DYNAMICAL MODELING AND DATA CLUSTERING WITH APPLICATIONS TO TICK POPULATION DYNAMICS AND FINANCE

JEMISA SADIKU

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS
YORK UNIVERSITY
TORONTO, ONTARIO

December 2020

Abstract

Mathematical models are valuable tools that assist to control the spread of infected ticks and analyze tick-host dynamics. In this work, a compartmental model is created to study the dynamics of lone star ticks and white-tailed deers, with a focus on migration effect of white-tailed deers. The second model focuses on the effect of grooming behaviour of the host on tick dynamics. Lastly, a data set is separated using K-means and Birch clustering techniques and the silhouette score has been evaluated.

Acknowledgements

I would like to express my very great appreciation to my supervisor Dr. Jianhong Wu for supporting me during my journey as a PhD student at York University. In addition, I wish to acknowledge the help provided by my committee members Dr. Jane Heffernan, Dr. Huaiping Zhu. My special thanks are extended to Mahnaz Alavinejad, Dr Zilong Song and Zarko Valtchev for their contributions on model formulation and analysis.

Table of Contents

List of Tables

List of Figures

1 Introduction

This thesis consists of two parts: biological modelling of ticks population dynamics and clustering analysis of time series of Bitcoin. It is hoped that the time series data clustering technique can be applied to analyze the time series of tick populations at different development stages and the time series of tick-borne disease incidence data in different geographical locations when quality data become available.

Due to climate change in general, and the increase of favourable habitats as a result of climate warming, ticks and tick-borne diseases are reported to expand to northern areas in north America. A large number of mathematical models and analyses have been developed in the last few decades, we refer to the recent monograph by Wu and Zhang (Transmission Dynamics of Tick-Borne Diseases with Co-Feeding, Developmental and Behavioural Diapause, Springer Nature, 2020) for relevant references.

In Chapter 2, we will consider the range expansion of lone star ticks in Canada due to the mobility of their hosts, white-tailed deers. In particular, we will formulate a deterministic compartmental model-a system of ordinary differential equations, to study the

population dynamics of lone star ticks and white-tailed deers, with a focus on effect of white-tailed deers mobility. We will use this model to explore the tick-host interaction and the effect of deer migration, analytically and numerically. We will first prove the positivity of solutions, as a first step to theoretically verify the well-posedness of the mathematical model. We will establish the existence and uniqueness of a positive equilibrium, and then prove this positive equilibrium is locally asymptotically stable. This justifies the focus on the dependence of the positive equilibrium on parameters influenced by climate change, and by potential mitigation measures. For this purpose, we conduct a sensitivity analysis on a set of parameters, revealing the correlation between the parameters and equilibrium populations. Numerical results show that the migration rate of white-tailed deer is one crucial parameter that increases the populations of (infected) ticks and (infected) hosts.

Our next focus is an important issue that has been hardly addressed using mathematical modelling. Namely, for a variety of tick species, the resistance, a behavioural and immunological response of hosts has been reported in the biological literature but its impact on tick population dynamics has not been mathematically formulated and analyzed using dynamical models reflecting the full biological development of ticks. In Chapter 3, we develop and simulate a delay differential equation model, with a particular focus on resistance resulting in grooming behaviour. We calculate the threshold parameter using the spectral analyses of delay differential equations with positive feedback, and establish

the existence and uniqueness of a positive equilibrium when the threshold parameter exceeds the unity. We also conduct numerical and sensitivity analyses about the dependence of this positive equilibrium on the parameters relevant to grooming behaviour. We numerically obtain the relationship between grooming behaviour and equilibrium value at different stages.

The thesis also applies various clustering techniques to volatility and market sentiment measures of historical Bitcoin prices in order to identify hidden structural patterns by separating the data into different classes known as regimes. Regimes are a common indicator for traders and asset managers to look at when deciding what actions they want to take. For instance, hedge funds might avoid trading in highly volatile times where-as option traders might prefer it. To separate the data, we perform K-means,Ward,Birch and complete-linkage Agglomerative clustering algorithms. We then analyze the performance of each clustering algorithm.

2 Modeling of Lone Star Ticks with Deer Migration to Canada

Lone star tick (Amblyomma americanum) is recognized for the first time by Linnaeus in 1758 [1]. It prefers damp forests and humid soil environment [2, 3]. Similar to other ticks, it has four life stages: egg, larvae, nymph and adult, and the transition from one stage to the next is done through questing, feeding and molting [4, 5, 6].

The associated tick-borne diseases and the allergic reaction have received a great attention on lone star tick in recent years. Lone star tick is a vector that can carry pathogens and transmit tick-borne diseases such as STARI (Southern tick-associated rash illness) [7, 8] and Human Monocytic Ehrlichiosis (HME) [9, 1, 5, 10]. The preferred host (here white-tailed deer) plays an important role in disease transmission, as it serves as a reservoir for the pathogens. Unlike other species of ticks, lone star tick exhibits *aggressive* [8, 3], which makes bites to humans more likely. In addition, lone star tick causes red meat allergy (delayed-onset allergy) [11, 12], first discovered in 2009 [13]. The allergic reaction

could be fatal [5, 14, 15], and is due to immunoglobulin E (IgE) antibody which is specific for galactose-α-1,3-galactose (alpha-gal). It is claimed that the cause of IgE antibodies is primary from bites of lone star tick [14].

The lone star tick is mostly found in the eastern, southeastern and south-central states of US [16]. However, its distribution and abundance have increased over the past decades, and according to the distribution map at the CDC, lone star tick has expanded to more northern and western areas in North America where it was absent in previous years [16]. In particular, it has migrated from endemic areas in the US to new regions in Southern Canada, such as British Columbia, Ontario, south-eastern Manitoba [17, 18, 5].

Two main factors responsible for the range expansion of lone star ticks are the climate change and the migration of white-tailed deer [16]. Due to climate change, more regions become inhabitable for lone star ticks. Since ticks are small species, they will spread to new regions mainly by migration of their host mammals. White-tailed deer is the main host for lone star tick of all life stages to get the majority of their blood meals, and their populations are positively correlated [17, 19, 4, 5]. The deer migration depends on the deer habitat suitability of the region, which is influenced by two factors: the need for food, shelter and water, and the disturbances from human activities (such as hunting) [17]. In order to better understand or predict the outbreaks of tick-borne diseases, there is a need to include the migration effect into the tick-host dynamics.

Various models have been proposed to investigate tick-host dynamics and to address a variety of issues on the tick-borne diseases [20, 21]. Specifically for Lone star ticks, there have been computer simulations based on age-structured difference equations [22, 23], agent-based models [24], predicative statistical models [25]. Recently, a metapopulation model with migration effect among patches and logistic-type birth term is adopted in [26, 27] to study the HME transmission and investigate various tick-control strategies. For another tick-borne disease (Lyme disease) [21], the range expansion of ticks and pathogens has been widely studied by various migration effects including movements of rodents and deer [28] as well as bird migration [29]. The migration of white-tailed deer has also been included in a single patch model with distributed delay (integral) term [30], which models that the deer travels out and then back in the patch. Since the logistic-type birth term may cause negative birth rate at high population density (which could occur with migration effect), other positive density-dependent birth rates have been adopted in modeling of ticks [6, 31].

In this work, we formulate a compartment model, where the hosts and ticks have been divided into susceptible and infected compartments. Our model adopts the Ricker function as the birth term to ensure positivity, which was used in a stage-structured model [31]. It also includes a simple migration term in the dynamics of hosts, with a focus to study the migration effects of white-tailed deer from US to Canada. The positivity of the pop-

ulations in the model is proved for the model with biologically meaningful parameters. With migration effect as a source term, there exists a unique positive equilibrium, which is asymptotically stable. Our numerical results confirm these features and a sensitivity analysis is carried out to reveal the effects of parameters. The migration rate is found to be a crucial parameter which will increase the populations of (infected) ticks and (infected) hosts.

The rest of the chapter is arranged as follows. The mathematical model is proposed in Section 2.1. After non-dimensionalization, Section 2.2 proves the positivity of the solutions and the stability of the unique positive equilibrium. In Section 2.3, numerical results and sensitivity analyses are carried out, and the effects of parameters including migration rate are studied. Finally, conclusions and future extensions are provided in Section 2.4.

2.1 The Mathematical Model

Given that lone star tick prefers white-tailed deer as the main host for all life stages [26], we formulate a mathematical model with a single host. We neglect the life stages of tick, and restrict ourselves to a single life stage, although the stage structure will be included in a later study. For simplicity, seasonality [32] has been ignored in the model. We concentrate on a single patch (region in Canada), and there is migration of host from the outside environment (US) where the host population is assumed as a constant.

Table 2.1: Definition and values of parameters

Symbol	Meaning	Value and Reference
b	Maximal birth rate of host	0.2/month [33, 26]
q	Strength of density dependence for host birth	1/100 [est]
β	Transmission rate from infected tick to susceptible host	0.02/month [26, 27]
μ	Death rate of host	0.01/month [26, 33]
m	Migration rate from outside region to current region	0.01/month [27]
$\hat{b}$	Maximal birth rate of ticks	0.75/month [26]
$\hat{q}$	Strength of density dependence for tick birth	1/20000 [27]
$\hat{\mu}$	Death rate of ticks	0.1/month [26]
$\hat{\beta}$	Transmission rate from infected host to susceptible tick	0.07/month [26, 27]
N_a	Average number of ticks per host	200 [26, 27]
r_s	Ratio of infected ticks over all ticks on susceptible host	0.05 [est]
r_I	Ratio of infected ticks over all ticks on infected host	0.1 [est]
H_s^o	Population of susceptible host in the outside region	800 [est]
H_I^o	Population of infected host in the outside region	200 [est]

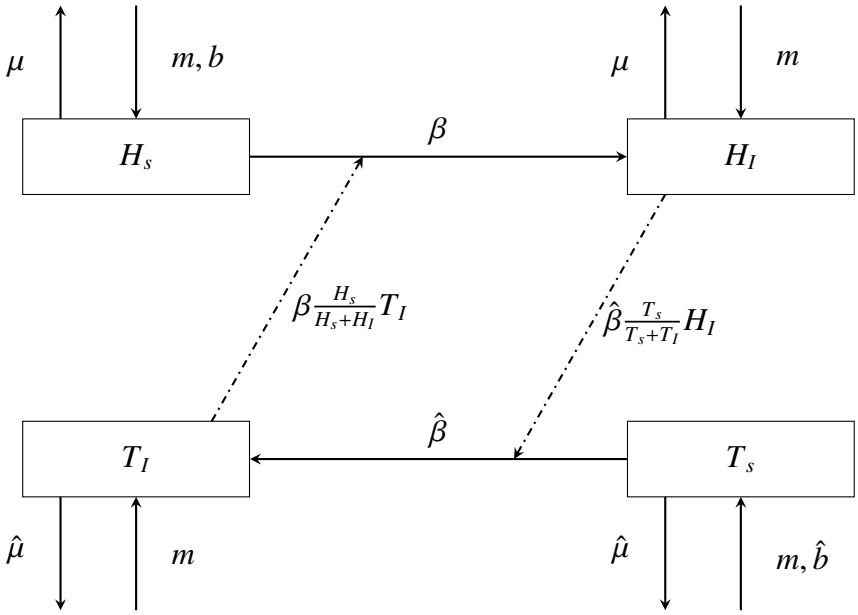


Figure 2.1: Flow chart for the compartmental model.

The susceptible host and susceptible tick are denoted by H_s and T_s, and the infected host and infected tick are denoted by H_I and T_I. A flow chart is depicted in Figure 2.1. Correspondingly, the model is given by

$$\frac{dH_s}{dt} = b(H_s + H_I)e^{-q(H_s+H_I)} - \mu H_s - \beta\frac{H_s}{H_s + H_I}T_I + mH_s^o, \tag{2.1a}$$

$$\frac{dH_I}{dt} = \beta\frac{H_s}{H_s + H_I}T_I - \mu H_I + mH_I^o, \tag{2.1b}$$

$$\frac{dT_s}{dt} = \hat{b}(T_s + T_I)e^{-\hat{q}(T_s+T_I)} - \hat{\mu}T_s - \hat{\beta}\frac{T_s}{H_s + H_I}H_I \tag{2.1c}$$

$$+ (1 - r_s)N_a mH_s^o + (1 - r_I)N_a mH_I^o$$

$$\frac{dT_I}{dt} = \hat{\beta}\frac{T_s}{T_s + T_I}H_I - \hat{\mu}T_I + r_s N_a mH_s^o + r_I N_a mH_I^o, \tag{2.1d}$$

where the physical meanings of the parameters are given in Table 2.1, with typical values based on other works in the literature. The Ricker function [31] is used for the birth term of both tick and host. The death rate for infected and susceptible hosts are assumed to be a common constant μ, and that for all ticks are assumed to be $\hat{\mu}$. Hereafter, a parameter with a hat symbol is associated with ticks. We assume that the transmission of disease occurs only from ticks to host or from host to tick, but not from tick to tick (such as cofeeding [34, 35]) or from host to host. The migration of host to the concerned region (Canada) is proportional to the host population in the outside region (US) with a migration rate m, since the outside region is assumed to have a much larger population size for host. The ratio of infected ticks over all ticks on susceptible host is r_s and that on infected hosts is r_I. The value of r_I should be greater than r_s since the host is more likely to be in the infected state if there are more infected ticks feeding on it.

