groups similar, the Wilcoxon test was considered also for the comparison between medians (H_0: $\lambda_{ST,a} = \lambda_{ST,b}$, α=0.05) [47].
- F-test [45] on variances σ^2_{ST} (H_0: $\sigma^2_{ST,a} = \sigma^2_{ST,b}$, α=0.05).

5.1.2. Tests on survival time: results

The survival curves of single- and co-infected rabbits are compared in Figure 5-1 and boxplots of their survival times are reported in Figure 5-2, for each strain.

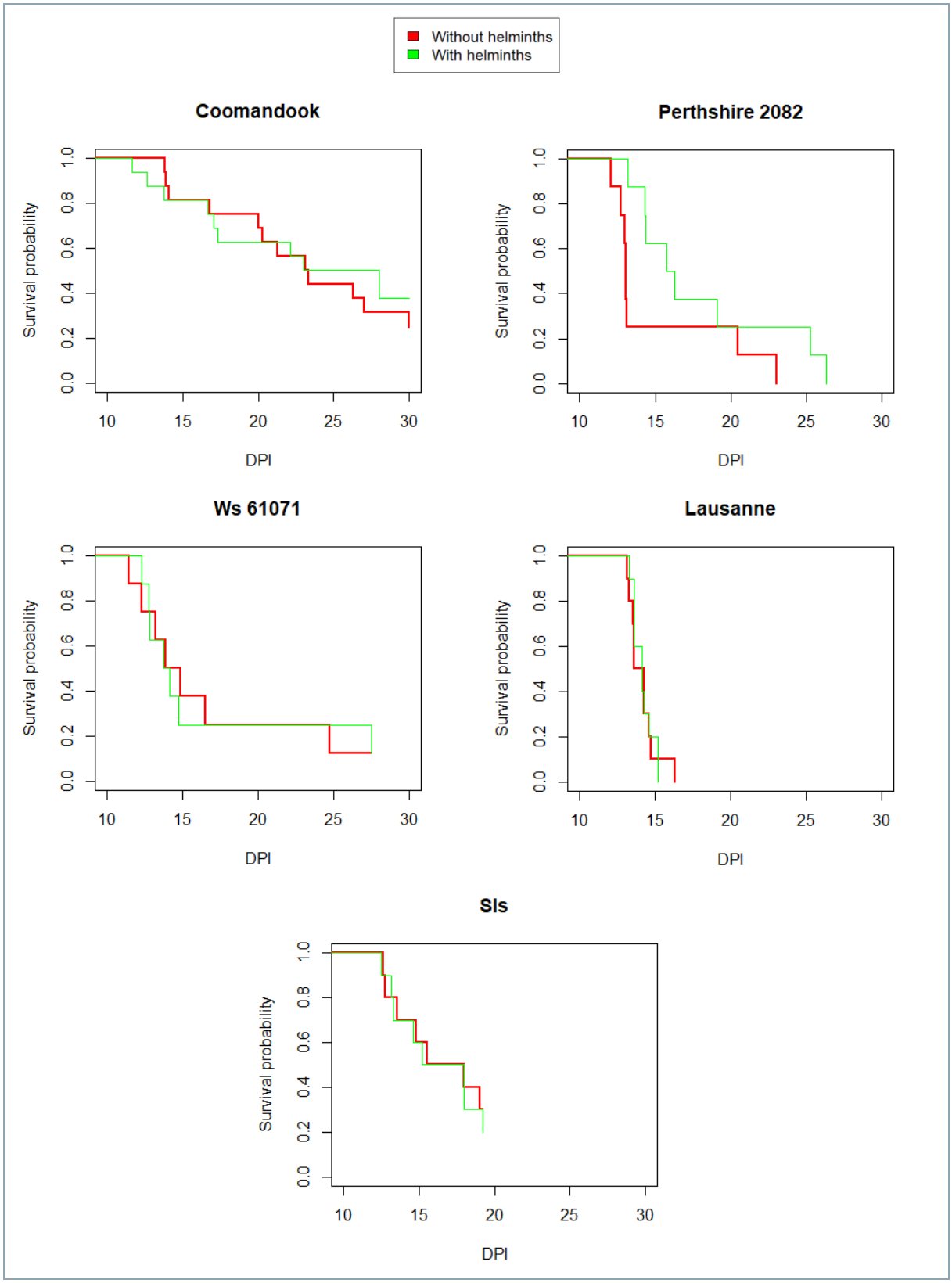


Figure 5-1: Comparison of survival curves between MYXV-helminths co-infected rabbits and MYXV single-infected rabbits, for different strains.

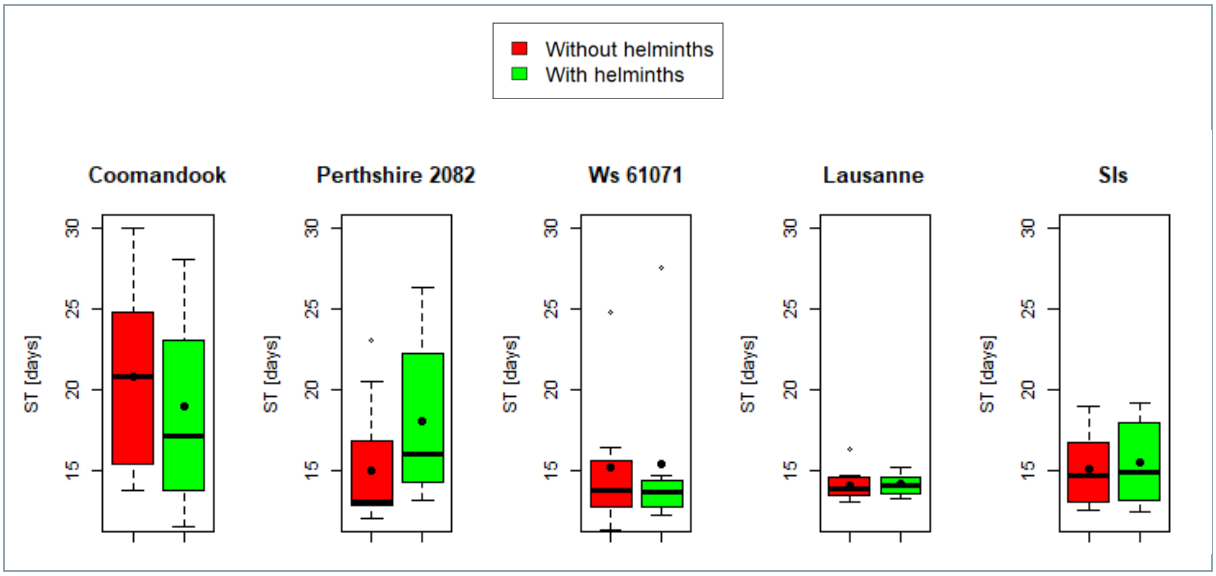


Figure 5-2: Boxplots to compare the survival times of MYXV-helminths co-infected rabbits and MYXV single-infected rabbits for each strain, excluding the survivors. Each box reports the minimum non-outlier, the first quartile, the median (black line), the average (black dot), the third quartile and the maximum non-outlier.

The results of the log-rank tests on the survival curves, t-tests on averages, Wilcoxon tests on populations and medians, and F-tests on variances of the two groups were listed in Table 5-1. The results highlight that in almost all the cases the empirical evidence is not sufficiently contrary to the null hypotheses of similar statistical indexes and curves. The only exception can be found for the strain *Perthshire 2082*, for which the Wilcoxon test suggests differences between the two groups and the median with helminths can be considered significantly larger than the one without helminths; a tendency for a difference, albeit not significant, was also found between the survival curves. Altogether, findings suggest that co-infection with helminths does not appear to cause underlinable differences in the host survival time, whether the survivors of the experiment were excluded or included (see Appendix B).

Table 5-1: Performances of log-rank tests on survival curves, t-tests on ST averages, Wilcoxon tests on ST populations and medians and F-tests on ST variances, between MYXV-helminths co-infected rabbits and MYXV single-infected ones (α=0.05), excluding survivors.

Strain (virulence grade)	Log-rank test pvalue	Average ST [days] without helminths	Average ST [days] with helminths	t-test pvalue	Median ST [days] without helminths	Median ST [days] with helminths	Wilcoxon test pvalue	Variance ST [days^2] without helminths	Variance ST [days^2] with helminths	F-test pvalue
Comandook (3.2)	0.633	20.818	19.013	0.467	20.790	17.165	0.410	29.502	35.835	0.749
Perthshire 2082 (3.1)	0.077	15.074	18.070	0.219	13.065	16.005	0.041	17.571	25.866	0.623
Ws 61071 (3.1)	0.944	15.254	15.429	0.949	13.830	13.750	1.00	20.301	29.123	0.672
Lausanne (1)	0.926	14.128	14.169	0.912	13.935	14.130	0.650	0.893	0.445	0.314
Sls (1)	0.744	15.159	15.485	0.808	14.770	14.895	0.862	6.322	6.618	0.972

5.1.3. Tests on viral load at death: methods

To look for significant differences in the viral load at death between the two groups, the following tests were performed. For all of them, a threshold value α of 0.05 was selected as level of significancy for the null hypothesis H_0.

- t-test [45] on averages $\mu_{log10(VLD)}$ (H_0: $\mu_{log10(VLD),a} = \mu_{log10(VLD),b}$, α=0.05). t-tests on the averages were carried out by considering whether the variances were similar or not, in relation to the F-test results.
- Wilcoxon test [46] on indistinguishability of populations (through population true location shift $\xi_{log10(VLD)}$) and medians $\lambda_{log10(VLD)}$ (H_0: $\xi_{log10(VLD)}$=0, α=0.05). Non-parametric Wilcoxon tests were performed to provide a different statistical view from the t-test in the comparison of populations, covering the case of non-normal distributions and trying to catch possible bias due to sample size. Assuming the distributions of the two groups similar, the Wilcoxon test was considered also a test on the comparison between medians (H_0: $\lambda_{log10(VLD),a} = \lambda_{log10(VLD),b}$, α=0.05) [47].
- F-test [45] on variances $\sigma^2_{log10(VLD)}$ (H_0: $\sigma^2_{log10(VLD),a} = \sigma^2_{log10(VLD),b}$, α=0.05)
- t-test on averages, Wilcoxon test on medians and F-test on variances with Jackknife [42] on outliers. To prevent atypical viral load values from affecting the tests, the tests were repeated removing the outliers one at a time, and the best obtained score was selected. The results of this approach are reported in Appendix C.

All the tests were conducted with the virus load in logarithmic form (base 10). For some rabbits the amount of virus detected in the lungs upon death was null. To proceed in the analysis, those '0' viral loads were converted to '1' before applying the log-transformation of the data.