For convenience, we denote total populations of ticks and hosts as

$$T = T_s + T_I, \quad H = H_s + H_I, \quad H^o = H_s^o + H_I^o, \tag{2.2}$$

where superscript o means the outside region. By (2.1), the equivalent system for (H, T, H_I, T_I)

is given by

$$\frac{dH}{dt} = bHe^{-qH} - \mu H + mH^o, \tag{2.3a}$$

$$\frac{dT}{dt} = \hat{b}Te^{-\hat{q}T} - \hat{\mu}T + N_a m H^o, \tag{2.3b}$$

$$\frac{dH_I}{dt} = \beta\frac{H - H_I}{H}T_I - \mu H_I + mH_I^o, \tag{2.3c}$$

$$\frac{dT_I}{dt} = \hat{\beta}\frac{T - T_I}{T}H_I - \hat{\mu}T_I + (r_s H_s^o + r_I H_I^o)N_a m. \tag{2.3d}$$

Based on the physical meanings of the parameters (see Table 2.1), we naturally have the following restrictions

$$\begin{aligned} &b > 0, \quad \mu > 0, \quad \hat{b} > 0, \quad \hat{\mu} > 0, \quad \beta > 0, \quad \hat{\beta} > 0, \\ &q > 0, \quad \hat{q} > 0, \quad N_a > 0, \quad 0 < r_s < 1, \quad 0 < r_I < 1, \\ &m > 0, \quad H^o > H_I^o > 0, \quad H^o > H_s^o > 0. \end{aligned} \tag{2.4}$$

2.2 Positivity and equilibrium analysis

In this section, we first do the non-dimensionalization to identify the effective combination of parameters. Then, we show the positivity of solutions, which is a crucial property of a biological system. Finally, we analyze the equilibrium of the system and its stability.

2.2.1 Non-dimensionlization

Now we simplify the system in (2.3) by using the scales

$$\tilde{H} = qH, \quad \tilde{H}_I = qH_I, \quad \tilde{T} = \hat{q}T, \quad \tilde{T}_I = \hat{q}T_I, \quad \tilde{t} = bt. \tag{2.5}$$

Then we can get the dimensionless system for $\tilde{H}, \tilde{T}, \tilde{H}_I, \tilde{T}_I$. To make the system easier to read we drop the tilde, and the dimensionless system is

$$\begin{aligned}
\frac{dH}{dt} &= He^{-H} - s_1 H + s_2, && \text{(2.6a)}\\
\frac{dT}{dt} &= s_3 T e^{-T} - s_4 T + s_5, && \text{(2.6b)}\\
\frac{dH_I}{dt} &= s_6(1 - \frac{H_I}{H})T_I - s_1 H_I + s_7, && \text{(2.6c)}\\
\frac{dT_I}{dt} &= s_8(1 - \frac{T_I}{T})H_I - s_4 T_I + s_9, && \text{(2.6d)}
\end{aligned}$$

where the dimensionless parameters are given by

$$\begin{aligned}
&s_1 = \frac{\mu}{b}, \quad s_2 = \frac{qmH^o}{b}, \quad s_3 = \frac{\hat{b}}{b}, \quad s_4 = \frac{\hat{\mu}}{b}, \quad s_5 = \frac{\hat{q}N_a m H^o}{b},\\
&s_6 = \frac{\beta q}{b\hat{q}}, \quad s_7 = \frac{qmH_I^o}{b}, \quad s_8 = \frac{\hat{\beta}\hat{q}}{bq}, \quad s_9 = \frac{\hat{q}N_a m(r_s H_s^o + r_I H_I^o)}{b}.
\end{aligned} \tag{2.7}$$

Then by (2.6) and the definitions of H_s and T_s, the corresponding equation for H_s and T_s will be

$$\begin{aligned}
\frac{dH_s}{dt} &= He^{-H} - s_6(1 - \frac{H_I}{H})T_I - s_1 H_s + s_2 - s_7,\\
\frac{dT_s}{dt} &= s_3 T e^{-T} - s_8(1 - \frac{T_I}{T})H_I - s_4 T_s + s_5 - s_9.
\end{aligned} \tag{2.8}$$

From the conditions of parameters in (2.4), we have the natural restrictions for the dimensionless parameters

$$s_i > 0, \quad (i = 1, .., 9), \quad s_2 > s_7, \quad s_5 > s_9. \tag{2.9}$$

2.2.2 Positivity of Solutions

In this subsection, we show the positivity of the solutions with initial positive data. To this end, we define the two sets:

$$\begin{aligned} S &= \{(T, H, T_I, H_I) | T > 0, H > 0, T_I > 0, H_I > 0\}, \\ \hat{S} &= \{(T_s, H_s, T_I, H_I) | T_s > 0, H_s > 0, T_I > 0, H_I > 0\}. \end{aligned} \tag{2.10}$$

The aim is to show the two sets are invariant. The invariance of $\hat{S}$ implies the invariance of S, but the reverse is not true, e.g., positive H and H_I can not imply positive $H_s = H - H_I$.

Theorem 2.2.1. *With the natural restrictions of parameters in* (2.9) *for the systems in* (2.6) *and* (2.8)*, or equivalently with conditions* (2.4) *for the original systems in* (2.3) *and* (2.1)*, the two sets S and $\hat{S}$ are invariant.*

Proof. First, we show that the set S is invariant. As Eq. (2.6)a is decoupled from the system, the positivity of H can be analyzed separately. We denote the right-hand side as

$$f(H) := He^{-H} - s_1 H + s_2, \tag{2.11}$$

and get

$$\frac{dH}{dt} = f(H) \to s_2 > 0, \quad \text{as} \quad H \to 0. \tag{2.12}$$

Thus, the quantity H remains positive. In fact, even for the case $s_2 = 0$ (or $m = 0$), one can show the positivity of H, by using the equivalent equation

$$\frac{d(He^{s_1 t})}{dt} = He^{-H+s_1 t} > 0, \quad \text{for} \quad H > 0. \tag{2.13}$$

Similarly, T remains positive by Eq. (2.6)b. We further show that the H and T are bounded, before proving the positivity of H_I and T_I. Clearly by (2.12), H is bounded from below, say $H \geq H_{min} > 0$. Note that

$$f(H) \to -\infty, \quad \text{as} \quad H \to \infty, \tag{2.14}$$

then there exists a $H_{max} \geq H(0) > 0$ such that

$$\frac{dH}{dt} = f(H) < 0, \quad \text{for} \quad H \geq H_{max}. \tag{2.15}$$

Therefore, we obtain $H \in [H_{min}, H_{max}]$, where in some cases H_{min} or H_{max} could be $H(0)$. Similarly T is bounded. With the boundedness, we get

$$\begin{aligned} H_I/H \to 0, \quad \text{as} \quad H_I \to 0, \\ T_I/T \to 0, \quad \text{as} \quad T_I \to 0. \end{aligned} \tag{2.16}$$

Subsequently in the first quadrant of the plane of (H_I, T_I), we obtain

$$\begin{aligned} \frac{dH_I}{dt} \to s_6 T_I + s_7 > 0, \quad \text{as} \quad H_I \to 0, \\ \frac{dT_I}{dt} \to s_8 H_I + s_9 > 0, \quad \text{as} \quad T_I \to 0. \end{aligned} \tag{2.17}$$

So, the two quantities H_I and T_I will remain positive.

Next, we prove the positivity of H_s and T_s, or equivalently the invariance of $\hat{S}$. We note that H and T are positive and bounded, independent of dynamics of H_s and T_s in (2.8). Each of the terms He^{-H} and Te^{-T} has a minimum positive value. Then, we have

$$\begin{aligned} \frac{dH_s}{dt} &\geq He^{-H} + s_2 - s_7 > 0, \quad \text{as} \quad H_s \to 0, \\ \frac{dT_s}{dt} &\geq s_3 Te^{-T} + s_5 - s_9 > 0, \quad \text{as} \quad T_s \to 0, \end{aligned} \tag{2.18}$$

which implies positivity of H_s and T_s and hence the invariance of $\hat{S}$. □

We remark that in the current model, we have used the Ricker's function He^{-H} for the birth term to ensure its positivity, which is essential for the positivity of H_s. A logistic birth term $H(1-H)$ (after non-dimensionlization) often appears in other tick models [26, 27], but with current migration effect such a term may give a negative birth rate, since $H > 1$ is possible with certain parameters.

2.2.3 Equilibrium analysis

Theorem 2.2.2. *With the natural conditions* (2.9) *for the system in* (2.6)*, there exists a unique positive equilibrium* (H^*, T^*, H_I^*, T_I^*)*, which is globally asymptotically stable.*

Proof. (1) As the equations for H and T are decoupled from the system, we first show the equilibrium of H and T. We take Eq. ((2.9)a) for example. With the definition (2.11), we

see that

$$f(0) = s_2 > 0, \quad f(\infty) = -\infty, \tag{2.19}$$

thus there exists at least one positive root (equilibrium). Taking the derivative gives

$$f'(H) = (1 - H)e^{-H} - s_1, \quad f'(0) = 1 - s_1. \tag{2.20}$$

There are two cases: $s_1 \geq 1$ and $0 < s_1 < 1$. When $s_1 \geq 1$, we always have

$$f'(H) < e^{-H} - s_1 < 0, \quad \text{for} \quad H > 0. \tag{2.21}$$

This implies $f(H)$ is monotone decreasing, and by (2.19) there exists a unique positive root. For the case $0 < s_1 < 1$, the $f'(H)$ is positive for small H by (2.20). Note that

$$f'(1) = -s_1 < 0, \tag{2.22}$$

so there exists a critical value $0 < H_c < 1$ such that

$$\begin{aligned} f'(H_c) &= (1 - H_c)e^{-H_c} - s_1 = 0, \\ f'(H) &> 0, \qquad \text{for} \quad 0 < H < H_c. \end{aligned} \tag{2.23}$$

Taking the second derivative gives

$$f''(H) = (H - 2)e^{-H} < 0, \quad \text{for} \quad H_c < H < 1, \tag{2.24}$$

which implies $f'(H)$ is decreasing in the above interval and

$$f'(H) < 0, \quad \text{for} \quad H_c < H < 1. \tag{2.25}$$

It is clear from (2.20) that

$$f'(H) = (1 - H)e^{-H} - s_1 < 0, \quad \text{for} \quad H \geq 1. \tag{2.26}$$

Combining ((2.23),(2.25),(2.28)), we conclude that $f(H)$ is increasing in $[0, H_c)$ and is decreasing in (H_c, ∞). Therefore, $f(H)$ has a unique positive root in (H_c, ∞). We denote the root (equilibrium) by H^*, then based on the above analysis, in either case we always have

$$\begin{aligned} f(H) > 0 \quad \text{for} \quad 0 < H < H^*, \\ f(H) < 0 \quad \text{for} \quad H > H^*. \end{aligned} \tag{2.27}$$

Therefore, the equilibrium H^* is (globally) asymptotically stable.

Similar arguments apply to the function for T, by defining

$$g(T) = Te^{-T} - \frac{s_4}{s_3}T + \frac{s_5}{s_3}, \tag{2.28}$$

where s_1 and s_2 are replaced by the two fractions s_4/s_3 and s_5/s_3. Therefore the unique equilibrium, denoted by T^*, is (globally) asymptotically stable.