5.1.4. Tests on viral load at death: results

A boxplot comparison of viral load at death between MYXV-helminths co-infected rabbits and MYXV single-infected rabbits is reported in Figure 5-3.

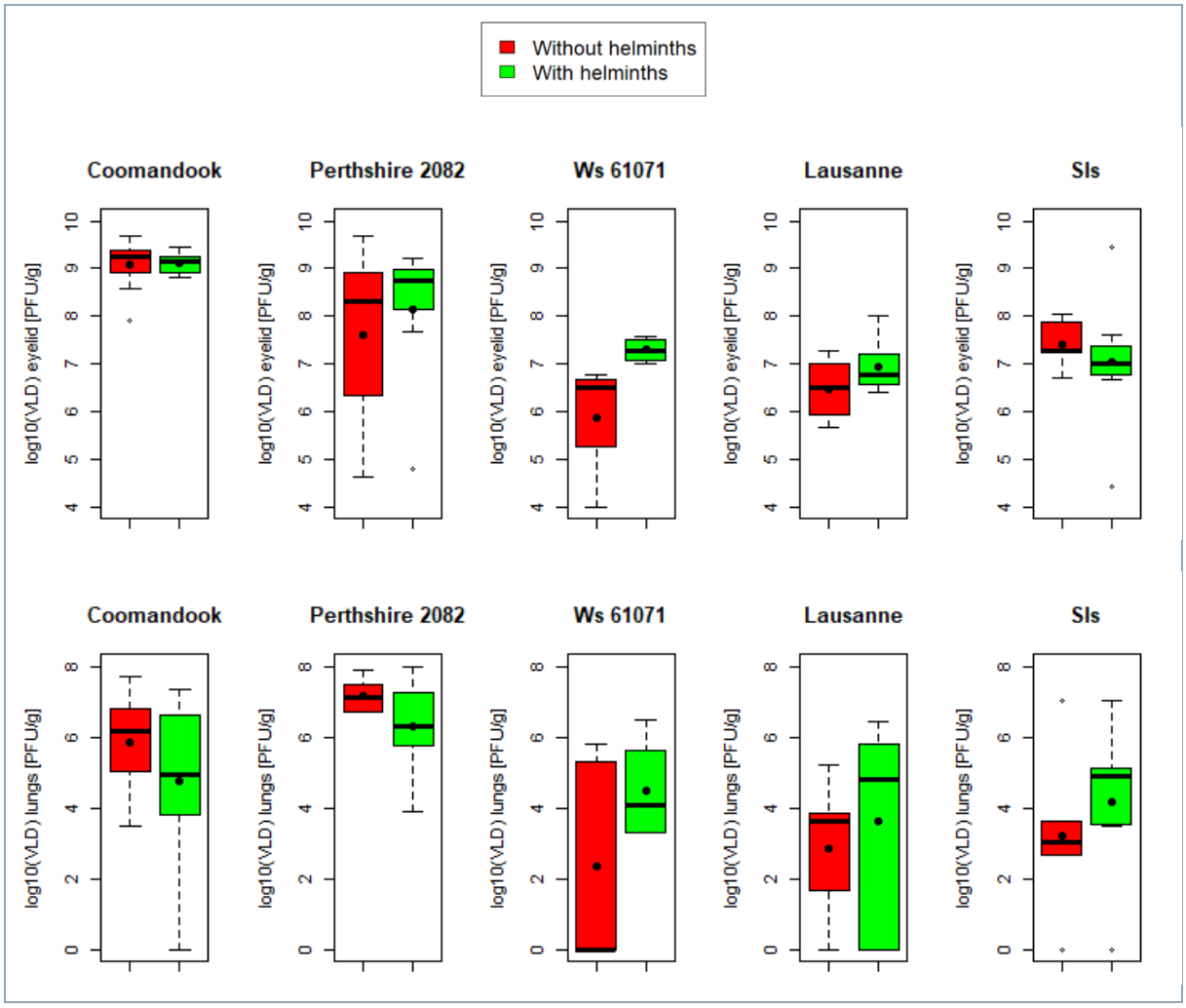


Figure 5-3: Boxplots to compare the viral load at death of MYXV-helminths co-infected rabbits and MYXV single-infected rabbits for each strain, excluding the survivors. Each box reports the minimum non-outlier, the first quartile, the median (black line), the average (black dot), the third quartile and the maximum non-outlier.

The results of the t-tests on averages, the Wilcoxon tests on populations and medians, and the F-tests on variances are reported in Table 5-2 and Table 5-3, respectively for the eyelid and the lungs. The results suggest that there is no significative evidence to refuse the hypothesis of similar statistical indexes for most of the strains. Only for strain *Ws 61071*, the means, medians, and variances of $\log_{10}(VLD)$ in the eyelid can be evaluated as significantly different between the two groups, with the viral load at death being larger in rabbits co-infected with helminths. This outcome is, however, not consistent with what results in the lungs. Altogether, findings suggest that co-infection with helminths does not appear to cause remarkable differences in the viral load at death of the hosts, whether the survivors of the experiment were excluded or included (see Appendix B).

Table 5-2: Performances of t-tests on $\log_{10}$(VLD) averages, Wilcoxon tests on $\log_{10}$(VLD) populations and medians and F-tests on $\log_{10}$(VLD) variances between MYXV-helminths co-infected rabbits and MYXV single-infected ones, in the eyelid (α=0.05), excluding survivors.

Strain (virulence grade)	Average $\log_{10}$(VLD) [PFU/g] without helminths	Average $\log_{10}$(VLD) [PFU/g] with helminths	t-test pvalue	Median $\log_{10}$(VLD) [PFU/g] without helminths	Median $\log_{10}$(VLD) [PFU/g] with helminths	Wilcoxon test pvalue	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ without helminths	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ with helminths	F-test pvalue
Comandook (3.2)	9.098	9.109	0.953	9.234	9.157	0.671	0.236	0.051	0.053
Perthshire 2082 (3.1)	7.594	8.136	0.573	8.307	8.745	0.609	3.685	2.430	0.626
Ws 61071 (3.1)	5.871	7.290	0.012	6.498	7.284	0.010	1.146	0.073	0.046
Lausanne (1)	6.475	6.952	0.168	6.500	6.761	0.272	0.370	0.351	0.988
Sls (1)	7.422	7.025	0.583	7.260	6.998	0.330	0.289	2.178	0.071

Table 5-3: Performances of t-tests on $\log_{10}$(VLD) averages, Wilcoxon tests on $\log_{10}$(VLD) populations and medians and F-tests on $\log_{10}$(VLD) variances between MYXV-helminths co-infected rabbits and MYXV single-infected ones, in the lungs (α=0.05) excluding survivors.

Strain (virulence grade)	Average $\log_{10}$(VLD) [PFU/g] without helminths	Average $\log_{10}$(VLD) [PFU/g] with helminths	t-test pvalue	Median $\log_{10}$(VLD) [PFU/g] without helminths	Median $\log_{10}$(VLD) [PFU/g] with helminths	Wilcoxon test pvalue	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ without helminths	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ with helminths	F-test pvalue
Comandook (3.2)	5.823	4.751	0.211	6.183	4.922	0.375	1.888	5.480	0.112
Perthshire 2082 (3.1)	7.164	6.318	0.171	7.102	6.301	0.283	0.233	1.920	0.035
Ws 61071 (3.1)	2.340	4.471	0.214	0.000	4.053	0.384	8.573	2.264	0.302
Lausanne (1)	2.845	3.638	0.574	3.629	4.785	0.346	4.119	8.349	0.415
Sls (1)	3.231	4.165	0.464	3.045	4.875	0.317	5.082	4.753	0.919

5.1.5. Comments

Overall, the application of multiple statistical tests does not evidence significant differences to distinguish single-infected and co-infected rabbits. Only a few cases reject the hypothesis of similarity, and the results were not consistent examining viral load at death in eyelid and lungs. In a larger number of cases the average MYXV load at death is higher in co-infected rabbits, but this is not sufficient to identify a pattern. The average survival times with or without helminths tend to become similar moving from low to high virulent strains. This suggests that for the most fatal strains, which quickly cause the death of the host, the co-presence of helminths is probably less impactful. Animals survive longer when infected with less virulent strains, allowing more time for the emergence of differences both within and between infection groups, although not statistically significant. The inclusion of survivors does not substantially change the general outcomes (Appendix B). Moreover, the Jackknife approach on outliers (Appendix C) show that some individuals consistently influence the tests, but the results still do not detect a clear effect caused by co-infection.

5.2. Predictions of survival time and viral load with helminth abundance at host death

The analyses in this chapter involve only rabbits co-infected with *Myxoma virus* and gastrointestinal helminths *Trichostrongylus retortaeformis* ('co-infection dataset', section 3.1.4). Here, the effects of the number of helminths inside the host on viral growth and host survival are investigated, for different strains. More specifically, a variable representing the number of helminths at death (*NHD*) was introduced in linear models of viral load at death ($\log_{10}(VLD)$) and host survival time (*ST*), as an additive and interacting term. The goal is to understand whether the interaction between *NHD* and $\log_{10}(VLD)$ is significant in *ST* predictions and whether the interaction between *NHD* and *ST* is significant in $\log_{10}(VLD)$ predictions.
Analyses were performed using the *glm* tool from the software R in accordance with the following procedure:

- choice of the probability distribution of the data
- comparison of predictive models by using Akaike Information Criterion [34]
- identification of the optimal value of their coefficients
- evaluation of the importance of each predictor based on the returned p-values
- visual and numeric research of patterns/differences between strains

The method was applied considering the amount of virus in the eyelid and excluding the survivors of the 30-days experiment, for which the actual survival time and viral load at death are unknown. Only for the strain *Ws 61071* it was necessary to include the data of the sole survivor, as the sample size would not have been sufficient to

perform the following analyses otherwise. The *ST* of that rabbit was set to 30 days and its *VLD* was set to the amount measured the 30th day post infection, assuming the end of the experiment as death condition.
In Appendix D the correlation matrices of *ST*, $\log_{10}(VLD)$ and *NHD* are reported for each strain to provide a general idea of the relationship between these variables. The matrices immediately highlight a strong variability of the results between strains.

5.2.1. Model candidates

Many aspects must be considered when analyzing the link between MYXV viral load, abundance of helminths and host survival time. The co-presence of helminths and MYXV biologically is expected to cause an antagonist response in the host, where a type 1 helper T cells is activated against the virus while a type 2 helper T cells is developed against helminths [24]. These contrasting reactions, which are expected to alter the rabbit survival, support the hypothesis that helminths are harmful to the host also because they could "divert" the immune response from fighting the virus. However, an immune system already warmed-up by the presence of a parasite may sometimes be able to better control an incoming infection [19], or the immunosuppressive action of the virus itself can influence in some cases the Th1-Th2 paradigm making it less significant. Furthermore, a factor that helps MYXV to raise faster implies a shorter survival time, which in turn could lead to a lower viral load at death by prematurely stopping the virus growth. The number of helminths can indeed affect the host immune system and its response to MYXV in many possible ways, even if it is still the effect of the virus that causes the decease.