(2) Next we show the existence of the equilibrium point (H_I^*, T_I^*) and its local stability. The theory of asymptotically autonomous systems [36] implies that one can substitute the equilibria H^*, T^* into the system for T_I, H_I for the equilibrium analysis. By coupling the

two equations (2.6)c and 2.6d, the equilibrium of H_I is determined by

$$\begin{aligned} &AH_I^2 + BH_I + C = 0, \\ &A = -\frac{s_6 s_8}{H^*} - \frac{s_1 s_8}{T^*} < 0, \\ &B = B_1 + B_2 = \left(\frac{s_7 s_8}{T^*} - \frac{s_6 s_9}{H^*}\right) + (s_6 s_8 - s_1 s_4), \\ &C = s_4 s_7 + s_6 s_9 > 0, \end{aligned} \tag{2.29}$$

where B_1 and B_2 are defined in the two brackets. Then, there exist two roots, one positive and one negative, denoted by $H_I^{(1)} > 0$ and $H_I^{(2)} < 0$. Similarly, the equilibrium of T_I is given by

$$\begin{aligned} &\hat{A}T_I^2 + \hat{B}T_I + \hat{C} = 0, \\ &\hat{A} = -\frac{s_6 s_8}{T^*} - \frac{s_4 s_6}{H^*} < 0, \\ &\hat{B} = -B_1 + B_2, \\ &\hat{C} = s_7 s_8 + s_1 s_9 > 0. \end{aligned} \tag{2.30}$$

The two roots (equilibria) are denoted by $T_I^{(1)} > 0$ and $T_I^{(2)} < 0$. Furthermore, by equations (2.6)c and (2.6)d, the two variables H_I and T_I at equilibrium are connected by a linear equation

$$B_1 + AH_I - \hat{A}T_I = 0, \tag{2.31}$$

with the same parameters $B_1, A, \hat{A}$ as above. Then, the two roots $T_I^{(1)}, T_I^{(2)}$ and those $H_I^{(1)}, H_I^{(2)}$ are related by this linear relation, so they form two pairs. Since $A < 0$ and $\hat{A} < 0$, the larger solution of H_I corresponds to the larger solution of T_I and hence the pos-

itive $T_I^{(1)}$ (or negative $T_I^{(2)}$) corresponds to the positive $H_I^{(1)}$ (or negative $H_I^{(2)}$). Therefore, we conclude that there is a unique pair of positive equilibria, denoted by $T_I^* = T_I^{(1)}$ and $H_I^* = H_I^{(1)}$.

For the system of (2.6) the local stability of the equilibrium (H_I^*, T_I^*) is determined by the eigenvalues of the matrix

$$D_I = \begin{pmatrix} -s_1 - s_6\frac{T_I^*}{H^*}, & s_6(1 - \frac{H_I^*}{H^*}) \\ s_8(1 - \frac{T_I^*}{T^*}), & -s_4 - s_8\frac{H_I^*}{T^*} \end{pmatrix}. \tag{2.32}$$

By the conditions (2.9) and positivity of equilibrium, we easily get

$$\mathrm{Tr}(D_I) = -s_1 - s_6\frac{T_I^*}{H^*} - s_4 - s_8\frac{H_I^*}{T^*} < 0. \tag{2.33}$$

Then, the equilibrium is asymptotically stable if $\mathrm{Det}(D_I) > 0$, which can be expressed by

$$-B_2 - AH_I^* - \hat{A}T_I^* > 0. \tag{2.34}$$

By (2.31), it is equivalent to

$$B + 2AH_I^* < 0, \quad (\text{or} \quad \hat{B} + 2\hat{A}T_I^* < 0). \tag{2.35}$$

Since H_I^* is the positive root (the larger one) of quadratic equation in (2.29), we get

$$B + 2AH_I^* = -\sqrt{B^2 - 4AC} < 0. \tag{2.36}$$

This means the condition (2.34) automatically holds, and hence the equilibrium (H^*, T^*) is asymptotically stable. □

Remark 3.2.1: Theorems 2.2.1 and 2.2.2 imply that there exists a unique positive equilibrium $(H_s^*, H_I^*, T_s^*, T_I^*)$ for the original system (2.1), where $H_s^* = H^* - H_I^*, T_s^* = T^* - T_I^*$.

Remark 3.2.2: The above theorem deals with the case $m > 0$, now we compare it with the critical case $m = 0$. With $m = 0$, from the definitions (2.7), (2.29),(2.30) we have

$$s_2 = s_5 = s_7 = s_9 = 0, \quad C = \hat{C} = 0. \tag{2.37}$$

With $s_1 < 1$ (i.e., $\mu < b$) and $s_3 < s_4$ (i.e., $\hat{\mu} < \hat{b}$), there still exists a unique positive equilibrium (H^*, T^*) for the system of (H, T), which is asymptotically stable. For the system of (T_I, H_I), clearly there is a disease free equilibrium, $T_I = 0, H_I = 0$. The stability of this equilibrium (0,0) is determined by see (2.34))

$$-B_2 > 0, \quad \Leftrightarrow \quad s_1 s_4 - s_6 s_8 > 0, \quad \Leftrightarrow \quad R_0 \equiv \frac{\beta\hat{\beta}}{\mu\hat{\mu}} < 1. \tag{2.38}$$

R_0 is therefore a threshold parameter. If $R_0 < 1$, the equilibrium (0,0) is stable, and there is no positive equilibrium with $m = 0$ (the other equilibrium is negative). As m increases from 0, the positive equilibrium in the theorem is a perturbation of the disease free equilibrium. If $R_0 > 1$, the equilibrium (0,0) is unstable, but there is the other positive equilibrium with $m = 0$, denoted by (T_{I0}^*, H_{I0}^*), which is asymptotically stable. As m increases from 0, the positive equilibrium in the theorem is a perturbation of (T_{I0}^*, H_{I0}^*). If $R_0 = 1$, this (T_{I0}^*, H_{I0}^*) coincides with (0,0). In brief, the positive equilibrium with $m > 0$ is always a perturbation of the stable equilibrium with $m = 0$.

2.3 Numerical results and sensitivity analyses

In this section, we simulate the solutions with typical parameters. Then, we conduct the sensitivity analysis for some parameters. Finally, the effects of migration rate and death rates are further analyzed.

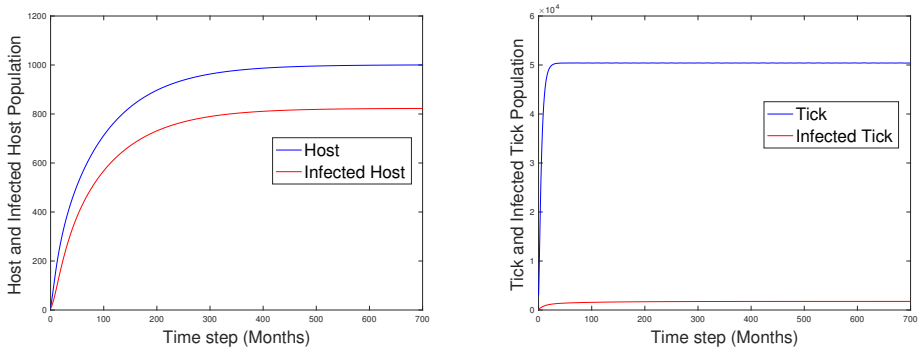
2.3.1 Solution curves

To simulate the dynamics of the system (2.3), we adopt the initial values as in [27]

$$H(0) = 10, \quad T(0) = 3000, \quad H_I(0) = 8, \quad T_I(0) = 80. \tag{2.39}$$

The initial values are based on one patch of 10,000 m^2 in [27], and other parameters used in simulation are shown in table 1. In Table 1, q is estimated from the carrying capacity of host, and we set $\hat{q} = q/M, N_a = M$, where $M = 200$ is maximum number of ticks per host in [26, 27]. Other estimated parameters are also consistent with this setting. If we were to consider a larger region, we could multiply a factor on these initial values and modify the estimated parameters $q, \hat{q}$ and H_I^o, H_s^o.

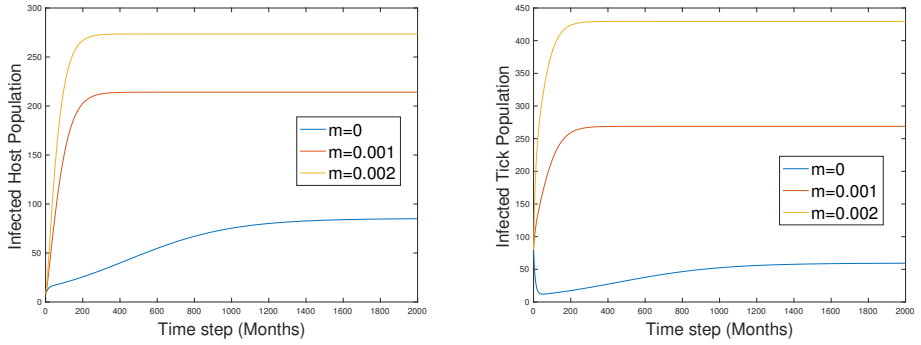
Figure 2.2 shows the dynamics for the populations of host, tick, infected host and infected tick, with parameters in Table 2.1. All the four populations quickly reach the positive equilibrium. To see more clearly the effect of migration rate m in the two different cases in Remark 3.2.2, the dynamics of the infected host and infected ticks are shown



(a) (b)

Figure 2.2: The dynamic curves for (a) host H and infected host H_I, (b) tick T and infected tick T_I.

in Figures 2.3 and 2.4 with varying parameters. With parameters in table 1, we have $R_0 = \frac{\beta\hat{\beta}}{\mu\hat{\mu}} = 1.4 > 1$, and Figure 2.3 shows the dynamics of H_I, T_I by varying m. Initially with $m = 0$, there is stable positive equilibrium, and as m increases the positive equilibrium increases. By changing μ from 0.01 to 0.02 (due to hunting or predators; alternatively changing $\hat{\mu}$ due to tick control measures), we get $R_0 = 0.7 < 1$, and Figure 2.4 shows the dynamics of H_I, T_I by varying m. Initially with $m = 0$, there is a stable disease free equilibrium (0,0), but as m increases this equilibrium increases as the positive equilibrium for the case $m > 0$.



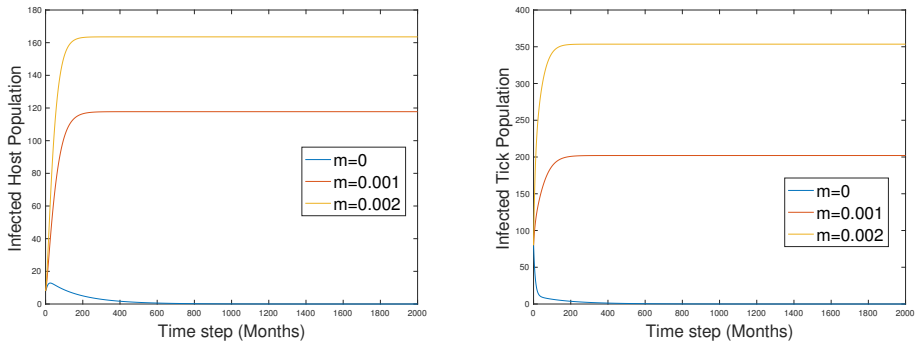
(a) (b)

Figure 2.3: The dynamic curves for infected host H_I and infected tick T_I with varying $m = 0, 0.001, 0.002$ ($R_0 = 1.4$ for the case $m = 0$).

2.3.2 Sensitivity Analyses

The mathematical analysis about the equilibrium and its stability can answer the questions about the long time behaviour, with fixed parameters. However, we have made assumptions about the parameters and there is always uncertainty in the estimation of parameters, which will result in uncertainty of the results. The Latin Hypercube Sampling (LHS) and Partial Rank Correlation Coefficient (PRCC) provides a useful tool to analyze a range of parameters and its effects on the dynamics/equilibrium of populations [37, 38, 39].

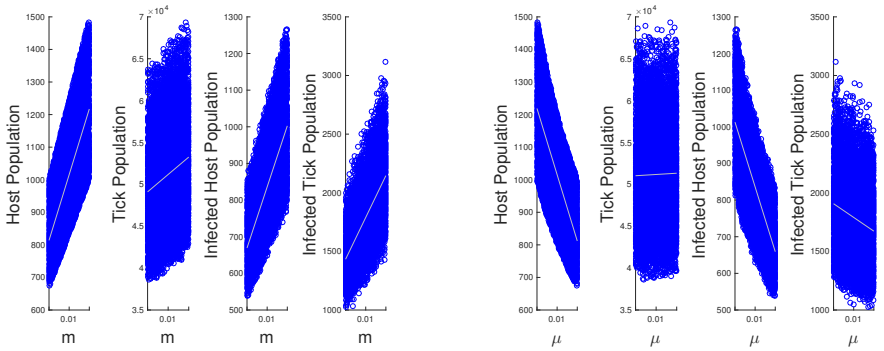
In the first step, we carry out the LHS, and verify the monotone relationship to ensure that the selected range of parameters are suitable for the PRCC analysis. We take m and



(a) (b)

Figure 2.4: The dynamic curves for infected host H_I and infected tick T_I with varying $m = 0, 0.001, 0.002$ ($R_0 = 0.7$ for the case $m = 0$)

μ for illustration, and other parameters such as r_s, r_I , $\hat{\mu}$ are similarly verified but omitted here for brevity. We set the range of the parameters to be a 20% change of the value in Table 2.1, i.e., $m \in [0.008, 0.012]$ and $\mu \in [0.008, 0.012]$. Uniform distribution is used in the LHS and 10,000 parameter sets are generated. Figure 2.5 shows the scatter plots for the equilibrium of four populations with respect to the two parameters m and μ in the parameter sets. Each point in the figure is from a simulation with one parameter set. One can see that all the subfigures show an increasing or decreasing pattern, except the tick population in Figure 2.5(b), since tick population T does not depend on μ by Eq. (2.3b). This suggests that the selected range of parameters can be used to perform the PRCC analysis.



(a) (b)

Figure 2.5: The scatter plot between the four equilibrium populations (H^*, T^*, H_I^*, T_I^*) and two parameters m and μ by LHS.

Next we calculate the PRCC, which determines how much each input parameter contributes to the output variable or measure. Figure 2.6 shows the PRCC results between the equilibrium of four populations and seven parameters $m, \mu, \hat{\mu}, q, \hat{q}, r_s, r_I$. The approximate p-values are shown inside the figure, and $p < 0.01$ means a significant result. For a significant result, the positive or negative PRCC value indicates a positive or negative correlation between the parameter and the output [38]. In Figure 2.6, the migration rate m is significant for all four equilibrium populations, and is positively correlated with them. The parameters μ (major effect) and q (minor effect) are negatively correlated with the equilibrium host population H^*, while $\hat{\mu}$ and $\hat{q}$ negatively affects the equilibrium tick population T^*. Both death rates μ and $\hat{\mu}$ are negatively correlated with the equilibrium infected host

and infected tick populations, as there is strong coupling in the two equations of H_I and T_I. The parameters r_I, r_s have major positive correlations with the equilibrium infected tick population T_I^*, as they directly appear in (2.3d), and they have minor positive effects on the equilibrium infected host population H_I^* due to the interactions of T_I and H_I.

2.3.3 The effects of migration rate and death rates

We further analyze effects of migration rate m and death rates $\mu, \hat{\mu}$ on the equilibrium populations for two reasons. The sensitivity analysis shows that these parameters are significant for the equilibrium populations.

Table 2.2: Effect of migration rate.

Population	slope at $m = 0.01$	change by increasing m by 10%
Host	9.92×10^4	9.94%
Tick	1.04×10^6	2.08%
Infected host	8.16×10^4	9.93%
Infected tick	1.74×10^5	9.88%

Figure 2.7 shows the dependence of the four equilibrium populations on migration rate m, indicating a positive correlation. This is consistent with the previous analysis and

PRCC results. Table 2.2 presents the slope of the curve at $m = 0.01$ and the relative change of population with increase of m by 10%. It shows about 10% increase for the three populations H^*, H_I^*, T_I^*, which is significant. The dependence on m can also be directly obtained based on previous analysis, for example, we get

$$\frac{dH^*}{dm} = \frac{H^o}{\mu + be^{-qH^*}(qH^* - 1)}, \quad \frac{dT^*}{dm} = \frac{H^o N_a}{\hat{\mu} + \hat{b}e^{-\hat{q}T^*}(\hat{q}T^* - 1)}. \tag{2.40}$$

With the chosen parameters and equilibrium values in Figure 2.2, we obtain exactly the results in Table 2.2. This verifies the results in this section. The formulas for the slopes dH_I^*/dm and dT_I^*/dm are complicated and hence omitted here.

Table 2.3: Effect of death rate of host.

Population	slope at $\mu = 0.01$	change by increasing μ by 10%
Host	-9.91×10^4	-8.99%
Tick	0	0%
Infected host	-8.62×10^4	-9.48%
Infected tick	-5.76×10^4	-2.97%

The dependence of the equilibrium populations on death rate of host μ is presented in Figure 2.8 and Table 2.3. The death rate μ has a negative correlation with the three populations H^*, H_I^* and T_I^*. It significantly affects H^* and H_I^*, which decrease by 8.96%

and 9.44% with an increase of μ by 10%. The T^*-μ subfigure (a constant) is omitted in Figure 2.8, since there is no correlation between T^* and μ, indicated by 0% in Table 2.3. This is because in the present model the birth term of ticks in (2.3b) does not depend on the host population H, although H could be included in the parameter $\hat{q}$ in future work.

Table 2.4: Effect of death rate of tick.

Population	slope at $\hat{\mu} = 0.1$	change by increasing $\hat{\mu}$ by 10%
Host	0	0%
Tick	-2.63×10^5	−4.91%
Infected host	-1.45×10^3	−1.73%
Infected tick	-1.84×10^4	−9.54%

Figure 2.9 and Table 2.4 show the effects of death rate of tick $\hat{\mu}$. The parameter $\hat{\mu}$ has a negative correlation with the three populations T^*, H_I^*, T_I^*, and has no relation with H^*. It has most significant impact on T_I^*, with a 9.54% decrease with increase of $\hat{\mu}$ by 10 %.

2.4 Climate Change and Implication for Public Health Interventions

The woodland habitats in Canada are suitable for survival of ticks off-host since the duff layer insulates ticks from deep freezing in winter and desiccation in the summer [40, 18,

41]. These studies also suggest the expansion of *I. scapularis* to may be accelerated northward in coming decades due to global warming. The death rate of ticks is affected by climate and temperature and can also be controlled by the use of acaricide [26, 42]. Climate plays an indirect role as it determines a suitable communities of hosts and vegetation to allow the creation of a duff layer that protects the tick from drowning and desiccation. Also, the ambient temperature and humidity directly affect the host-seeking activities. Lastly, the temperature affects the rate of development of ticks from one life stage to the next, in case of 0° C there is no development and in higher temperature the development is faster [43]

The migration and death rate of deer, also the death rate of tick are related to practical control measures [44], and can provide insights for future policy making. The migration rate of deer is affected by border control and deer habitat suitability [17] while the death rate is influenced by its predators and the hunting activities [3].

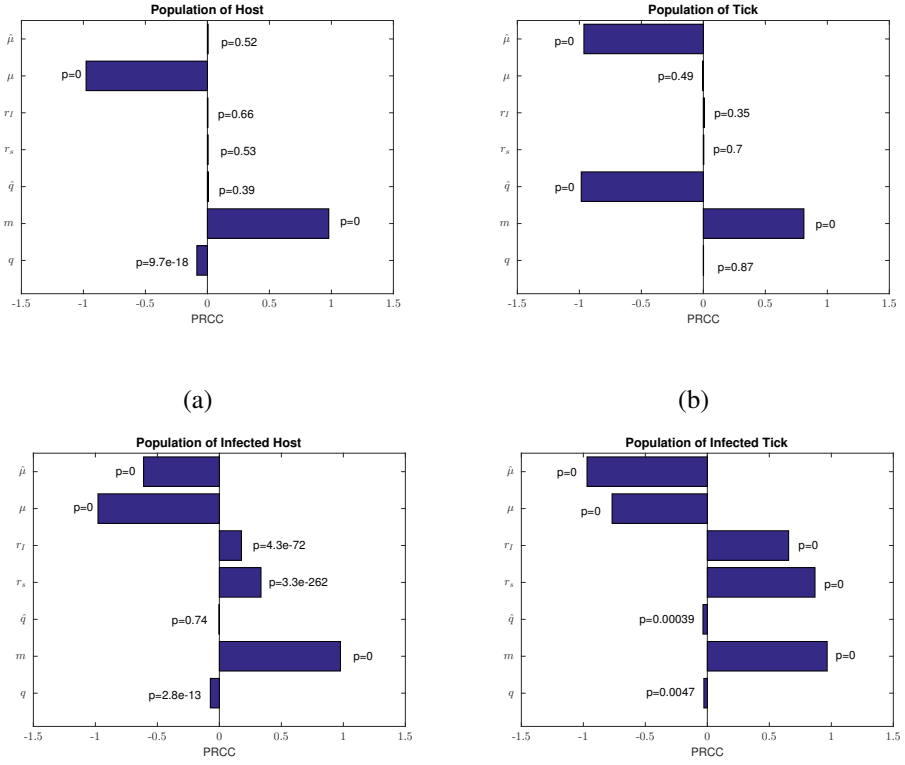
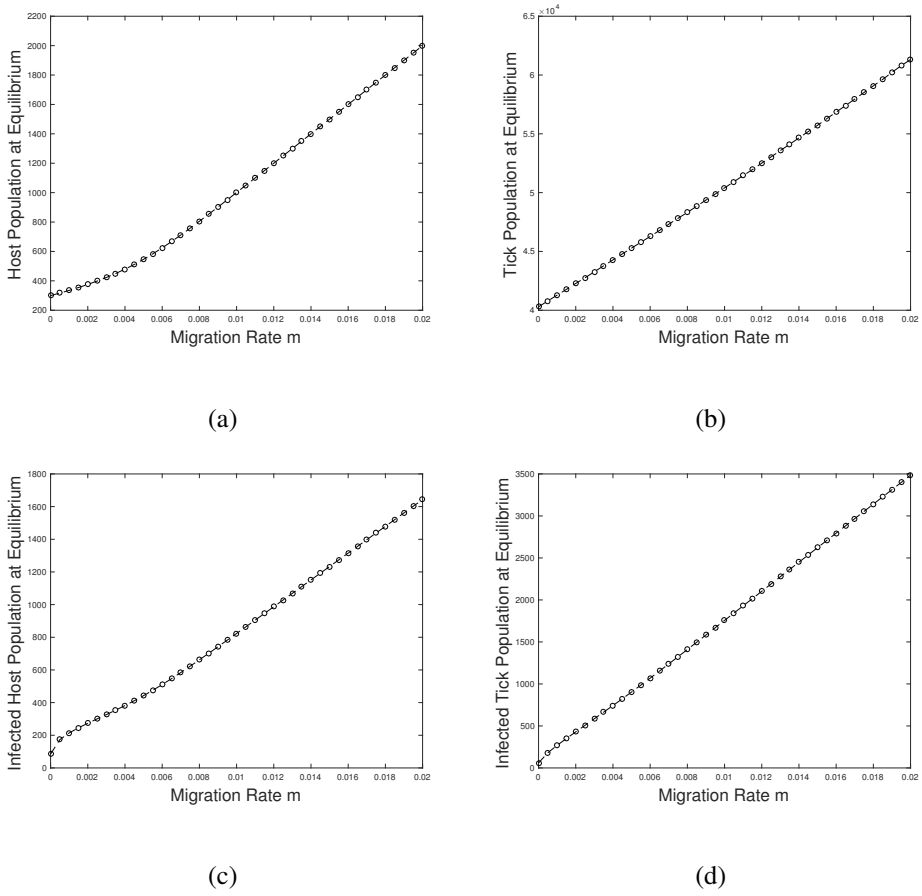


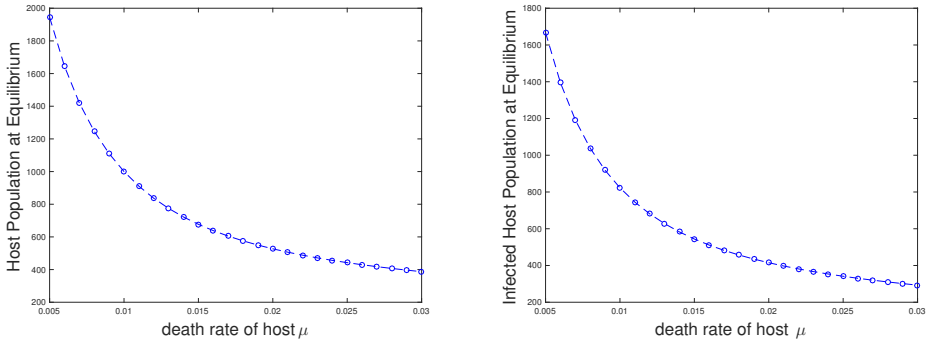
Figure 2.6: The PRCC results for four equilibrium populations (H^*, T^*, H_I^*, T_I^*) against parameters m, μ, $\hat{\mu}$, q, $\hat{q}$, r_s and r_I.



(a) (b)

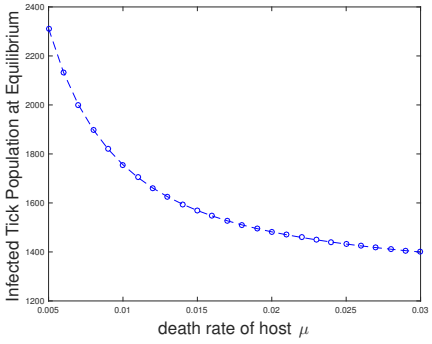
(c) (d)

Figure 2.7: The dependence of equilibrium populations on migration rate m: (a) host H^*, (b) tick T^*, (c) infected host H_I^*, (d) infected tick T_I^*.



(a)

(b)



(c)

Figure 2.8: The dependence of equilibrium populations on death rate of host μ: (a) host H^*, (b) infected host H_I^*, (c) infected tick T_I^*.