Different predictive model versions were developed and tested to investigate these interaction dynamics, evaluating in the same way both the prediction of $\log_{10}(VLD)$ with data of *ST* and *NHD* and the prediction of *ST* with data of $\log_{10}(VLD)$ and *NHD*. Initially, the responses were modelled as linear combination of predictors, with the effect of helminths included as an additive component. From this framework, three model options were compared: a linear model with no extra interaction, a model with a direct extra interaction and a model with an inverse extra interaction.

5.2.1.1. Linear model with no interaction

The first developed option is a basic linear model structure, as reported below, where the role of each predictor is simply additive. This model version supports a system in which helminths and virus have their individual effects and these effects do not significantly influence each other.

$$log_{10}(VLD) = a + b\,ST + c\,NHD \qquad (5.1)$$

$$ST = e + f\ log_{10}(VLD) + g\ NHD \quad (5.2)$$

5.2.1.2. Model with direct interaction

The second option implies that helminths abundance interacts with survival time and viral load at death in a direct way, through a multiplicative term:

$$log_{10}(VLD) = a + b\ ST + c\ NHD + d\ ST\ NHD \quad (5.3)$$

$$ST = e + f\ log_{10}(VLD) + g\ NHD + h\ NHD\ log_{10}(VLD) \quad (5.4)$$

5.2.1.3. Model with inverse interaction

The last option considers that helminths abundance interacts with survival time and viral load at death in an inverse way, through a partition term:

$$log_{10}(VLD) = a + b\ ST + c\ NHD + d\ \frac{ST}{NHD} \quad (5.5)$$

$$ST = e + f\ log_{10}(VLD) + g\ NHD + h\frac{log_{10}(VLD)}{NHD} \quad (5.6)$$

5.2.2. Methods

For all the considered options, a Gaussian family distribution of ST and $\log_{10}(VLD)$ was selected among the available possibilities. Some comments about this choice are reported in Appendix E. The model candidates were evaluated by applying a Akaike Information Criterion approach with the formulation shown in Table 5-4.

Table 5-4: Model selection of viral load at death and survival time of co-infected hosts.

Response X= $log_{10}(VLD),\ ST$
Predictors Y = $NHD,\ log_{10}(VLD),\ ST,\ f(NHD, VLD, ST)$
Parameters $\theta = (a, b, c, d, e, f, g, h)$; $m = count(\theta)+1$
Objective function $J(\theta) = AIC = 2m + n \ln(\frac{\sum_{i=1}^{n}(X_{i,obs}-X_{i,sim})^2}{n})$
Optimization φ^* such that $J(\varphi^*) = J^* = \min_{\varphi}(J(\theta))$; $\varphi \in \theta$

All the coefficients (*a*,*b*,*c*,*d*,*e*,*f*,*g*,*h*) of the predictors can assume positive or negative real values. The optimized coefficients of the best model option were compared between strains and inspected by looking at the assigned p-values, to assess the importance of each single predictor. The 3D model outcomes were built by a predicted grid (with the tool *expand.grid* from the software R) from a data frame containing 128 values between minimum and maximum of each predictor.

5.2.3. Results

The helminth abundance in relation to viral load at host death and host survival time are illustrated in Figure 5-4, for each strain.

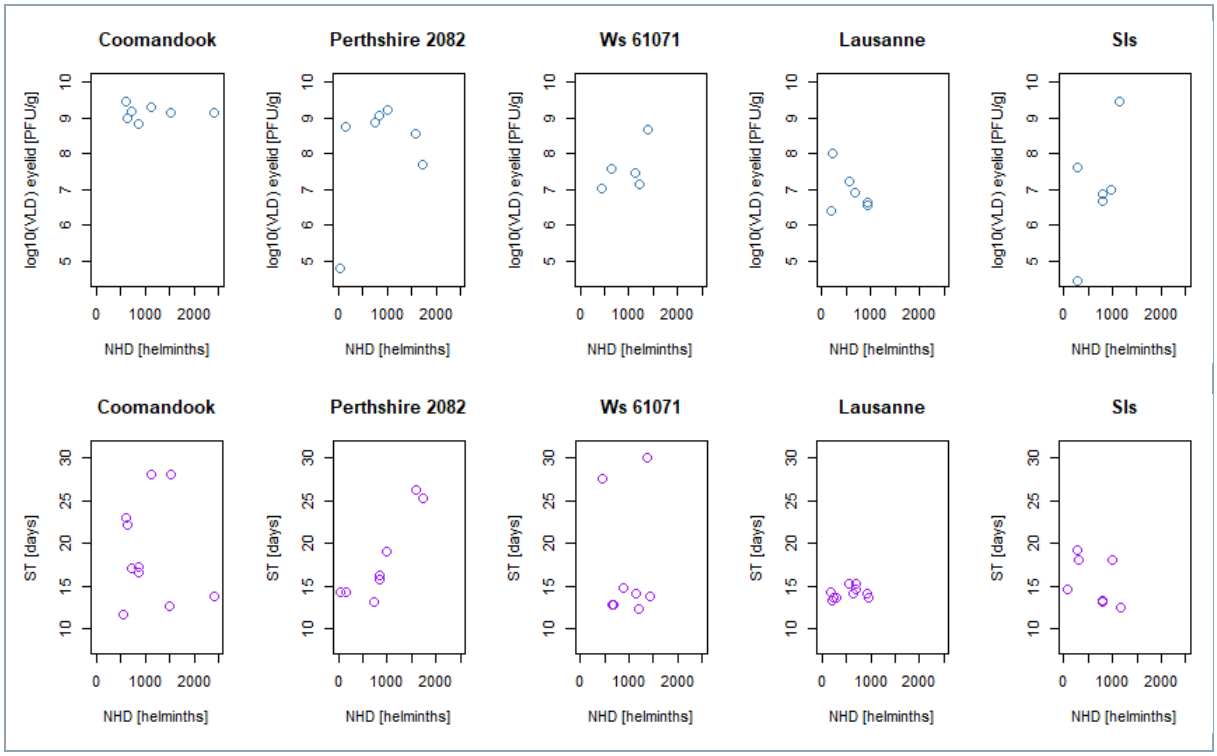


Figure 5-4: Abundance of helminths at death in relation to host survival time and viral load at death in the eyelid and in the lungs, for different strains, excluding the survivors (except from strain Ws 61071, for which the only survivor is included and set as ST=30).

The number of helminths at death seems to decrease with the virulence: *Lausanne* and *Sls* (both grade 1) show values close to or below 10^3, in average smaller than the ones of *Ws 61071*, *Perthshire 2082* (both grade 3.1) and *Coomandook* (grade 3.2). This trend can be explained by considering that a less virulent strain means higher survival time and so a longer period for helminths to survive and reproduce. However, the temporal distribution of helminths abundance at host death does not show a simple pattern. In one hand, the periodical administration of helminths reflects the natural infection

mechanism but can also influence the analysis. In fact, helminths doses were administered with a week between one and the other and only some larvae move to the intestine and become easily detected after days [48]. On the other hand, even for the same day of death, the number of helminths in different hosts varies widely, suggesting non-uniform responses of the hosts.

The AIC scores obtained with the model selection process are shown in Table 5-5 for $\log_{10}(VLD)$ predictions and in Table 5-6 for *ST* predictions. In most cases, the inclusion of an inverse interacting term (5. 5) (5. 6) leads to the best predictions. The addition of the inverse interaction appears to be a relevant effect in all cases except only for the viral load at death of *Sls*. In one case, *ST* of *Ws 61071*, the model with direct interaction produces the best predictions, but it predicts negative survival times for some realistic combinations of helminths abundance and viral load at host death. This option is indeed discarded and model (5. 6) is selected also for *Ws 61071*.

Table 5-5: $\log_{10}$(VLD) predictive model selection, with the lowest AIC score marked in bold.

Strain	Grade	Model (5. 1) AIC	Model (5. 3) AIC	Model (5. 5) AIC
Coomandook	3.2	3.618	2.731	**1.581**
Perthshire 2082	3.1	29.701	29.803	**7.412**
Ws 61071	3.1	12.682	-4.836	**-14.491**
Lausanne	1	16.057	12.333	**1.457**
Sls	1	**26.035**	27.373	26.294

Table 5-6: ST predictive model selection, with the lowest AIC score marked in bold.

Strain	Grade	Model (5. 2) AIC	Model (5. 4) AIC	Model (5. 6) AIC
Coomandook	3.2	54.139	53.564	**50.718**
Perthshire 2082	3.1	36.862	38.846	**12.843**
Ws 61071	3.1	40.452	**2.378**	24.357
Lausanne	1	20.824	26.432	**5.395**
Sls	1	33.025	35.017	**17.121**

In Table 5-7 and Table 5-8 the selected models are described. The tables contain the p-values assigned to each predictor and the optimized value of their coefficients (representing how the response changes for a unit increase in the relative predictor). Looking at the p-values, the results highlight that for both $\log_{10}(VLD)$ and ST the inverse interaction terms result generally relevant.

Table 5-7: P-values and coefficients of the selected $\log_{10}$(VLD) predictive model.

Strain	a intercept (pvalue)	b Coeff. ST (pvalue)	c Coeff. NWD (pvalue)	d Coeff. ST/NWD (pvalue)
Coomandook	7.913 (1.445E-04)	0.005 (0.789)	4.192E-04 (0.171)	27.396 (0.180)
Perthshire 2082	9.563 (0.001)	0.100 (0.235)	-0.002 (0.055)	-18.893 (0.002)
Ws 61071	8.492 (0.012)	0.159 (0.029)	-0.002 (0.054)	-81.583 (0.034)
Lausanne	23.522 (0.017)	-0.660 (0.048)	-0.007 (0.027)	-83.516 (0.032)
Sls	7.752 (0.309)	-0.144 (0.676)	0.002 (0.482)	-

Table 5-8: P-values and coefficients of the selected ST predictive model.

Strain	a intercept (pvalue)	b Coeff. $\log_{10}$(VLD) (pvalue)	c Coeff. NWD (pvalue)	d Coeff $\log_{10}$(VLD)/NWD (pvalue)
Coomandook	-113.680 (0.172)	19.140 (0.082)	-0.016 (0.103)	-2237.167 (0.120)
Perthshire 2082	-37.370 (0.005)	3.751 (0.003)	0.019 (3.337E-04)	301.825 (0.002)
Ws 61071	-133.305 (0.098)	-0.411 (0.903)	0.084 (0.119)	8138.838 (0.105)
Lausanne	21.694 (0.009)	-0.054 (0.845)	-0.007 (0.032)	-170.822 (0.028)
Sls	4.711 (0.281)	-6.751 (0.024)	0.048 (0.031)	1956.695 (0.026)

The optimized coefficients vary both in intensity and sign between strains. This suggests that, as discussed before, the influence of helminths on viral growth (and consequently on host survival time) does not act in a univocal way. Both viral growth and survival of hosts are characterized by a strong variability and, with the available sample size, each observation considerably impacts the model optimization.
The predictive surfaces of MYXV viral load at death and host survival time obtained from the selected models are displayed in Figure 5-5 and Figure 5-6.