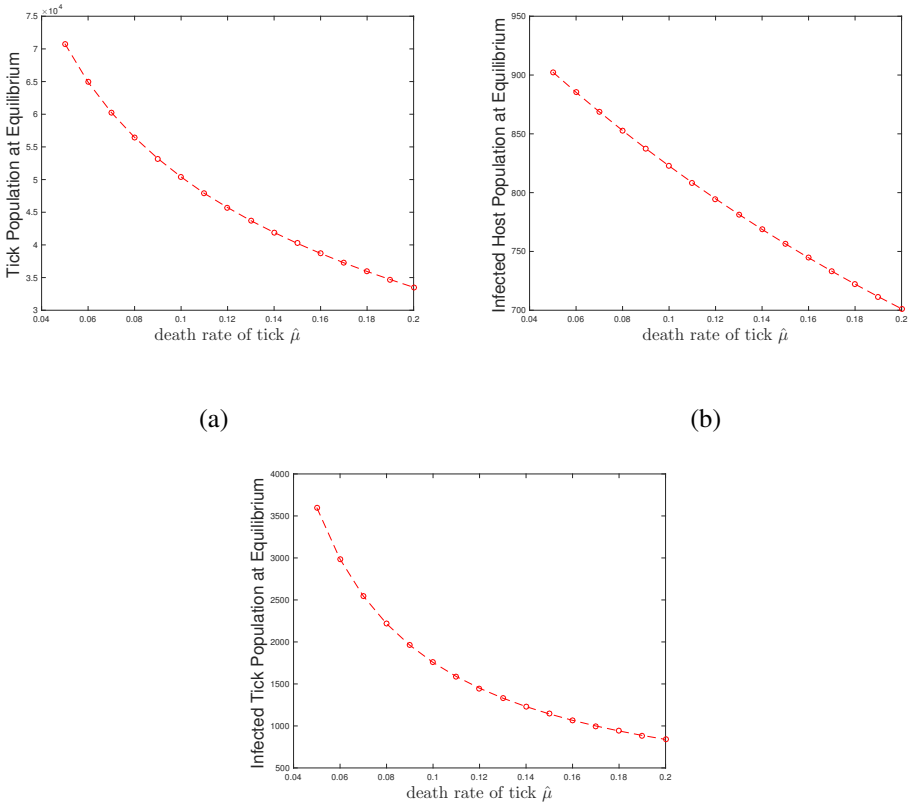


Figure 2.9: The dependence of equilibrium populations on death rate of tick $\hat{\mu}$: (a) tick T^*, (b) infected host H_I^*, (c) infected tick T_I^*.

3 Modeling the Impact of Host Resistance on Structured Tick Population Dynamics

Lyme Disease is the most reported athropod-borne illness and it was first recognized in 1976 in Lyme, Connecticut USA [45]. *Borrelia burgdorferi* is a tick-borne spirochete responsible for Lyme disease which is found in nymphal *Ixodes dammini* and has the highest chance to be transmitted to the host if the infected tick feeds for a duration of 72 hours or more [46, 47, 48]. Once an infected tick bites the host, a skin lesion called erythema migrans (EM) starts emerging and more than 95% of those patients diagnosed with Lyme disease have EM on the tick biting site [49, 45]. Once the bacterium enters the body it starts spreading in many organs and tissues through the lymph system and blood [49]. As time progresses the patient will experience headache, neck pain, fever, fatigue, and migratory musculoskeletal pain [47, 48, 49]. The government of Canada has data representing an increase of 144 cases in 2009 to 2025 cases in 2017 [50]. The *I.scopularis* also known as a black-legged tick is the main carrier of *B. burgdorferi* and has a life cycle

of nearly two years [45]. The tick population undergoes three main stages: L-larvae, N-nymph, A-adult and to move from one stage to the other ticks will quest feed and molt [51, 43, 52, 6]. Larvae and nymph feed on small rodents such as mice while adult ticks are more selective when it comes to their host since their body is larger compared to larvae and nymph and therefore the host must be a large mammal such as a deer. For ticks to move from one stage to the next it requires three hosts per stage and often the tick may use the same host for all three blood meals [51, 43, 52, 6]. Female ticks lay eggs in the spring and larvae hatch during late summer. The larvae that feeds during the late summer starts molting to nymph during winter. The nymph then starts feeding in the spring of the following year and molts into adult on the same year. Adult ticks die shortly right after they lay their eggs in the early spring [43, 52].

When a tick bites a host the expression of immunity varies depending on different hosts and tick species.The effects on ticks can vary from a simple rejection of the tick to interfering with the duration of feeding, inhibition of egg laying, also decreasing their viability to death of the tick while feeding. In addition, studies reveal that when female ticks feed on immune cattle their body of fully engorged tick was reduced by 30% [53, 54, 55]. According to Brown [56] hosts with resistance respond to tick bites with an intensified *grooming behaviour* and the attachment site is marked by serous exudes which could engulf the tick. In an experiment conducted on resistant guinea pigs bitten by Dermacentor andersoni, ba-

sophilia is present on the biting site. The attachment of a tick on a tick-sensitizes host is characterized by packs of basophils located in the intraepidermal vesicles. When ticks' extracts are injected into tick-sensitized host it causes a skin reaction and the plasma of the host expresses anti-tick antibodies which suggests a present mediated immune response. In case of unbitten animals, the reaction starts with neutrophils and the feeding site is characterized by an hemorrhagic as feeding progresses. Basophils start to also accumulate to the feeding site, however little degranulation occurs. In an experiment to study the effect of resistance of guinea pigs to ticks, basophil degranulation at tick feeding sites, resulted in tick rejection after tick-attachment: 29% after 6 hours, 18% after 12 hours, 22% after 24 hours, 37% after 48 hours and 7.3% after 72 and 96 hours. This shows that ticks are most susceptible to the resistance at 6, 12, 24 and 48 hours after attachment which are corresponding to the attachment time and fast feeding period [57].

There have been intensive studies modelling the dynamics of tick-host interaction and the transmission of various pathogens. Different aspects have also been included such as: seasonality, environmental changes, geographical heterogeneity and so on. On the other hand, few models incorporate delays in the development of tick from each life stage to the next [58, 59, 31]. Jennings et al. [60] studied the effect of host resistance on tick population dynamics. They developed a mathematical model, described by a system of ordinary differential equations, focusing on tick-host interaction where the tick's life cycle

was divided into two main stages, adult and juvenile, and the host was subdivided into host with no immunity and host with immunity. Their focus is to show how immunity affects the extinction or persistence of tick dynamics. However, their model does not include all biological stages (and sub-stages) of ticks and the possibility of different immunological response for each stage is ignored.

Here, we consider a stage-structured model that involves the full biological dynamics of tick and the emphasis is on the grooming behaviour of the host and its impact on tick population dynamics. We analyze the grooming behaviour in the mathematical model as a reduction in the successful attachment rates of ticks on the host i.e., the host-finding rates are reduced by a fraction for the host that shows resistance to tick bites. The model studies three main stages of tick's life cycle in which the ticks interact with hosts during questing-feeding-molting process. There is one more stage that we consider between Adult and Egg which is egg laying female. The host is divided into two compartments: host with resistance (host has been bitten by ticks before) and host with no resistance (host that has not been exposed to ticks). We observe that the basic reproduction number does not change with the resistance factor, however, numerical simulations show that the value of the positive equilibrium for different stages of tick population, and the dynamical behaviour of the solutions change with varying the resistance factor. Also, the sensitivity analysis demonstrates the dependence of the solutions on different parameters.

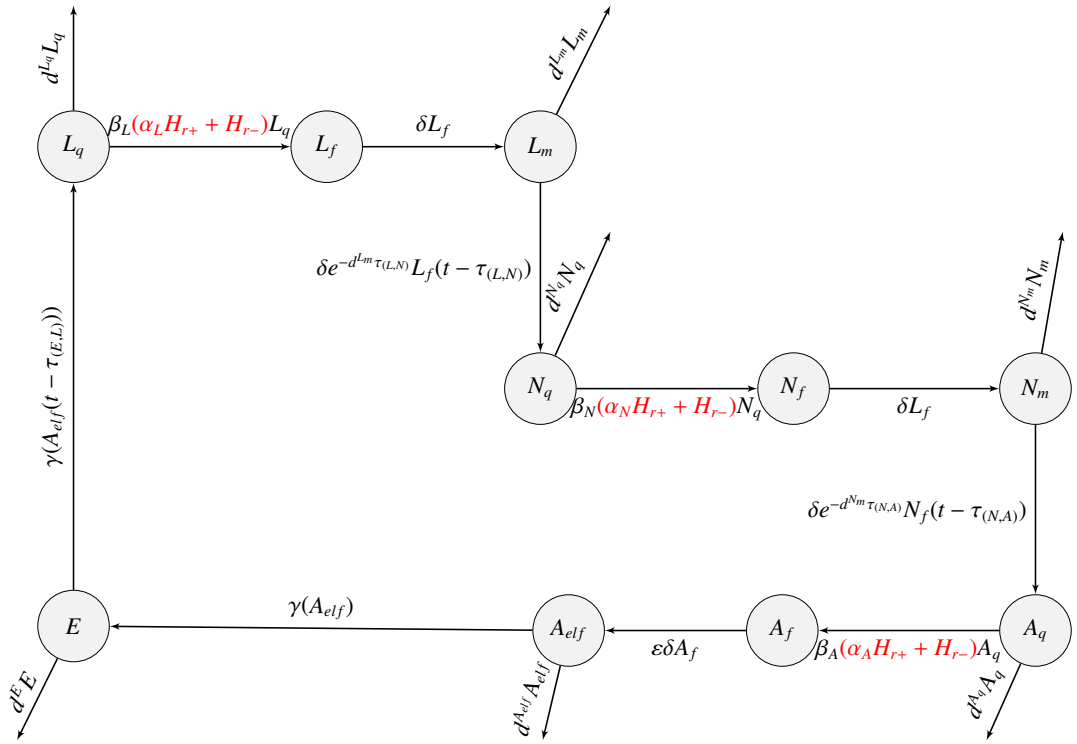


Figure 3.1: Flow diagram for the ticks' life cycle and their interaction with hosts. The variables are: larvae questing (L_q), larvae feeding (L_f), larvae molting (L_m), nymph questing (N_q), nymph feeding (N_f), nymph molting (N_m), egg (E), adult questing (A_q), adult feeding (A_f) and adult egg laying female (A_{elf}).

Table 3.1: Definition of parameters and their values

Symbol	Meaning	Value	Reference
d^{L_q}	Per capita mortality rate of L_q	0.6×10^{-2} per day	[61]
d^{L_m}	Per capita mortality rate of L_m	0.3×10^{-2} per day	[61]
d^{N_q}	Per capita mortality rate of N_q	0.6×10^{-2} per day	[61]
d^{N_m}	Per capita mortality rate of N_m	0.2×10^{-2} per day	[61]
d^{A_q}	Per capita mortality rate of A_q	0.6×10^{-2} per day	[61]
$d^{A_{elf}}$	Per capita mortality rate of A_{elf}	1 per day	[62]
d^{E}	Per capita mortality rate of E	0.2×10^{-2} per day	[61]
β_L	Successful attachment rate of questing larva to host	0.6×10^{-3} per day per host	[63]
β_N	Successful attachment rate of questing nymph to host	0.6×10^{-3} per day per host	[63]
β_A	Successful attachment rate of questing adult to host	0.2×10^{-2} per day per host	[63]
β_L^*	Rate of developing resistance to larva	$\kappa \times \beta_L$ per day per tick	Calculated
β_N^*	Rate of developing resistance to nymph	$\kappa \times \beta_N$ per day per tick	Calculated
β_A^*	Rate of developing resistance to adult	$\kappa \times \beta_A$ per day per tick	Calculated

Table 3.2: Definition of parameters and their values

Symbol	Meaning	Value	Reference
δ	Detachment rate	0.01 per day	[64]
α_L	Host grooming effect for larva	0.4, [0, 1] unitless	Assumed
α_N	Host grooming effect for nymph	0.6, [0, 1] unitless	Assumed
α_A	Host grooming effect for adult	0.5, [0, 1] unitless	Assumed
ϵ	Female proportion	0.5 unitless	[65]
$\tau_{(E,L)}$	The delay of development form egg to larvae	21 days	[61]
$\tau_{(L,N)}$	The delay of development form larvae to nymph	$101.18 \times Temp^{-2.25}$, 200 days	[61]
$\tau_{(N,A)}$	The delay of development form nymph to adult	$1596 \times Temp^{-1.21}$, 61 days	[61]
b	Birth rate of the host	0.66×10^{-3} per day	[59]
μ	Death rate of the host	0.33×10^{-3} per day	[59]
c	Crowding	3.5×10^{-4} per day	Calculated
K	Carrying Capacity of deers	20	[61]
p	Maximum number of eggs per female adult tick	3000	[61]
q	The strength of density dependence	0.001 unitless	[66]
κ	Constant factor for resistance development	0.0001 unitless	Assumed

Table 3.3: Modified parameter values to get different values for $\mathcal{R}_0$

Symbol	Modified Value	Comments
d^E	$1.2 \times 0.2 \times 10^{-2}$	p+20%p
d^{L_q}	$1.2 \times 0.6 \times 10^{-2}$	p+20%p
d^{L_m}	$1.2 \times 0.3 \times 10^{-2}$	p+20%p
d^{N_q}	$1.2 \times 0.6 \times 10^{-2}$	p+20%p
d^{N_m}	$1.2 \times 0.2 \times 10^{-2}$	p+20%p
d^{A_q}	$1.2 \times 0.6 \times 10^{-2}$	p+20%p
β_L	$0.1 \times 0.6 \times 10^{-3}, 0.2 \times 0.6 \times 10^{-3}$	$10\%p, 20\%p$
β_N	$0.3 \times 0.6 \times 10^{-3}, 0.5 \times 0.6 \times 10^{-3}$	$30\%p, 50\%p$
β_A	$0.5 \times 0.2 \times 10^{-2}$	50%p fixed
β_L^*	$\kappa \times \beta_L$	changed by changing β_L
β_N^*	$\kappa \times \beta_N$	changed by changing β_N
β_A^*	$\kappa \times \beta_A$	changed by changing β_A

Table 3.4: Modified parameter values to get different values for $\mathcal{R}_0$

Symbol	Modified Value	Comments
α_L	0.4	varied in [0, 1]
α_N	0.6	varied in [0, 1]
α_A	0.5	varied in [0, 1]
c	$1.2 \times 3.5 \times 10^{-4}$	$p + 20\%p$ fixed
q	0.001	not changed

3.1 Model Formulation

We start the model aiming to focus on the grooming behaviour. We model the three stages of larvae, nymph, and adult. The larvae and nymph populations are subdivided into questing (L_q, N_q), feeding (L_f, N_f) and molting (L_m, N_m). On the other hand, for the adult population we consider adult egg laying female A_{elf}, adult questing (A_q) and feeding (A_f). Since a single female tick lays several thousands eggs the birth rate is the entry into population which is represented by Ricker function, $\gamma(A) = pAe^{-qA}$ [66, 61]. Tick dynamics are regulated by insufficient resources for blood meal and this is illustrated in parameter q as a constant.

The delay functions, demonstrating the time delays of ticks molting from one stage to another, are constants. In the model, $\tau_{(E,L)}$, $\tau_{(L,N)}$, $\tau_{(N,A)}$ represent the time it takes for ticks to molt from egg to larvae, larvae to nymph and nymph to adult, respectively. The host population is divided into two compartments: H_{r+} represents the bitten host compartment that have developed with immunity; H_{r-} represents the compartment of hosts that have never been bitten before and therefore without immunity. H is the total host population with a birth rate b and a mortality rate μ. The density-dependent regulations of the host population is described by K, c, and $b-\mu$. The variables and parameters and their meaning are given in Tables 3.1 and 3.2. The life cycle of ticks and their interaction with hosts is illustrated in Figure 3.1. We suppose the successful attachment rates are reduced by a fraction α_L for larvea, α_N for nymph and α_A for adult ticks. Based on the results from [57] we can assume that α is in the range $[0.6, 0.95]$, however we will study the effect of α values in $[0, 1]$. We also assume the hosts with no resistance develop resistance to ticks at a rate, denoted by κ, that depends on the tick densities, tick attachment rates and the immune system response.

In order to make the model comprehensible we neglect few biological factors of tick dynamics. There are multiple blood meals that take place during molting procedures however in our model we consider only a homogeneous molting process, that is, ticks feed once, drop and molt with a constant time delay. Also we only consider same host for three

stages. The death rate depends on the stage of the tick (egg, larvae, nymph, adult) and also on whether the tick is questing or feeding. However, we consider a constant mortality rate. Impact of climate change on development of ticks having a non linear relationship with increasing ambient temperature has not also been modelled. In addition, the questing rate is considered constant, even though it decreases as the temperatures and the day light decreases. The model is described by the following system of delay differential equations:

$$
\begin{aligned}
\frac{dL_q}{dt} &= e^{-d^E \tau_{(E,L)}} \gamma(A_{elf}(t-\tau_{(E,L)})) - \beta_L L_q(t)[\alpha_L H_{r+}(t) + H_{r-}(t)] - d^{L_q} L_q(t) \\
\frac{dL_f}{dt} &= \beta_L L_q(t)(\alpha_L H_{r+}(t) + H_{r-}(t)) - \delta L_f(t) \\
\frac{dL_m}{dt} &= \delta L_f(t) - d^{L_m} L_m(t) - \delta e^{-d^{L_m}\tau_{(L,N)}} L_f(t-\tau_{(L,N)}) \\
\frac{dN_q}{dt} &= \delta e^{-d^{L_m}\tau_{(L,N)}} L_f(t-\tau_{(L,N)}) - \beta_N N_q(t)[\alpha_N H_{r+}(t) + H_{r-}(t)] - d^{N_q} N_q(t) \\
\frac{dN_f}{dt} &= \beta_N N_q(t)(\alpha_N H_{r+}(t) + H_{r-}(t)) - \delta N_f(t) \\
\frac{dN_m}{dt} &= \delta N_f(t) - d^{N_m} N_m(t) - \delta e^{-d^{N_m}\tau_{(N,A)}} N_f(t-\tau_{(N,A)}) \\
\frac{dA_q}{dt} &= \delta e^{-d^{N_m}\tau_{(N,A)}} N_f(t-\tau_{(N,A)}) - \beta_A A_q(t)[\alpha_A H_{r+}(t) + H_{r-}(t)] - d^{A_q} A_q(t) \\
\frac{dA_f}{dt} &= \beta_A A_q(t)(\alpha_A H_{r+}(t) + H_{r-}(t)) - \delta A_f(t) \\
\frac{dA_{elf}}{dt} &= \varepsilon \delta A_f(t) - d^{A_{elf}} A_{elf}(t) \\
\frac{dE}{dt} &= \gamma(A_{elf}(t)) - d^E E(t) - e^{-d^E \tau_{(E,L)}} \gamma(A_{elf}(t-\tau_{(E,L)})) \\
\frac{dH_{r-}}{dt} &= bH(t) - \mu H_{r-}(t) - \frac{c}{K} H(t) H_{r-}(t) - [\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t)] H_{r-}(t) \\
\frac{dH_{r+}}{dt} &= -\mu H_{r+}(t) - \frac{c}{K} H(t) H_{r+}(t) + [\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t)] H_{r-}(t)
\end{aligned}
\tag{3.1}
$$

where $\gamma(A) = pAe^{-qA}$ is the birth function. The variables are: larvae questing (L_q), larvae feeding (L_f), larvae molting (L_m), nymph questing (N_q), nymph feeding (N_f), nymph molting (N_m), egg (E), adult questing (A_q), adult feeding (A_f) and adult egg laying female (A_{elf}), unbitten host (H_{r-}) and bitten host (H_{r+}). Here, we use the following equation for the host population dynamics

$$\frac{dH(t)}{dt} = (b-\mu)H(t) - \frac{c}{K}(H(t))^2 \tag{3.2}$$

where $H(t) = H_{r-}(t) + H_{r+}(t)$. Note that the positive equilibrium of this equation is given by $\bar{H} = \frac{(b-\mu)}{c}K$. Interpreting K as an environmental constraint, and in order to have $\bar{H} \leq K$ we assume $c \geq (b-\mu)$, with $\bar{H} = K$ when the equality holds.

From System (3.1) we can get the following integral equations for $L_m(t)$, $N_m(t)$ and $E(t)$

$$\begin{aligned}
L_m(t) &= L_m(0) - \int_{-\tau_{(L,N)}}^{0} e^{-d^{L_m}(-s)}\delta L_f(s)ds + \int_{t-\tau_{(L,N)}}^{t} e^{-d^{L_m}(t-s)}\delta L_f(s)ds \\
N_m(t) &= N_m(0) - \int_{-\tau_{(N,A)}}^{0} e^{-d^{N_m}(-s)}\delta N_f(s)ds + \int_{t-\tau_{(N,A)}}^{t} e^{-d^{N_m}(t-s)}\delta N_f(s)ds \\
E(t) &= E(0) - \int_{-\tau_{(E,L)}}^{0} e^{-d^{E}(-s)}\gamma(A_{elf}(s))ds + \int_{t-\tau_{(E,L)}}^{t} e^{-d^{E}(t-s)}\gamma(A_{elf}(s))ds
\end{aligned} \tag{3.3}$$

Therefore System (3.1) is equivalent to the following

$$
\begin{aligned}
\frac{dL_q}{dt} &= e^{-d^E\tau_{(E,L)}}\gamma(A_{elf}(t-\tau_{(E,L)})) - \beta_L L_q(t)(\alpha_L H_{r+}(t) + H_{r-}(t)) - d^{L_q}L_q(t)\\
\frac{dL_f}{dt} &= \beta_L L_q(t)(\alpha_L H_{r+}(t) + H_{r-}(t)) - \delta L_f(t)\\
\frac{dN_q}{dt} &= \delta e^{-d^{L_m}\tau_{(L,N)}}L_f(t-\tau_{(L,N)}) - \beta_N N_q(t)(\alpha_N H_{r+}(t) + H_{r-}(t)) - d^{N_q}N_q(t)\\
\frac{dN_f}{dt} &= \beta_N N_q(t)(\alpha_N H_{r+}(t) + H_{r-}(t)) - \delta N_f(t)\\
\frac{dA_q}{dt} &= \delta e^{-d^{N_m}\tau_{(N,A)}}N_f(t-\tau_{(N,A)}) - \beta_A A_q(t)(\alpha_A H_{r+}(t) + H_{r-}(t)) - d^{A_q}A_q(t)\\
\frac{dA_f}{dt} &= \beta_A A_q(t)(\alpha_A H_{r+}(t) + H_{r-}(t)) - \delta A_f(t)\\
\frac{dA_{elf}}{dt} &= \varepsilon\delta A_f(t) - d^{A_{elf}}A_{elf}(t)\\
\frac{dH_{r-}}{dt} &= bH(t) - \mu H_{r-}(t) - \frac{c}{K}H(t)H_{r-}(t) - (\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t))H_{r-}(t)\\
\frac{dH_{r+}}{dt} &= -\mu H_{r+}(t) - \frac{c}{K}H(t)H_{r+}(t) + (\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t))H_{r-}(t)
\end{aligned}
\tag{3.4}
$$

together with (3.3).

For further analyses of this model we use the theory of monotone dynamical systems [67]. Let $\tau = (\tau_1, \cdots, \tau_{12})$ where $\tau_i \geq 0$, $\tau_2 = \tau_{(L,N)}$, $\tau_5 = \tau_{(N,A)}$), $\tau_9 = \tau_{(E,L)}$ are non zero and $\tau_i = 0$ for $i \neq 2, 5, 9$. Assume $|\tau| = \max\{\tau_i\}$.

Let C_τ be the product of Banach spaces $C_{\tau_i} = C([-\tau_i, 0], \mathbb{R})$, i.e.,

$$C_\tau = \prod_{i=1}^{12} C([-\tau_i, 0], \mathbb{R}).$$

Let $X_t = (X_t^1, \cdots, X_t^{12}) \in C_\tau$ be given by

$$X_t^i(\theta) = X^i(t+\theta), \quad i = 1, \cdots, 12.$$

where

$$X(t) = (X^1(t), \cdots, X^{12}(t)) = (L_q, L_f, L_m, N_q, N_f, N_m, A_q, A_f, A_{elf}, E, H_{r-}, H_{r+}).$$

Then the right hand side of the Equation (3.1) is given by

$$X'(t) = f(X_t). \tag{3.5}$$

We assume the initial data is non-negative. So we will assume the initial data X_0 is in the Banach space C_τ^+ defined below

$$C_\tau^+ = \{\phi \in C_\tau : \phi_i(\theta) \geq 0, -\tau_i \leq \theta \leq 0\}.$$

Also, for the initial data to be continuous and positive we assume:

$$\begin{aligned} L_m(0) &\geq \int_{-\tau_{(L,N)}}^{0} e^{-d^{Lm}(-s)} \delta L_f(s) ds \\ N_m(0) &\geq \int_{-\tau_{(N,A)}}^{0} e^{-d^{Nm}(-s)} \delta N_f(s) ds \\ E(0) &\geq \int_{-\tau_{(E,L)}}^{0} e^{-d^{E}(-s)} \gamma(A_{elf}(s)) ds. \end{aligned} \tag{3.6}$$

The fundamental theory of functional differential equations implies that the solutions exist and are unique for all $t \geq 0$. We now show that the solutions will be positive and remain bounded.