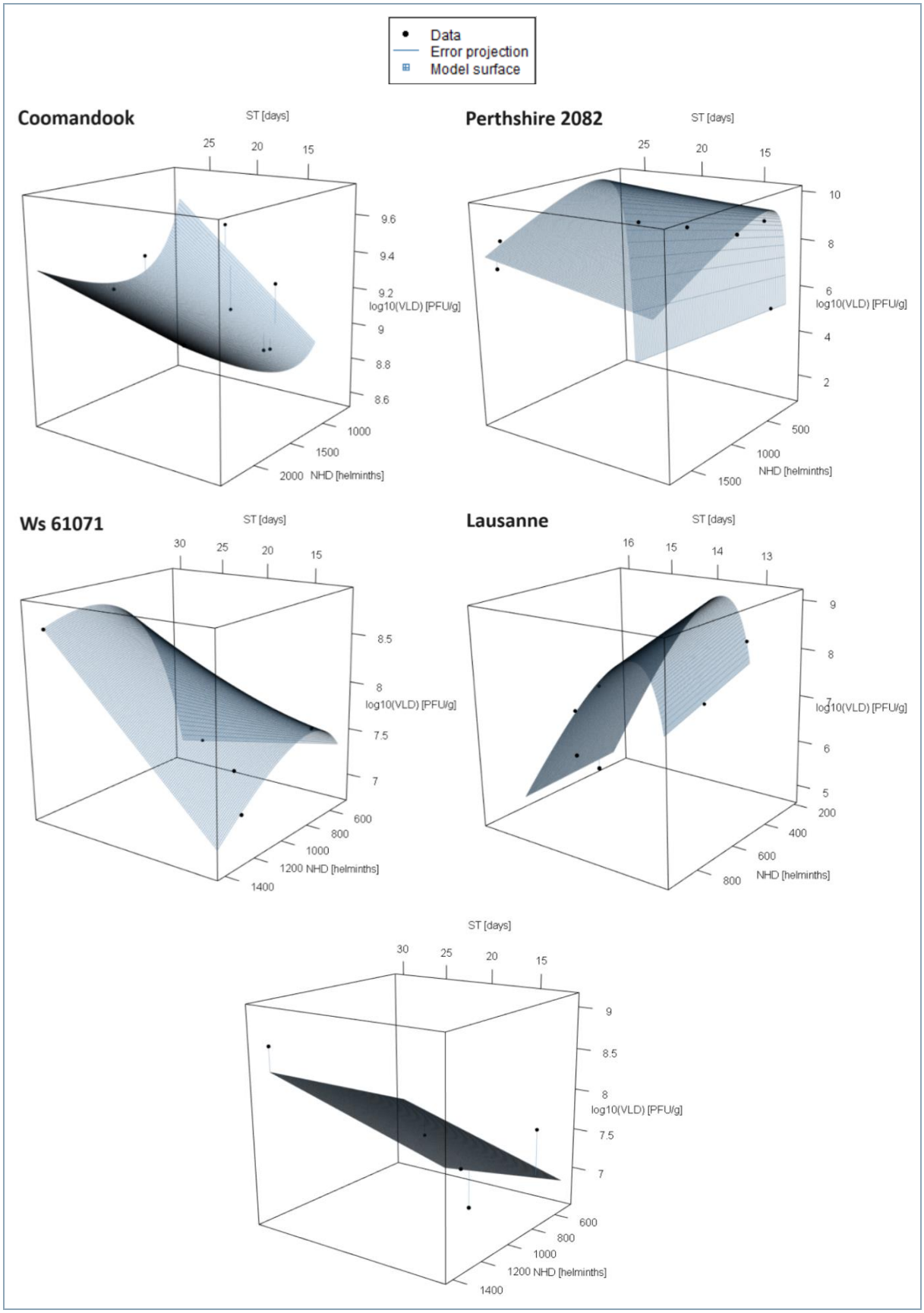


Figure 5-5: Viral load at death (log_{10}(VLD)) predictive surface obtained from selected model.

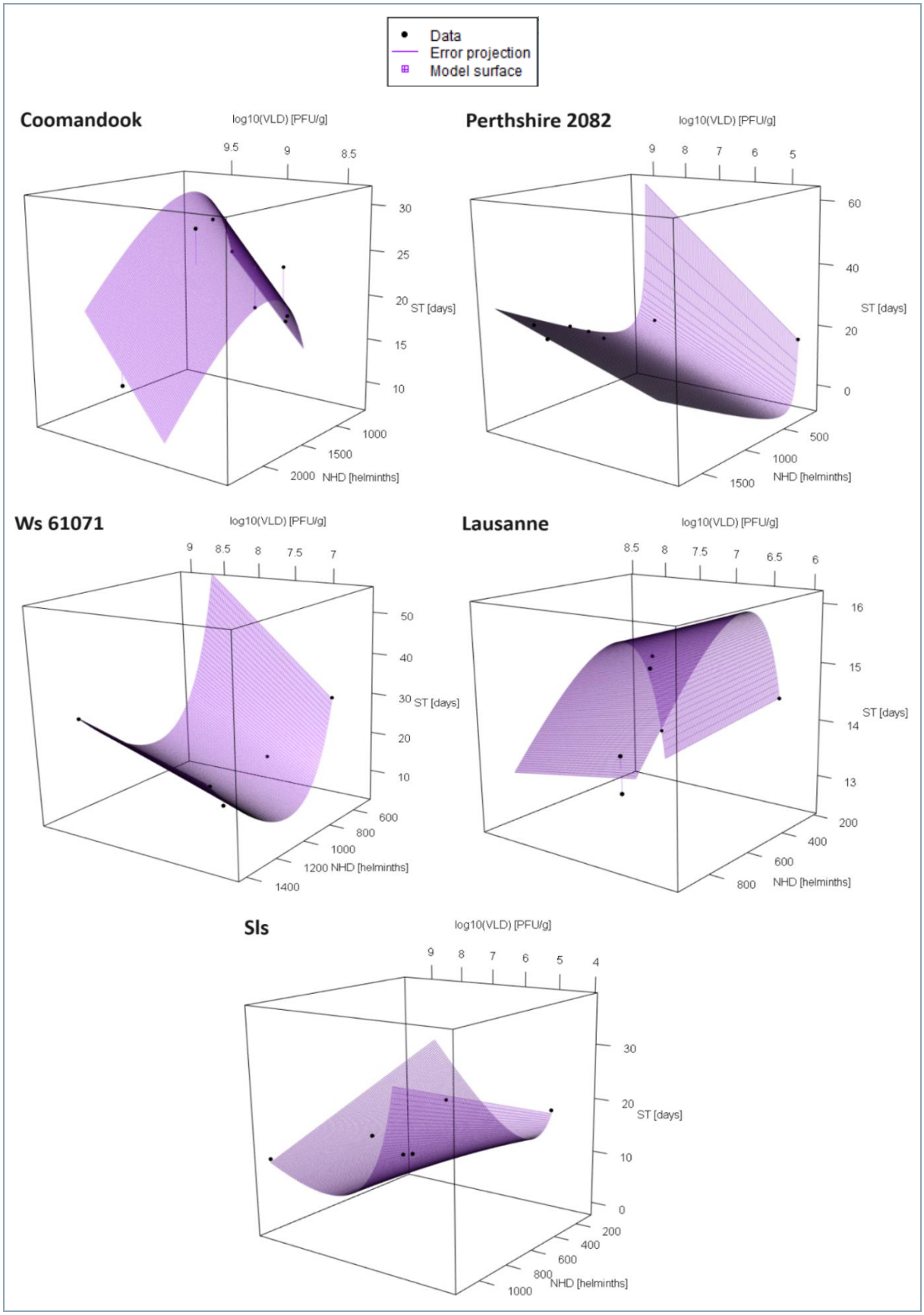


Figure 5-6: Survival time (ST) predictive surface obtained from selected model.

The obtained surfaces show very different features. The grade of virulence does not seem to be a solid way to classify the co-infection dynamics: *Sls* was expected to be akin to *Lausanne* (both grade 1), while *Coomandook* was expected to be similar to *Perthshire 2082* and *Ws 61071* (grade 3.2, grade 3.1 and grade 3.1). Overall, these results stress the large biological variability of the *Myxoma virus- Trichostrongylus retortaeformis* co-infection system in the European rabbit, caused by the ways different MYXV strains interact with host and helminths via the immune response and by the individual characteristics of each rabbit. Here, with the available sample size, the allocation of each single data point becomes a very strong factor in surface modelling and the predictions tend to diverge at the extremes of the ranges. The resulting frameworks cannot therefore be considered as high-performing predictive models, but they still suggest that the inclusion of an interaction representing co-infection dynamics may significantly improve their performance.

5.2.4. Comments

The model performance and accuracy of predictions could be further improved. In Appendix F two additional model options are reported, developed by modifying the predictor element representing the abundance of helminths in two different ways:

- By applying a logarithmic transformation to the number of helminths, which brings the value scale closer to the MYXV values (expressed a $\log_{10}(VLD)$).
- By modulating the helminths abundance with the number of helminth doses administered to the rabbits.

However, in both cases, the performance of the models does not generally improve. Other modifications could be taken into account to build more powerful predictive models, by considering other types of interaction or by including other non-linear predictor terms. These additional analyses are left for future analyses. The current study has identified that the interaction between MYXV and helminths produces a significant contribution to predictive models, but without a common effect on the strains under study. Different traits of the strains, possible different effects of MYXV on the immune responses, and different characteristics of the animals in respect to the sample size made it difficult to detect a clear scheme. It is possible, however, that a pattern could be pinpointed with new research.

6 Conclusions and future directions

With the work performed in this thesis, some extra steps in the comprehension of the MYXV within-host dynamics in the European rabbit were taken. Two main topics were explored: i- the development of mathematical models to describe the viral growth in the host and the consequence of infection on host survival, and ii- the investigation of the effects of helminths co-infection on host survival time and MYXV viral load at host death. These topics were examined using different strains of MYXV, representing diverse grades of virus virulence.