Therefore System (3.1) is equivalent to the following

$$
\begin{aligned}
\frac{dL_q}{dt} &= e^{-d^E\tau_{(E,L)}}\gamma(A_{elf}(t-\tau_{(E,L)})) - \beta_L L_q(t)(\alpha_L H_{r+}(t) + H_{r-}(t)) - d^{L_q}L_q(t)\\
\frac{dL_f}{dt} &= \beta_L L_q(t)(\alpha_L H_{r+}(t) + H_{r-}(t)) - \delta L_f(t)\\
\frac{dN_q}{dt} &= \delta e^{-d^{Lm}\tau_{(L,N)}}L_f(t-\tau_{(L,N)}) - \beta_N N_q(t)(\alpha_N H_{r+}(t) + H_{r-}(t)) - d^{N_q}N_q(t)\\
\frac{dN_f}{dt} &= \beta_N N_q(t)(\alpha_N H_{r+}(t) + H_{r-}(t)) - \delta N_f(t)\\
\frac{dA_q}{dt} &= \delta e^{-d^{Nm}\tau_{(N,A)}}N_f(t-\tau_{(N,A)}) - \beta_A A_q(t)(\alpha_A H_{r+}(t) + H_{r-}(t)) - d^{A_q}A_q(t)\\
\frac{dA_f}{dt} &= \beta_A A_q(t)(\alpha_A H_{r+}(t) + H_{r-}(t)) - \delta A_f(t)\\
\frac{dA_{elf}}{dt} &= \varepsilon\delta A_f(t) - d^{A_{elf}}A_{elf}(t)\\
\frac{dH_{r-}}{dt} &= bH(t) - \mu H_{r-}(t) - \frac{c}{K}H(t)H_{r-}(t) - (\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t))H_{r-}(t)\\
\frac{dH_{r+}}{dt} &= -\mu H_{r+}(t) - \frac{c}{K}H(t)H_{r+}(t) + (\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t))H_{r-}(t)
\end{aligned}
\tag{3.7}
$$

together with (3.3).

Since the host population stabilizes quickly at $\bar{H} = (b-\mu)K/c$, the limiting system is

as follows

$$
\begin{aligned}
\frac{dL_q}{dt} &= e^{-d^E \tau_{(E,L)}} \gamma(A_{elf}(t-\tau_{(E,L)})) + \beta_L(1-\alpha_L)L_q(t)H_{r+}(t) - (\beta_L \bar{H} + d^{L_q})L_q(t) \\
\frac{dL_f}{dt} &= -\beta_L(1-\alpha_L)L_q(t)H_{r+}(t) + \beta_L \bar{H} L_q(t) - \delta L_f(t) \\
\frac{dN_q}{dt} &= \delta e^{-d^{L_m} \tau_{(L,N)}} L_f(t-\tau_{(L,N)}) + \beta_N(1-\alpha_N)N_q(t)H_{r+}(t) - (\beta_N \bar{H} + d^{N_q})N_q(t) \\
\frac{dN_f}{dt} &= -\beta_N(1-\alpha_N)N_q(t)H_{r+}(t) + \beta_N \bar{H} N_q(t) - \delta N_f(t) \\
\frac{dA_q}{dt} &= \delta e^{-d^{N_m} \tau_{(N,A)}} N_f(t-\tau_{(N,A)}) + \beta_A(1-\alpha_A)A_q(t)H_{r+}(t) - (\beta_A \bar{H} + d^{A_q})A_q(t) \\
\frac{dA_f}{dt} &= -\beta_A A_q(t)(1-\alpha_A)H_{r+}(t) + \beta_A \bar{H} A_q(t) - \delta A_f(t) \\
\frac{dA_{elf}}{dt} &= \varepsilon \delta A_f(t) - d^{A_{elf}} A_{elf}(t) \\
\frac{dH_{r+}}{dt} &= -\mu H_{r+}(t) - \frac{c}{K}\bar{H} H_{r+}(t) + (\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t))(\bar{H} - H_{r+}(t))
\end{aligned}
\tag{3.8}
$$

3.2 Analyses

In this section we give the necessary condition for existence and uniqueness of the positive equilibrium point and the conditions for local stability of the tick free equilibrium.

3.2.1 Equilibria

Let X^* denote the vector $(L_q, L_f, N_q, N_f, A_q, A_f, A_{elf}, H_{r+})$ in $\mathbb{R}^8$, and let $f(X)$ be the right hand side of (3.8). In order to find all equilibria we need to solve the system $f(X) = 0$:

$$
\begin{aligned}
0 &= e^{-d^E \tau_{(E,L)}} \gamma(A_{elf}(t - \tau_{(E,L)})) + \beta_L(1 - \alpha_L)L_q(t)H_{r+}(t) - (\beta_L \bar{H} + d^{L_q})L_q(t) \\
0 &= -\beta_L(1 - \alpha_L)L_q(t)H_{r+}(t) + \beta_L \bar{H} L_q(t) - \delta L_f(t) \\
0 &= \delta e^{-d^{L_m} \tau_{(L,N)}} L_f(t - \tau_{(L,N)}) + \beta_N(1 - \alpha_N)N_q(t)H_{r+}(t) - (\beta_N \bar{H} + d^{N_q})N_q(t) \\
0 &= -\beta_N(1 - \alpha_N)N_q(t)H_{r+}(t) + \beta_N \bar{H} N_q(t) - \delta N_f(t) \\
0 &= \delta e^{-d^{N_m} \tau_{(N,A)}} N_f(t - \tau_{(N,A)}) + \beta_A(1 - \alpha_A)A_q(t)H_{r+}(t) - (\beta_A \bar{H} + d^{A_q})A_q(t) \\
0 &= -\beta_A A_q(t)(1 - \alpha_A)H_{r+}(t) + \beta_A \bar{H} A_q(t) - \delta A_f(t) \\
0 &= \varepsilon \delta A_f(t) - d^{A_{elf}} A_{elf}(t) \\
0 &= -\mu H_{r+}(t) - \frac{c}{K} \bar{H} H_{r+}(t) + (\beta_L^* L_q(t) + \beta_N^* N_q(t) + \beta_A^* A_q(t))(\bar{H} - H_{r+}(t))
\end{aligned}
\tag{3.9}
$$

At the tick-free equilibrium, where all tick stages are equal to zero, we have $H_{r+} = 0$. Let $H_{r+} \neq \bar{H}$ so that $(\bar{H} - (1 - \alpha_L)H_{r+}), (\bar{H} - (1 - \alpha_N)H_{r+}), (\bar{H} - (1 - \alpha_A)H_{r+}) > 0$. We want to derive conditions for existence and uniqueness of a (strongly) positive equilibrium point

($X_i > 0$ for all $i = 1, \cdots, n$). From the equations in (3.9) we get the following

$$
\begin{aligned}
L_q &= \frac{d^{A_{elf}}(\beta_N(\bar{H} - (1-\alpha_N)H_{r+}) + d^{N_q})(\beta_A(\bar{H} - (1-\alpha_A)H_{r+}) + d^{A_q})}{s_2 s_3 \epsilon \beta_L \beta_N \beta_L (\bar{H} - (1-\alpha_L)H_{r+})(\bar{H} - (1-\alpha_N)H_{r+})(\bar{H} - (1-\alpha_A)H_{r+})} A_{elf} \\
L_f &= \frac{d^{A_{elf}}(\beta_N(\bar{H} - (1-\alpha_N)H_{r+}) + d^{N_q})(\beta_A(\bar{H} - (1-\alpha_A)H_{r+}) + d^{A_q})}{s_2 s_3 \delta \epsilon \beta_N \beta_A (\bar{H} - (1-\alpha_N)H_{r+})(\bar{H} - (1-\alpha_A)H_{r+})} A_{elf} \\
N_q &= \frac{d^{A_{elf}}(\beta_A(\bar{H} - (1-\alpha_A)H_{r+}) + d^{A_q})}{s_3 \epsilon \beta_N \beta_A (\bar{H} - (1-\alpha_N)H_{r+})(\bar{H} - (1-\alpha_A)H_{r+})} A_{elf} \\
N_f &= \frac{d^{A_{elf}}(\beta_A(\bar{H} - (1-\alpha_A)H_{r+}) + d^{A_q})}{s_3 \delta \epsilon \beta_A (\bar{H} - (1-\alpha_A)H_{r+})} A_{elf} \\
A_q &= \frac{d^{A_{elf}}}{\epsilon \beta_A (\bar{H} - (1-\alpha_A)H_{r+})} A_{elf} \\
A_f &= \frac{d^{A_{elf}}}{\epsilon \delta} A_{elf}
\end{aligned}
\tag{3.10}
$$

where $s_1 = e^{-d^E \tau_{(E,L)}}$, $s_2 = e^{-d^{Lm} \tau_{(L,N)}}$ and $s_3 = e^{-d^{Nm} \tau_{(N,A)}}$. From the first equation in the system (3.9) we get

$$
L_q = \frac{s_1 \gamma(A_{elf})}{(\beta_L(\bar{H} - (1-\alpha_L)H_{r+}) + d^{L_q})} \tag{3.11}
$$

and therefore

$$\gamma(A_{elf}) = \frac{d^{A_{elf}}(\beta_L(\bar{H}-(1-\alpha_L)H_{r+})+d^{Lq})(\beta_N(\bar{H}-(1-\alpha_N)H_{r+})+d^{Nq})(\beta_A(\bar{H}-(1-\alpha_A)H_{r+})+d^{Aq})}{s_1 s_2 s_3 \epsilon\beta_L\beta_N\beta_A(\bar{H}-(1-\alpha_L)H_{r+})(\bar{H}-(1-\alpha_N)H_{r+})(\bar{H}-(1-\alpha_A)H_{r+})} A_{elf}. \tag{3.12}$$

Since $\gamma(A_{elf}) = pA_{elf}e^{-qA_{elf}}$ we have the following cases: $A_{elf} = 0$ or

$$pe^{-qA_{elf}} = \frac{d^{A_{elf}}(\beta_L(\bar{H}-(1-\alpha_L)H_{r+})+d^{Lq})(\beta_N(\bar{H}-(1-\alpha_N)H_{r+})+d^{Nq})(\beta_A(\bar{H}-(1-\alpha_A)H_{r+})+d^{Aq})}{s_1 s_2 s_3 \epsilon\beta_L\beta_N\beta_A(\bar{H}-(1-\alpha_L)H_{r+})(\bar{H}-(1-\alpha_N)H_{r+})(\bar{H}-(1-\alpha_A)H_{r+})} \tag{3.13}$$

Finally, we reduce the system (3.9) to the following system

$$0 = \Gamma(H_{r+}) - pe^{-qA_{elf}} \tag{3.14a}$$

$$0 = -bH_{r+} + \Omega(H_{r+})(\bar{H} - H_{r+})A_{elf} \tag{3.14b}$$

where

$$\Gamma(H_{r+}) = \frac{d^{A_{elf}}(\beta_L(\bar{H} - (1-\alpha_L)H_{r+}) + d^{Lq})(\beta_N(\bar{H} - (1-\alpha_N)H_{r+}) + d^{Nq})(\beta_A(\bar{H} - (1-\alpha_A)H_{r+}) + d^{Aq})}{s_1 s_2 s_3 \epsilon\beta_L\beta_N\beta_A(\bar{H} - (1-\alpha_L)H_{r+})(\bar{H} - (1-\alpha_N)H_{r+})(\bar{H} - (1-\alpha_A)H_{r+})}$$

$$\begin{aligned}\Omega(H_{r+}) = \beta_L^* &\frac{d^{A_{elf}}(\beta_N(\bar{H} - (1-\alpha_N)H_{r+}) + d^{Nq})(\beta_A(\bar{H} - (1-\alpha_A)H_{r+}) + d^{Aq})}{s_2 s_3 \epsilon\beta_L\beta_N\beta_A(\bar{H} - (1-\alpha_L)H_{r+})(\bar{H} - (1-\alpha_N)H_{r+})(\bar{H} - (1-\alpha_A)H_{r+})} \\ + \beta_N^* &\frac{d^{A_{elf}}(\beta_A(\bar{H} - (1-\alpha_L)H_{r+}) + d^{Aq})}{s_3 \epsilon\beta_N\beta_A(\bar{H} - (1-\alpha_N)H_{r+})(\bar{H} - (1-\alpha_A)H_{r+})} + \beta_A^* \frac{d^{A_{elf}}}{\epsilon\beta_A(\bar{H} - (1-\alpha_A)H_{r+})}\end{aligned}$$

From (3.14b) we have

$$A_{elf} = \frac{bH_{r+}}{\Omega(H_{r+})(\bar{H} - H_{r+})}$$

given that $H_{r+} \neq \bar{H}$ and $\Omega(H_{r+}) \neq 0$ (it can be proved that this holds). Substituting this in

the equation (3.14a) we get the following

$$\Gamma(H_{r+}) = pe^{-q\frac{bH_{r+}}{\Omega(H_{r+})(\bar{H}-H_{r+}).}} \tag{3.15}$$

This is a nonlinear equation for $H_{r+}$and we need to determine under what conditions this equation has a unique positive solution. Let $G(H_{r+})$ be the right hand side of Equation (3.15). The functions Γ and G have the following properties:

(i) Γ is a rational function and is strictly increasing for $0 < H_{r+} < \bar{H}$;

(ii) $\Gamma(0) > 0$ is given by

$$\frac{d^{A_{elf}}(\beta_L\bar{H} + d^{L_q})(\beta_A\bar{H} + d^{A_q})(\beta_N\bar{H} + d^{N_q})}{s_1 s_2 s_3 \epsilon\beta_L\beta_N\beta_A\bar{H}^3};$$

(iii) G is a negative exponential function and it approaches zero (exponentially) as H_{r+} approaches $\bar{H}$;

(iv) $G(0) = p$.