As far as the first is concerned, mathematical models were developed by adapting to our system few classical functions of species growth and by selecting the best framework through the Akaike Information Criterion, which takes into account model accuracy and complexity. MYXV growth was modelled with good accuracy using simple two- or three-parameters model options. Specifically, the strains that grow over the entire experimental period and whose host immune system is uncapable of control are well described by a Malthusian or a logistic model. The choice between the two functions depends on whether the carrying capacity of the virus is approached before the fatality interval (logistic case) or the rabbits die before that capacity is approached (Malthusian case). In contrast, for the strains that the host immune response is able to control by reducing the viral load after an initial growth, the host immune response is explicitly modelled as a dynamic killing rate of the virus. The model structure is therefore identified as a differential system based on Malthusian or logistic viral growth. This model structure could be applied also for virulent strains in a population-based analysis aimed to describe the percentage of animals that succeed in eradicating the virus and heal.

The speed of viral growth is not the only factor that influences the virulence of the strains, characterized by different disease severities. Models of host mortality rate were developed assuming that the mortality rate increase as a function of the viral load $\mu(V_t)$ and selected optimizing the consequent survival probability on the survival curves built with survival time data. The selection process identified two models with viral threshold as the best options. In particular, a model with a post-threshold proportional mortality rate increase was selected for rabbits infected with high virulent strains and that perished in a short time interval. In contrast, a model with on-off mortality was selected for rabbits infected with mid-virulent MYXV strains, whose survival times are distributed over a wider period, better described by a constant

mortality rate. In both cases, host fatalities are described by the intensity of the disease, (β or ω, depending on the model), and by a viral threshold after which the severity of the infection can lead to the death of the host (*thr*). Strains with a lower virulence (hence a higher grade) are characterized by lower β or ω. Instead, the resulting values of *thr* do not appear to show a pattern correlated with the virulence of the strains.
The uncertainty of the models was examined by building confidence intervals for both the longitudinal and the survival simulations. The two sets of models provide together a characterization of the infection process that could be a precious tool in new applications. This thesis proposes models with reduced complexity, undemanding of big data sizes and that produce good fit performances with a small computational effort. The system could be potentially further investigated with the availability of new data, such as longitudinal measurements of some immune system components. For example, the immunosuppressive properties of highly virulent strains against target components of the immune response could be examined, as well as the role of fundamental immune variables on virus control and clearance.

The second part of this thesis moved to a focus on co-infection between *Myxoma virus* and gastrointestinal helminth *Trichostrongylus retortaeformis* in the European rabbit. In this perspective, this work investigated the effect of helminths on the MYXV growth in the hosts and their survival times.
The application of various statistical tests (t-test, Wilcoxon test, F-test, log-rank test) did not produce significant outcomes to distinguish single-infected from co-infected cases. Only a few strains reject the hypotheses of similarity and only in one of the considered sites (lungs or eyelid). In highly virulent strains, the gap of average survival times with or without helminths seems to be narrow, suggesting that the co-presence of helminths is apparently even less impactful when the virus quickly causes the death of the host. Moreover, the inclusion or exclusion of survivors in their respective groups seem not to be a major factor in the analysis.
For the available samples, the abundance of gastrointestinal helminths in the host at the time of death appears to decrease with the virulence. The MYXV strain *Coomandook* (grade 3.2) shows number of helminths up to three times larger than the ones in the highly virulent strains, like *Sls* and *Lausanne* (grade 1). However, the distribution of helminths abundance at host death appears to be very variable also inside each strain group, suggesting a largely heterogeneous response of the hosts.
The estimated number of helminths at host death was introduced to identify the role of co-infection dynamics in the prediction of survival time and MYXV viral load at death. The inclusion in linear models of an inverse term representing the interaction of helminths with virus load or host survival time was selected among the considered model options by the application of the Akaike Information Criterion. This addition leads to considerably better predictions, and the interacting term turned out to be a significant predictor for most of the strains. However, the optimized coefficients vary widely between strains, with consequent different features of the predictive surfaces.

The reasons behind these differences are to be found in different biological traits of the strains, different effects of MYXV on the antagonist immune responses and the impacts of different characteristics of the sampled animals with respect to the sample size. These predictive models could be further improved by testing new types of interactions and probability distributions. Larger and more detailed datasets, which are currently not available, could facilitate the identification of patterns. In conclusion, helminth co-infection appears to impact on the MYXV dynamics with a low-profile intensity and with different possible outcomes, due to diverse factors whose role still needs to be fully understood.

Bibliography

[1] P. J. Kerr and S. M. Best, "Myxoma virus in rabbits," *Revue scientifique et technique (International Office of Epizootics),* vol. 17(1), pp. 256-268, 1998.

[2] F. Fenner, M. F. Day and G. M. Woodroofe, "Epidemiological consequences of the mechanical transmission of Myxomatosis by mosquitoes," *Epidemiology & Infection,* vol. 54(2), pp. 284-303, 1956.

[3] P. J. Kerr, I. M. Cattadori, M. B. Rogers, A. Fitch, A. Geber, J. Liu, D. G. Sim, B. Bog, J.-S. Eden, E. Ghedin, A. F. Read and E. C. Holmes, "Genomic and phenotypic characterization of Myxoma virus from Great Britain reveals multiple evolutionary pathways distinct from those in Australia," *PLoS pathogens,* vol. 13(3), 2017.

[4] National Geographic Society, "How European Rabbits Took over Australia," 20 May 2022. [Online]. Available: https://education.nationalgeographic.org/resource/how-european-rabbits-took-over-australia/. [Accessed 23 September 2023].

[5] J. M. Alves, M. Carneiro, J. Y. Cheng, A. Lemos de Matos, R. M. M, L. Loog and F. M. Jiggins, "Parallel adaptation of rabbit populations to myxoma virus," *Science,* vol. 363(6433), pp. 1319-1326, 2019.

[6] M. P. Kain, I. M. Cattadori, B. M. Bolker, P. J. Kerr, I. M. Cattadori, D. Sim, J. Liu, E. C. Holmes and A. F. Read, "Divergent evolutionary pathways of Myxoma virus in Australia: virulence phenotypes in susceptible and partially resistant rabbits indicate possible selection for transmissibility," *Journal of virology,* vol. 96(20), 2022.

[7] I. D. Marshall and F. Fenner, "Studies in the epidemiology of infectious myxomatosis of rabbits. V. Changes in the innate resistance of Australian wild rabbits exposed to myxomatosis," *Epidemiology & Infection,* vol. 56(2), pp. 288-302, 1958.

[8] P. J. Kerr, J. Liu, I. Cattadori, E. Ghedin, A. F. Read and E. C. Holmes, "Myxoma virus and the Leporipoxviruses: an evolutionary paradigm," *Viruses* , vol. 7(3), pp. 1020-1061, 2015.

[9] P. J. Kerr, I. M. Cattadori, J. Liu, D. G. Sim, J. W. Dodds, J. W. Brooks, M. J. Kennet, E. C. Holmes and A. F. Read, "Next step in the ongoing arms race between Myxoma virus and wild rabbits in Australia is a novel disease phenotype," *Proceedings of the National Academy of Sciences,* vol. 114(35), pp. 9397-9402, 2017.

[10] F. Fenner and I. D. Marshall, "A comparison of the virulence for european rabbits (or Yctolagus Cuniculus) of strains of Myxoma virus recovered in the field in Australia, Europe and America," *Epidemiology & Infection,* vol. 55(2), pp. 149-191, 1957.

[11] P. J. Kerr, "Myxomatosis in Australia and Europe: a model for emerging infectious diseases," *Antiviral research,* vol. 93(3), pp. 387-415, 2012.

[12] B. M. Bolker, A. Nanda and D. Shah, "Transient virulence of emerging pathogens," *Journal of the Royal Society Interface,* vol. 7(46), pp. 811-822, 2009.

[13] G. Dwyer, S. A. Levin and L. Buttel, "A simulation model of the population dynamics and evolution of Myxomatosis," *Ecological monographs,* vol. 60(4), pp. 423-447, 1990.

[14] P. Nash, J. Barret, J. X. Cao, S. Hota-Mitchell, A. S. Lalani, H. Everett and G. McFadden, "Immunomodulation by viruses: the myxoma virus story," *Immunological Reviews,* vol. 168(1), pp. 103-120, 1999.

[15] V. G. da Costa, M. L. Morelli and M. V. Saivish, "The emergence of SARS, MERS and novel SARS-2 coronaviruses in the 21st century," *Archives of Virology,* vol. 165(7), pp. 1517-1526, 2020.

[16] A. D. Blackwell, M. Martin, H. Kaplan and M. Gurven, "Antagonism between two intestinal parasites in humans: the importance of co-infection for infection risk and recovery dynamics," *Proceedings of the Royal Society B: Biological Sciences,* vol. 280(1769), 2013.

[17] P. Devi, A. Khan, P. Chattopadhyay, P. Mehta, S. Sahni, S. Sharma and R. Pandey, "Co-infections as modulators of disease outcome: minor players or major players?," *Frontiers in Microbiology,* vol. 12, 2021.

[18] A. Fenton and S. E. Perkins, "Applying predator-prey theory to modelling immune-mediated, within-host interspecific parasite interactions," *Parasitology,* vol. 137(6), pp. 1027-1038, 2010.

[19] A. J. McArdle, A. Turkova and A. J. Cunnington, "When do co-infections matter?," *Current Opinion in Infectious Diseases,* vol. 31(3), pp. 209-215, 2018.

[20] M. Nacher, "Interactions between worms and malaria: good worms or bad worms?," *Malaria Journal,* vol. 10, pp. 1-6, 2011.

[21] World Health Organization, "Soil-transmitted helminth infections," 18 January 2013. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs366/en/. [Accessed 1 August 2023].

[22] J. Lello, B. Boag, A. Fenton, I. R. Stevenson and P. J. Hudson, "Competition and mutualism among the gut helminths of a mammalian host," *Nature,* vol. 428(6985), pp. 840-844, 2004.

[23] D. P. Benesh, G. Parker and J. C. Chubb, "Life-cycle complexity in helminths: What are the benefits?," *Evolution,* vol. 75(8), pp. 1936-1952, 2021.

[24] I. M. Cattadori, R. Albert and B. Boag, "Variation in host susceptibility and infectiousness generated by co-infection: the Myxoma–Trichostrongylus Retortaeformis case in wild rabbits," *Journal of the Royal Society Interface,* vol. 4(16), pp. 831-840, 2007.

[25] J. E. Allen and R. M. Maizels, "Th1-Th2: reliable paradigm or dangerous dogma?," *Immunology today,* vol. 18(8), pp. 387-392, 1997.

[26] A. O. G. Hoarau, P. Mavingui and C. Lebarbenchon, "Coinfection in wildlife: focus on a neglected aspect of infectious disease epidemiology," *PLoS Pathogens,* vol. 16(9), 2020.

[27] A. Fenton, "Dances with worms: the ecological and evolutionary impacts of deworming on coinfecting pathogens," *Parasitology,* vol. 140(9), pp. 1119-1132, 2013.

[28] E. Hake, "The impact of co-infections with medium and high virus virulence on the length and fecundity of Trichostrongylus Retortaeformis," Penn State University , 2017.

[29] M. P. Kain, I. M. Cattadori and B. M. Bolker, "The evolutionary response of virulence to host heterogeneity: a general model with application to Myxomatosis in rabbits co-infected with intestinal helminths," *Evolutionary Ecology Research,* vol. 19(3), pp. 257-278, 2018.

[30] E. Vynnyck and R. White, An introduction to infectious disease modelling, Oxford: OUP oxford, 2010.

[31] A. Fenton, "Worms and germs: the population dynamic consequences of microparasite-macroparasite co-infection," *Parasitology* , vol. 135(13), pp. 1545-1560, 2007.

[32] N. C. Grassly and C. Fraser, "Mathematical models of infectious disease transmission," *Nature Reviews Microbiology* , vol. 6(6), pp. 477-487, 2008.

[33] P. Kerr and G. McFadden, "Immune responses to myxoma virus," *Viral Immunology,* vol. 15(2), pp. 229-246, 2002.

[34] J. E. Cavanaugh and A. A. Neath, "The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements," *Wiley Interdisciplinary Reviews: Computational Statistics,* vol. 11(3), 2019.

[35] J. D'Errico, "fminsearchbnd," MATLAB Central File Exchange , 2020.

[36] T. J. DiCiccio and B. Efron, "Bootstrap confidence intervals," *Statistical Science,* vol. 11(3), pp. 189-228, 1996.

[37] M. Martcheva, N. Tuncer and C. St Mary, "Coupling within-host and between-host infectious diseases models," *Biomath,* vol. 4(2), 2015.

[38] K. M. Tjørve and E. Tjørve, "The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family," *PloS one,* vol. 12(06), 2017.

[39] S. Sakanoue, "Extended logistic model for growth of single-species populations," *Ecological Modelling,* Vols. 205(1-2), pp. 159-168, 2007.

[40] Y. Pei, G. Zeng and L. Chen, "Species extinction and permanence in a prey–predator model with two-type functional responses and impulsive biological control," *Nonlinear Dynamics* , vol. 52, pp. 71-81, 2008.

[41] C. Vanalli, L. Mari, R. Casagrandi, M. Gatto, B. Boag and I. M. Cattadori, "Modeling the contribution of antibody attack rates to single and dual helminth infections in a natural system," *Mathematical Biosciences,* vol. 360, 2023.

[42] C. F. J. Wu, "Jackknife, bootstrap and other resampling methods in regression analysis," *The Annals of Statitics,* vol. 14(4), pp. 1261-1295, 1986.

[43] I. M. Cattadori, B. Boag and P. J. Hudson, "Parasite co-infection and interaction as drivers for host heterogeneity," *International journal for Parasitology,* Vols. 38(3-4), pp. 371-380, 2008.

[44] J. M. Bland and D. G. Altman, "The logrank test," *Bmj,* vol. 328(7447), 2004.

[45] R. J. Freund and W. J. Wilson, Statistical methods, Elsevier, 2003.

[46] W. Haynes, "Wilcoxon Rank Sum Test," in *Encyclopedia of Systems Biology*, New York, Springer, 2013, pp. 2345-2355.

[47] N. Feltovich, "Nonparametric tests of differences in medians: comparison of the Wilcoxon–Mann–Whitney and robust rank-order tests," *Experimental Economics,* vol. 6, pp. 273-297, 2003.

[48] F. Audebert, J. Cassone, D. Kerboeuf and M. C. Durette-Desset, "The life cycle of Nematodiroides zembrae (Nematoda, Trichostrongylina) in the rabbit," *Journal of Parasitology ,* vol. 88(5), pp. 898-904, 2002.

[49] N. Khatun, "Applications of normality test in statistical analysis," *Open Journal of Statistics ,* vol. 11(01), p. 113, 2021.

[50] A. Saumard and F. Navarro, "Finite sample improvement of Akaike's Information Criterion," *IEEE Transactions on Information Theory,* vol. 67(10), pp. 6328-6343, 2021.

A Appendix A: Confidence intervals of the myxoma virus growth model parameters

This appendix contains, in Table A-1, the 5%-95% confidence intervals of the parameters of the growth model with the lowest AIC score for each strain analyzed in chapter 4.1. Each confidence interval was obtained with a bootstrap approach: the data of each strain were randomly resampled 100 times and the model calibration was repeated for each obtained series of data. From the results the 5% - 95% confidence intervals were generated and the MYXV growth was simulated 100 times to obtain the confidence intervals of the simulations shown in chapter 4.1.4.

Table A-1: Bootstrap 5%-95% confidence interval of the parameters of the model with lowest AIC score, between Malthusian model (V_0, r), logistic model (V_0, r, K) and model with immune response (V_0, r, a or V_0, r, d, a)

Strain (virulence grade)	Dataset	V_0 [[V]] 5%-95% CI	r [$\frac{1}{days}$] 5%-95% CI	K [[V]] 5%-95% CI	d [$\frac{1}{days[V]}$] 5%-95% CI	a [$\frac{1}{days^2[V]}$] 5%-95% CI
Swh 8/2/93 (2)	Kerr	6.143E+06 1.929E+07	0.182 0.291	-	-	-
Ws6 346 (2)	Kerr	1.559E+07 6.870e+07	0.060 0.209	-	-	-
Bd23 (1-2)	Kerr	4.909E+07 3.060E+8	0.253 1.172	1.540E+08 3.060E+08	-	-
Swh 1209 (3)	Kerr	3.248E+07 5.610E+07	0.111 0.152	-	-	-
Brk 12/2/93 (1-2)	Kerr	4.074E+04 3.987E+07	0.143 1.390	1,445E+08 5.987E+08	-	-
Km13 5)	Kerr	4.915E+06 4.050E+07	0.086 0.258	-	-	1.773E-13 9.004E-11
Uriarra (5)	Kerr	1.862E+06 5.987E+08	1.346E-14 0.459	-	-	2.643E-24 7.282E-10
Sls 1)	Fenner	4.585E+03 9.277E+05	0.627 2.384	2.847E+07 1.118E+08	-	-
Km13 II (3)	Fenner	1.200E+05 1.627E+06	0.324 0.740	-	-	-
Uriarra III (3-4)	Fenner	1.024E+05 6.806E+05	0.724 1.117	3.314E+08 1.119E+09	-	-
Loiret 55 (4)	Fenner	2.439E+08 7.301E+05	0.398 0.887	-	-	-
Lausanne (1)	Fenner	6.713 1.256+05	1.105 3.781	5.016E+07 2.690E+08	-	-
Neuromyxoma (5)	Fenner	3.443E-04 5.608E+05	1.001 9.469	-	4.534E-07 6.758E-06	2.814E-07 2.909E-06

B Appendix B: Statistical tests on survival time and viral load at death of co-infected and single-infected rabbits, including survivors

This appendix contains a series of statistical tests to compare the survival curves and the averages, medians, and variances of survival time (*ST*) and viral load at death ($\log_{10}(VLD)$) between rabbits co-infected with *Myxoma virus* (MYXV) and *Trichostrongylus retortaeformis* and rabbits single-infected with *Myxoma virus* ('co-infection dataset', section 3.1.4). All tests were performed in accordance with the methods discussed in chapter 5.1, with the only difference that here the survivors of the 30-days experiment are included. The experimental percentage of survivors of each strain is reported in Table B-1. The *ST* of the survivors was set to 60 days (averaging their fate after the end of the experiment) whereas their *VLD* was set equal to the virus amount measured the 30th day, assuming the end of the experiment as death condition. Single- and co-infected rabbits are compared in Figure B-1 and the results of the tests are reported in Table B-2, Table B-3 and Table B-4.

Table B-1: Comparison of experimental percentages of dead hosts, between MYXV-helminths co-infected rabbits and MYXV single-infected ones.

Strain	Virulence grade	Sampled survival percentage (without helminths)	Sampled survival percentage (with helminths)
Coomandook	3.2	25%	37.5%
Perthshire 2082	3.1	0%	0%
Ws 61071	3.1	12.5%	12.5%
Lausanne	1	0%	0%
Sls	1	30%	20%

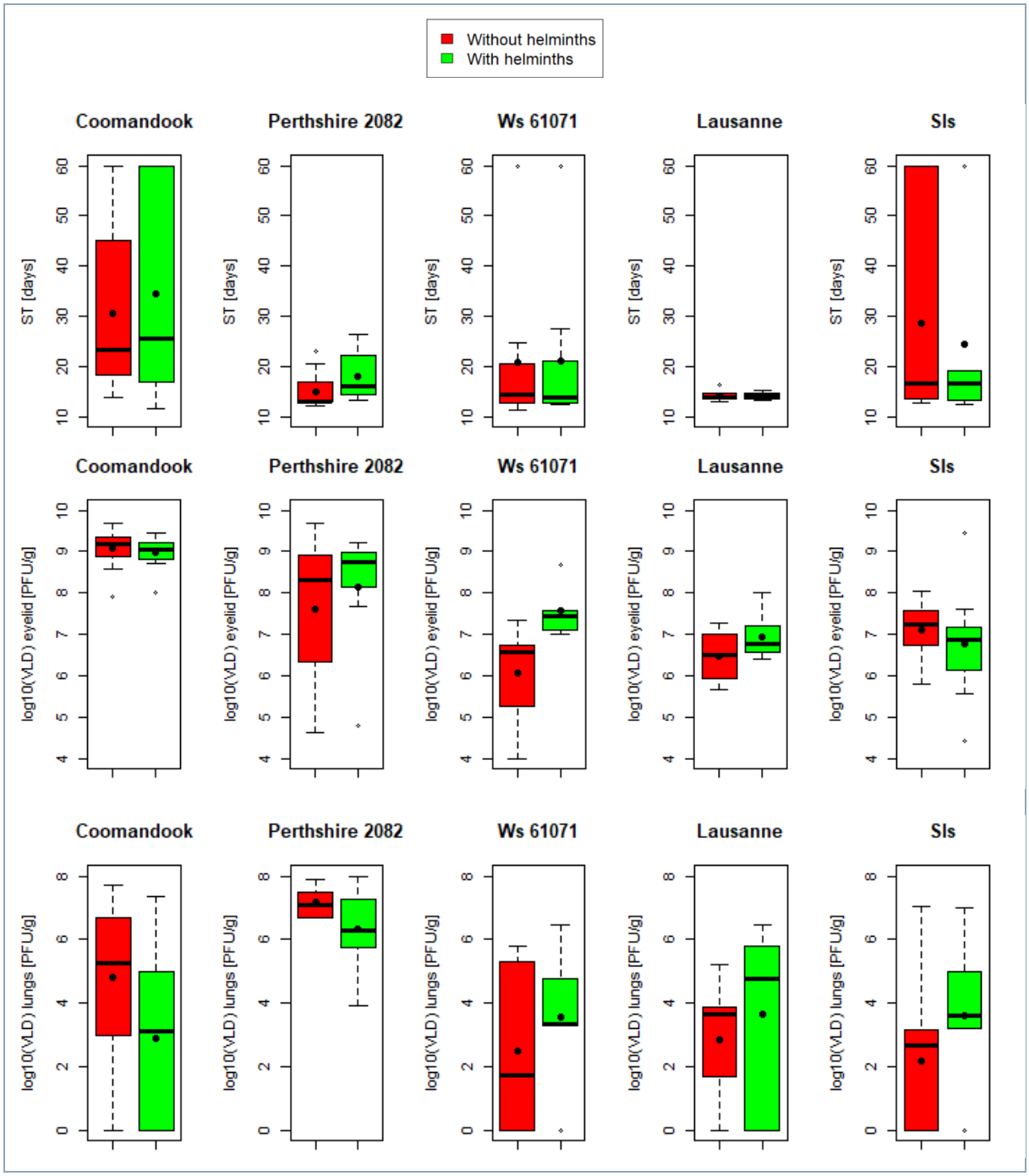


Figure B-1: Boxplots to compare the viral load at death of MYXV-helminths co-infected rabbits and MYXV single-infected rabbits for each strain, excluding the survivors. Each box is built with the minimum non-outlier, the first quartile, the median (black line), the average (black dot), the third quartile and the maximum non-outlier.

Table B-2: Performances of log-rank tests on survival curves, t-tests on ST averages, Wilcoxon tests on ST populations and medians and F-tests on ST variances (α=0.05), between MYXV-helminths co-infected rabbits and MYXV single-infected ones including survivors.

Strain	Average ST [days] without helminths	Average ST [days] with helminths	t-test pvalue	Median ST [days] without helminths	Median ST [days] with helminths	Wilcoxon test pvalue	Variance ST [days^2] without helminths	Variance ST [days^2] with helminths	F-test pvalue
Comandook (3.2)	30.613	34.383	0.591	23.210	25.500	0.924	473.390	357.429	0.575
Perthshire 2082 (3.1)	15.074	18.070	0.129	13.065	16.005	0.041	0.893	0.445	0.623
Ws 61071 (3.1)	20.848	21.000	0.985	14.330	13.940	1.00	267.674	273.289	0.979
Lausanne (1)	14.128	14.169	0.912	13.935	14.130	0.650	17.571	25.866	0.314
Sls (1)	28.611	24.388	0.649	16.730	16.575	0.849	328.688	441.485	0.682

Table B-3: Performances of t-tests on $\log_{10}$(VLD) averages, Wilcoxon tests on $\log_{10}$(VLD) populations and medians and F-tests on $\log_{10}$(VLD) variances between MYXV-helminths co-infected rabbits and MYXV single-infected ones, in the eyelid (α=0.05), including survivors.

Strain	Average $\log_{10}$(VLD) [PFU/g] without helminths	Average $\log_{10}$(VLD) [PFU/g] with helminths	t-test pvalue	Median $\log_{10}$(VLD) [PFU/g] without helminths	Median $\log_{10}$(VLD) [PFU/g] with helminths	Wilcoxon test pvalue	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ without helminths	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ with helminths	F-test pvalue
Comandook (3.2)	9.084	8.993	0.535	9.178	9.062	0.371	0.185	0.127	0.502
Perthshire 2082 (3.1)	7.594	8.136	0.573	8.307	8.745	0.610	3.685	2.430	0.626
Ws 61071 (3.1)	6.053	7.566	0.020	6.557	7.447	0.010	1.246	0.435	0.326
Lausanne (1)	6.474	6.952	0.168	6.500	6.761	0.272	0.370	0.351	0.988
Sls (1)	7.097	6.763	0.577	7.224	6.860	0.397	0.583	1.925	0.163

Table B-4: Performances of t-tests on $\log_{10}$(VLD) averages, Wilcoxon tests on $\log_{10}$(VLD) populations and medians and F-tests on $\log_{10}$(VLD) variances between MYXV-helminths co-infected rabbits and MYXV single-infected ones, in the lungs (α=0.05), including survivors.

Strain	Average $\log_{10}$(VLD) [PFU/g] without helminths	Average $\log_{10}$(VLD) [PFU/g] with helminths	t-test pvalue	Median $\log_{10}$(VLD) [PFU/g] without helminths	Median $\log_{10}$(VLD) [PFU/g] with helminths	Wilcoxon test pvalue	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ without helminths	Variance $\log_{10}$(VLD) $[(PFU/g)]^2$ with helminths	F-test pvalue
Comandook (3.2)	4.806	2.902	0.051	5.248	3.101	0.090	4.976	8.299	0.342
Perthshire 2082 (3.1)	7.164	6.318	0.171	7.102	6.301	0.283	0.233	1.920	0.035
Ws 61071 (3.1)	2.482	3.577	0.478	1.739	3.328	0.707	7.510	5.697	0.833
Lausanne (1)	2.845	3.638	0.574	3.629	4.785	0.346	4.119	8.349	0.415
Sls (1)	2.154	3.594	0.217	2.653	3.605	0.150	5.786	5.485	0.942

B | Appendix B: Statistical tests on survival time and viral load at death of co-infected and single-infected rabbits, including survivors

The results highlight that the inclusion of survivors' does not substantially change the outcomes of the tests. The application of statistical tests does not show significant evidence to distinguish single-infected and co-infected cases, with only a few cases rejecting the hypotheses of similarity.

C Appendix C: Tests of viral load at death of single-infected and co-infected rabbits with outliers Jackknife

This appendix shows the results of statistical tests on averages, medians, and variances of survival time (*ST*) and viral load at death ($\log_{10}(VLD)$) between rabbits co-infected with *Myxoma virus* (MYXV) and *Trichostrongylus retortaeformis* and rabbits single-infected with *Myxoma virus* ('co-infection dataset', section 3.1.4) applying a Jackknife technique [42] on the outliers. The tests of section 5.1.3 were repeated removing the outliers one at a time and the best obtained scores were selected. The outlier identification was performed over the entire dataset of each strain, without distinguishing between co-infected and single-infected rabbits and excluding the surviving rabbits. A data is classified as outlier if it is included in the set *O*, as defined below, where Q_1 is the first quartile, Q_3 the third quartile and IQR is the interquartile range of the dataset of the strain.

$$log_{10}(VLD_i) \in O \; if \; log_{10}(VLD_i) < Q_1 - 1.5\, IQR \cup log_{10}(VLD_i) > Q_3 + 1.5\, IQR \quad (C.1)$$

A visual identification of the outliers is shown in Figure C-1. Table C-1 and Table C-2 report the results of the Jackknife technique applied on the outliers for eyelid and lungs. The comparison with the results of chapter 5.1 show that some individual data consistently influence the tests. After removing the outliers, more cases approach significant p-values, but the overall results still do not detect a clear effect caused by co-infection.

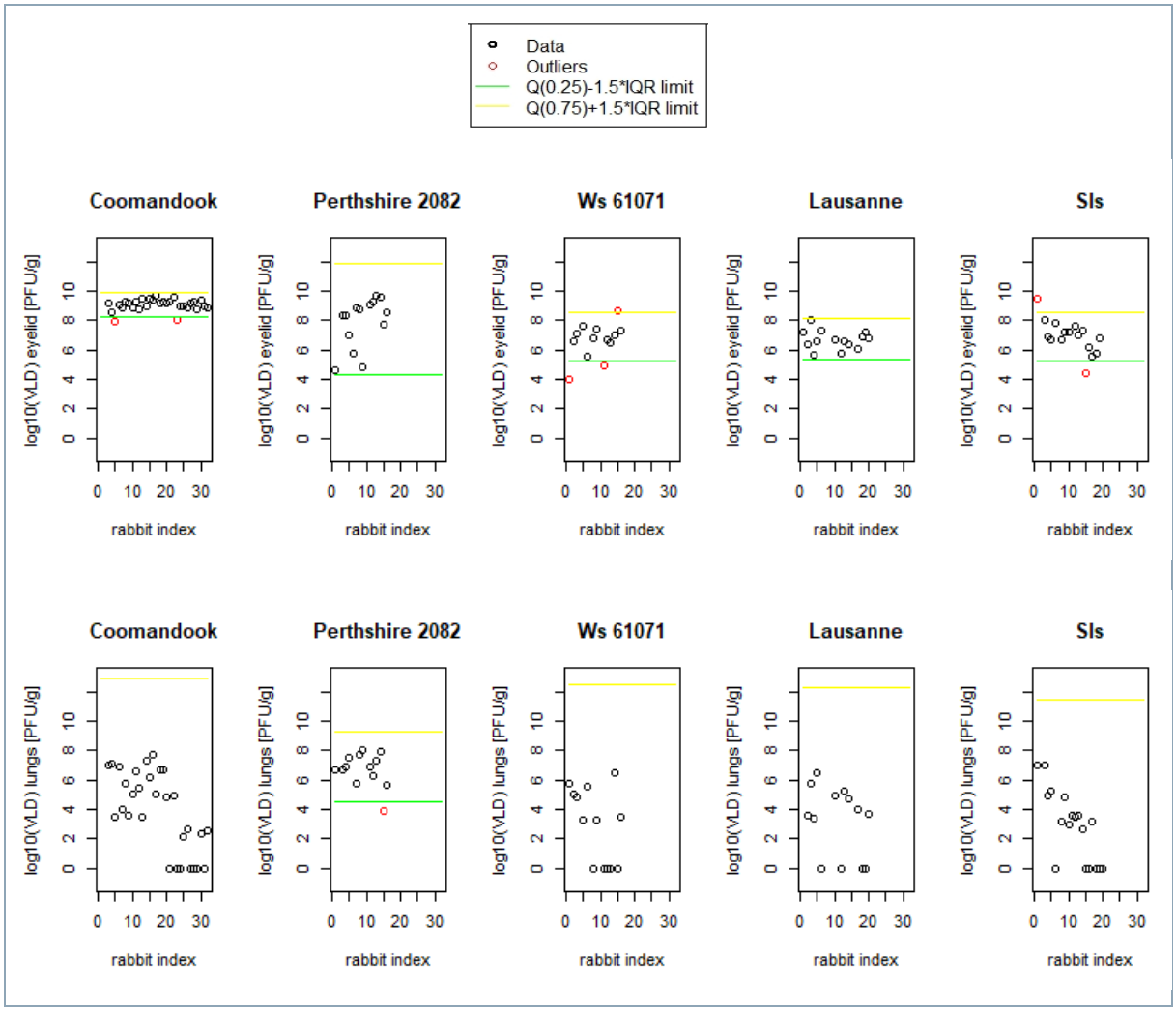


Figure C-1: Identification of viral load at death outliers for MYXV-helminths co-infected and MYXV single-infected rabbits for different strains.

Table C-1: Best outliers Jackknife performances of t-tests on $\log_{10}$(VLD) averages, Wilcoxon tests on $\log_{10}$(VLD) populations and medians and F-tests on $\log_{10}$(VLD) variances, between MYXV-helminths co-infected rabbits and MYXV single-infected ones, in the eyelid (α=0.05).

Strain	Virulence grade	t-test pvalue	Wilcoxon test pvalue	F-test pvalue
Coomandook	3.2	0.396	0.230	0.455
Perthshire 2082	3.1	0.190	0.284	0.014
Ws 61071	3.1	0.017	0.014	0.040
Lausanne	1	0.143	0.213	0.218
Sls	1	0.179	0.183	0.127

Table C-2: Best outliers Jackknife performances of t-tests on $\log_{10}$(VLD) averages, Wilcoxon tests on $\log_{10}$(VLD) populations and medians and F-tests on $\log_{10}$(VLD) variances, between MYXV-helminths co-infected rabbits and MYXV single-infected ones, in the lungs (α=0.05).

Strain	Virulence grade	t-test pvalue	Wilcoxon test pvalue	F-test pvalue
Coomandook	3.2	0.021	0.036	0.302
Perthshire 2082	3.1	0.072	0.092	0.020
Ws 61071	3.1	0.212	0.437	0.353
Lausanne	1	0.809	0.620	0.373
Sls	1	0.058	0.044	0.391

D Appendix D: Correlation analysis of survival time, viral load at death and helminth abundance at death of co-infected rabbits

This appendix contains the correlation matrices of survival time (*ST*), viral load at death ($\log_{10}$(*VLD*)) and number of helminths at death (*NHD*) of *Myxoma virus* (MYXV) and *Trichostrongylus retortaeformis* co-infected hosts, divided by strain ('co-infection dataset', section 3.1.4). The results, reported in Table D-1, Table D-2, Table D-3, Table D-4 and Table D-5 highlight a strong variability between the strains under analysis in subchapter 5.2.

Table D-1: Correlation matrix of survival time, viral load at death (eyelid and lungs) and abundance of helminths at death of MYXV(*Coomandook*)-helminths co-infected rabbits.

Correlation ϱ	ST	$\log_{10}$(VLD) eyelid	$\log_{10}$(VLD) lungs	NHD
ST	-	0.460	-0.410	-0.187
$\log_{10}$(VLD) eyelid	0.460	-	-0.163	0.088
$\log_{10}$(VLD) lungs	-0.410	-0.163	-	0.159
NHD	-0.187	0.088	0.159	-

Table D-2: Correlation matrix of survival time, viral load at death (eyelid and lungs) and abundance of helminths at death of MYXV(*Perthshire 2082*)-helminths co-infected rabbits.

Correlation ϱ	ST	log_{10}(VLD) eyelid	log_{10}(VLD) lungs	NHD
ST	-	0.126	-0.720	0.895
log_{10}(VLD) eyelid	0.126	-	-0.302	0.380
log_{10}(VLD) lungs	-0.720	-0.302	-	-0.908
NHD	0.895	0.380	-0.908	-

Table D-3: Correlation matrix of survival time, viral load at death (eyelid and lungs) and abundance of helminths at death of MYXV(*Ws 61071*)-helminths co-infected rabbits.

Correlation ϱ	ST	log_{10}(VLD) eyelid	log_{10}(VLD) lungs	NHD
ST	-	0.474	-0.240	-0.022
log_{10}(VLD) eyelid	0.474	-	-0.964	0.578
log_{10}(VLD) lungs	-0.240	-0.964	-	-0.695
NHD	-0.022	0.578	-0.695	-

Table D-4: Correlation matrix of survival time, viral load at death (eyelid and lungs) and abundance of helminths at death of MYXV(*Lausanne*)-helminths co-infected rabbits.

Correlation ϱ	ST	log_{10}(VLD) eyelid	log_{10}(VLD) lungs	NHD
ST	-	-0.235	-0.955	0.113
log_{10}(VLD) eyelid	-0.235	-	-0.048	0.445
log_{10}(VLD) lungs	0.955	-0.048	-	-0.020
NHD	0.113	-0.445	-0.020	-

Table D-5: Correlation matrix of survival time, viral load at death (eyelid and lungs) and abundance of helminths at death of MYXV(*Sls*)-helminths co-infected rabbits.

Correlation ϱ	ST	log_{10}(VLD) eyelid	log_{10}(VLD) lungs	NHD
ST	-	-0.583	-0.877	-0.679
log_{10}(VLD) eyelid	-0.583	-	0.875	0.646
log_{10}(VLD) lungs	-0.877	0.875	-	0.787
NHD	-0.679	0.646	0.787	-

E Appendix E: Gaussian family distributions of survival time and viral load at death

This appendix provides some details on the selection of the family distribution of the response necessary for prediction analyses of survival times and viral load at death of rabbits co-infected with *Myxoma virus* (MYXV) and *Trichostrongylus retortaeformis* ('co-infected dataset', section 3.1.4), performed in subchapter 5.2. A Gaussian family was chosen for this task after the performance of a Lillie test [49], with null hypothesis H_0 that data comes from a distribution of a normal family and with a level of significance α of 0.05. The results are reported in Table E-1.

Table E-1: Lillie test performances on survival times and viral load at death of MYXV-helminths co-infected rabbits, for different strains.

Strain	ST [days] Lillie test pvalue	Log_{10}(VLD) [PFU/g] eyelid Lillie test pvalue
Coomandook	0.323	0.299
Perthshire 2082	0.013	0.022
Ws 61071	2.252E-4	0.073
Lausanne	0.051	0.897
Sls	0.155	0.317

It can be observed that, in most of the cases, there is no sufficient evidence to refuse the hypothesis of a normal distribution, with some exceptions. Nevertheless, the aim of the prediction analysis was not to build the most precise predictive model, but to use response modelling as a variable selection method to understand the role of different agents in the mechanism. For this reason, the gaussian distribution was however selected as family for the model, to exploit its simplicity and adaptability.

F Appendix F: Predictions of survival time and viral load with modified helminth abundance at host death

This appendix further investigates the structures the models for survival time and *Myxoma virus* (MYXV) viral load at death developed in chapter 5.2, by modifying the predictor element representing the abundance of helminths. Two types of modifications were considered. First, a base 10 logarithmic transformation was applied to number of helminths, bringing the value scale closer to the MYXV values expressed as $\log_{10}$(VLD). Second, the number of helminths was corrected by modulating the abundance with the number of doses administered to the rabbits before death. The helminth infections of experiments with strains *Ws 61071* and *Perthshire 2082* occurred on days -17, -10, -3, +4, +11 and +18 in relation to MYXV infection, while for strains *Coomandook*, *Lausanne* and *Sls* only on days -17, -10, -3, +4 and +11, with the methodologies described in section 3.1.4. The helminth abundance was corrected as follows, where *numd* is the number of doses injected to the host.

$$NHD_{corr} = \frac{NHD}{numd} \qquad (F.\ 1)$$

These two modifications were applied, in the series of equations below, only for the model structure with inverse interaction identified as the best option in chapter 5.2.

$$log_{10}(VLD) = a + b\,ST + c\,log_{10}(NHD) + d\,\frac{ST}{log_{10}(NHD)} \qquad (F.\ 2)$$

$$log_{10}(VLD) = a + b\,ST + c\,NHD_{corr} + d\,\frac{ST}{NHD_{corr}} \qquad (F.\ 3)$$

$$ST = e + f\,log_{10}(VLD) + g\,log_{10}(NHD) + h\,\frac{log_{10}(VLD)}{log_{10}(NHD)} \qquad (F.\ 4)$$

$$ST = e + f\ log_{10}(VLD) + g\ NHD_{corr} + h\ \frac{log_{10}(VLD)}{NHD_{corr}} \qquad (F.5)$$

Table F-1 and Table F-2 show, respectively for log10(*VLD*) and *ST* models, the comparison of the AIC scores between the original structures and the modified models.

Table F-1: Results of log10(VLD) predictive models AIC selection with modified number of helminths, with the lowest AIC score marked in bold.

Strain	Grade	Model (5. 5) AIC	Model (F. 2) AIC	Model(F. 3) AIC
Coomandook	3.2	**1.581**	2.403	**1.581**
Perthshire 2082	3.1	**7.412**	18.107	8.363
Ws 61071	3.1	**-14.491**	-6.198	-8.664
Lausanne	1	1.457	**-9.514**	1.457
Sls	1	**26.294**	27.958	**26.294**

Table F-2: Results of ST predictive models AIC selection with modified number of helminths, with the lowest AIC score marked in bold.

Strain	Grade	Model (5. 6) AIC	Model (F. 4) AIC	Model (F. 5) AIC
Coomandook	3.2	**50.718**	50.738	**50.718**
Perthshire 2082	3.1	**12.843**	40.754	28.292
Ws 61071	3.1	24.357	33.126	**17.225**
Lausanne	1	**5.395**	6.820	**5.395**
Sls	1	**17.121**	34.287	**17.121**

In only one case, $\log_{10}(VLD)$ model of *Lausanne*, the AIC score is lower with the number of helminths in log-form. This transformation therefore does not generally improve predictions. All the rabbits infected with *Coomandook*, *Lausanne* and *Sls* die after the same number of helminth doses, so the correction made to the abundance of helminths is irrelevant and results only in a coefficient scaling. The AIC scores change only in the

cases of *Perthshire 2082* and *Ws 61071*, for which some rabbits die after five doses of helminths and some others after six, but again this does not cause a clear improvement of performances.

List of Figures

List of Tables

Glossary

- AIC: Akaike Information Criterion
- ANU: Australian National University
- AST: average survival time
- CFR: case fatality rate
- CIDD: Center for Infectious Diseases
- DDPI: death day post infection
- DNA: deoxyribonucleic acid
- DPI: day post infection
- HIV: Human Immunodeficiency Virus
- ID: identification number
- IQR: interquartile range
- JCSMR: John Curtin School of Medical Research
- MYXV: myxoma virus
- NHD: number of helminths at death
- PBS: phosphate buffered saline
- PCR: polymerase chain reaction
- PFU: plaque forming units
- PSU: Pennsylvania State University
- qPCR: quantitative polymerase chain reaction
- RID: rabbit infectious doses
- ST: survival time
- Th1: type 1 helper T cell
- Th2: type 2 helper T cell
- VLD: viral load at death