From these properties we can see that the equation (3.15) has at least one solution $0 < H_{r+} < \bar{H}$, if and only if $G(0) > \Gamma(0)$, i.e.,

$$p > \frac{d^{A_{elf}}(\beta_L\bar{H} + d^{L_q})(\beta_A\bar{H} + d^{A_q})(\beta_N\bar{H} + d^{N_q})}{s_1 s_2 s_3 \epsilon\beta_L\beta_N\beta_A\bar{H}^3}.$$

This solution is unique if $G(H_{r+})$ is monotonically decreasing, and this holds if and only if

$$\frac{d}{dH_{r+}}\left(\frac{H_{r+}}{\Omega(H_{r+})(\bar{H} - H_{r+})}\right) > 0$$

for all $H_{r+} \in (0, \bar{H})$. Let

$$\mathcal{R}_0^v = \{\frac{ps_1s_2s_3\epsilon\beta_L\beta_N\beta_A\bar{H}^3}{d^{A_{elf}}(\beta_L\bar{H} + d^{L_q})(\beta_A\bar{H} + d^{A_q})(\beta_N\bar{H} + d^{N_q})}\}^{1/7}.$$

If $\mathcal{R}_0^v > 1$, then system (3.8) has a positive equilibrium point. If additionally

$$\frac{d}{dH_{r+}}(\frac{H_{r+}}{\Omega(H_{r+})(\bar{H} - H_{r+})}) > 0$$

holds, then the positive equilibrium is unique.

3.2.2 Stability of the tick-free Equilibrium

First we linearize System (3.8) about a given equilibrium point using the Fréchet derivative of the functional operator $F(X)$, given by the right hand side of the System (3.8):

$$DF(X^*)X = \lim_{h\to 0}\Big(\frac{F(X^* + hX) - F(X^*)}{h}\Big)$$

The linearized system is given by

$$
\begin{aligned}
Df_1(X^*)X &= pe^{-d^E\tau_{(E,L)}}A_{elf}(t-\tau_{(E,L)})e^{-qA^*_{elf}(t-\tau_{(E,L)})} \\
&\quad - pqe^{-d^E\tau_{(E,L)}}A_{elf}(t-\tau_{(E,L)})A^*_{elf}(t-\tau_{(E,L)})e^{-qA^*_{elf}(t-\tau_{(E,L)})} \\
&\quad + (1-\alpha_L)\beta_L(L^*_q(t)H_{r+}(t)+L_q(t)H^*_{r+}(t)) - (\beta_L\bar{H}+d^{L_q})L_q(t) \\
Df_2(X^*)X &= -(1-\alpha_L)\beta_L(L^*_q(t)H_{r+}(t)+L_q(t)H^*_{r+}(t)) + \beta_L\bar{H}L_q(t) - \delta L_f(t) \\
Df_4(X^*)X &= \delta e^{-d^{Lm}\tau_{(L,N)}}L_f(t-\tau_{(L,N)}) \\
&\quad + (1-\alpha_N)\beta_N(N^*_q(t)H_{r+}(t)+N_q(t)H^*_{r+}(t)) - (\beta_N\bar{H}+d^{N_q})N_q(t) \\
Df_5(X^*)X &= -(1-\alpha_N)\beta_N(N^*_q(t)H_{r+}(t)+N_q(t)H^*_{r+}(t)) + \beta_N\bar{H}N_q(t) - \delta N_f(t) \\
Df_7(X^*)X &= \delta e^{-d^{Nm}\tau_{(N,A)}}N_f(t-\tau_{(N,A)}) \\
&\quad + (1-\alpha_A)\beta_A(A^*_q(t)H_{r+}(t)+A_q(t)H^*_{r+}(t)) - (\beta_A\bar{H}+d^{A_q})A_q(t) \\
Df_8(X^*)X &= -(1-\alpha_A)\beta_A(A^*_q(t)H_{r+}(t)+A_q(t)H^*_{r+}(t)) + \beta_A\bar{H}A_q(t) - \delta A_f(t) \\
Df_9(X^*)X &= \varepsilon\delta A_f(t) - d^{A_{elf}}A_{elf}(t) \\
Df_{12}(X^*)X &= -(\mu+\frac{c}{K}\bar{H})H_{r+}(t) + \bar{H}(\beta^*_L L^*_q(t)+\beta^*_N N^*_q(t)+\beta^*_A A^*_q(t)) \\
&\quad - \Big((\beta^*_L L_q(t)+\beta^*_N N_q(t)+\beta^*_A A_q(t))H^*_{r+}(t) + (\beta^*_L L^*_q(t)+\beta^*_N N^*_q(t)+\beta^*_A A^*_q(t))H_{r+}(t)\Big)
\end{aligned}
\tag{3.16}
$$

The linearized system about the tick-free equilibrium point is as follows:

$$
\begin{aligned}
Df_1(X^*)X &= ps_1A_{elf}(t-\tau_{(E,L)}) - (\beta_L\bar{H} + d^{L_q})L_q(t)\\
Df_2(X^*)X &= \beta_L\bar{H}L_q(t) - \delta L_f(t)\\
Df_4(X^*)X &= \delta s_2 L_f(t-\tau_{(L,N)}) - (\beta_N\bar{H} + d^{N_q})N_q(t)\\
Df_5(X^*)X &= \beta_N\bar{H}N_q(t) - \delta N_f(t)\\
Df_7(X^*)X &= \delta s_3 N_f(t-\tau_{(N,A)}) - (\beta_A\bar{H} + d^{A_q})A_q(t)\\
Df_8(X^*)X &= \beta_A\bar{H}A_q(t) - \delta A_f(t)\\
Df_9(X^*)X &= \varepsilon\delta A_f(t) - d^{A_{elf}}A_{elf}(t)\\
Df_{12}(X^*)X &= -(\mu + \frac{c}{K}\bar{H})H_{r+}(t)
\end{aligned}
\tag{3.17}
$$

Using the theory of monotone dynamical systems we can see that system (3.8) is cooperative [67] and therefore stability of the zero equilibrium of system (3.17) is given by the stability of the corresponding ODE system. If $\mathcal{R}_0^v < 1$, then $X = 0$ is the only equilibrium point of the system (3.8) and is locally asymptotically stable. When $\mathcal{R}_0^v > 1$, there exists a positive equilibrium point and $X = 0$ is unstable.

Proof. We use the method of next generation matrix for the ODE system given by $X'(t) =$

$JX(t)$ where the matrix J is obtained from system (3.17):

$$J = \begin{bmatrix} -\beta_L\bar{H} - d^{L_q} & 0 & 0 & 0 & 0 & 0 & ps_1 & 0 \\ \beta_L\bar{H} & -\delta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \delta s_2 & -\beta_N\bar{H} - d^{N_q} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_N\bar{H} & -\delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta s_3 & -\beta_A\bar{H} - d^{A_q} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta_A\bar{H} & -\delta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \epsilon\delta & -d^{A_{elf}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b \end{bmatrix}$$

The matrix J can be written as $J = F - V$. The zero equilibrium is locally asymptotically stable if $\rho(FV^{-1}) < 1$ (ρ is the spectral radius of FV^{-1}) and it is unstable if $\rho(FV^{-1}) > 1$. We can see that

$$\rho(FV^{-1}) = \left(\frac{ps_1s_2s_3\epsilon\beta_L\beta_N\beta_A\bar{H}^3}{d^{A_{elf}}(\beta_L\bar{H} + d^{L_q})(\beta_A\bar{H} + d^{A_q})(\beta_N\bar{H} + d^{N_q})}\right)^{1/7}.$$

Finally, we note that $\rho(FV^{-1}) < 1$ is equivalent to $\mathcal{R}_0^v < 1$. This completes the proof.

□

3.3 Numerical Simulations

In this section we study the long-term dynamical behaviour of the system using numerical simulations and perform a sensitivity analysis for different parameters.

3.3.1 Model parametrization and validation

The observation of the dynamical behaviour of each stage of the tick population is demonstrated by applying DDE23 packages in Matlab to System (3.8). The model is parameterized using parameter values available in mathematical and ecological literature ([65, 63, 61, 64, 66, 62]). Parameter values and initial conditions are given in Tables 3.1-3.2. We note that the grooming behaviour does not impact the initial growth of the tick population, since parameters reflecting the grooming factor do not change the value of the basic reproduction number. We consider three cases to illustrate the dynamics of tick population in the presence of grooming factor. However, in these cases we fix the values for parameters related to the grooming behaviour. In the first case (Figure 3.2) the basic reproduction number is below the threshold value i.e., $\mathcal{R}_0^v < 1$, the tick-free equilibrium is locally asymptotically stable and therefore all stages of ticks go extinct.

In case 2 (Figure 3.3) the basic reproduction number is slightly greater than one, the

tick-free equilibrium point becomes unstable and the solutions approach the positive equilibrium without any initial oscillatory behaviour. In case 3 (Figure 3.4) the solutions oscillate initially and then approach the positive equilibrium. When the resistance related parameter values are fixed and the rest of the parameters vary, the positive equilibrium becomes unstable and a limit cycle appears. Therefore, the solutions oscillate about the equilibrium point. The limit cycle appears as the value of α_A increases from 0 to 1.

To study the population behaviour without grooming factor we set $\alpha_L = \alpha_N = \alpha_A = 1$ and $\kappa = 0$ and for intense grooming behaviour the $\alpha_L = \alpha_N = \alpha_A = 0$. In addition, we observe the dynamics for a mild grooming behaviour where $\alpha_L = 0.4, \alpha_N = 0.6, \alpha_A = 0.5$ and $\kappa = 0.1 \times 10^{-5}$. The equilibrium value for all stages are higher when there is no grooming behaviour. In particular, the value of the adult egg laying females at the equilibrium is 693 for a mild resistance behaviour and 1.9×10^3, when there is no resistance (Figure 3.3 and the left side of Figure 3.5). We also see that by decreasing the resistance solutions with non-oscillatory behaviour show damped oscillation. In a maximum intensified grooming behaviour the tick attachment rates to hosts with resistance are reduced to 0, therefore high resistance of hosts affects the tick equilibrium values significantly. For instance, in Figure 3.5 the equilibrium value for A_{elf} reduces from 1.9×10^3, when there is no resistance, to 78 when the resistance is very high. Comparing the right side of Figure 3.4 with 3.6, demonstrates the effect of resistance factors on the dynamical behaviour of the solutions.

Reducing the resistance from high to a mild resistance results in an increase in the value of the equilibrium of A_{elf} from 78 to 1600. However, in the absence of host resistance, the tick population at different stages oscillate about a positive equilibrium ($A_{elf} \approx 2.7 \times 10^3$). In other words, by decreasing the grooming behaviour (increasing the value of α_L, α_N and α_A from 0 to 1), there is more available resources for ticks to feed on. Therefore, the dynamical behaviour of tick population at different stages changes from solutions converging to the positive equilibrium to oscillatory solutions. The dynamics of the feeding ticks are similar to those of questing ticks and therefore we exclude the figures here. When we ignore the resistance behaviour in case 2 and 3, the host population with resistance H_{r+} is equal to 0 and it reaches a positive equilibrium point when $\alpha_L = \alpha_N = \alpha_A = 0$.

3.3.2 LHS and PRCC

We perform Latin Hypercube Sampling to further analyze the effects of each parameter on the dynamics of each life stage of the ticks [68, 37] Before we proceed to performing PRCC a verification of monotonicity is necessary to ensure the correct range of the parameters for PRCC analysis. Next, we calculate the PRCC, which determines the contribution of each parameter to the output variable such the population of larvae questing. A PRCC value significantly greater than 0 indicates a positive correlation and for PRCC significantly less than 0, a negative correlation between the parameter and the output [69]. In

Figure 3.7, the PRCC for the larvae questing population demonstrates the negative correlation with the death rates d^{Aelf}, d^{Nm}, d^{Lm}, d^{Nq}, d^{Aq}, d^{Lq}, d^{E} and d^{Lq} having the highest effect on this stage. The detachment rate δ does not have an impact, however the parameters related ticks' biological characteristics, p, q, ϵ, have significant effects. We also observe that the host finding rates β_A, β_N,β_L, have positive correlations with larvae questing dynamics. For the values of most parameters that are taken from the literature, we would expect to see a reasonable correlation between the parameter and the output (in a range where the output is monotonically increasing or decreasing with parameter). For instance the output value of L_q (and therefore L_f) at the equilibrium is supposed to decrease with an increase of the larvae questing death rate (negative correlation).

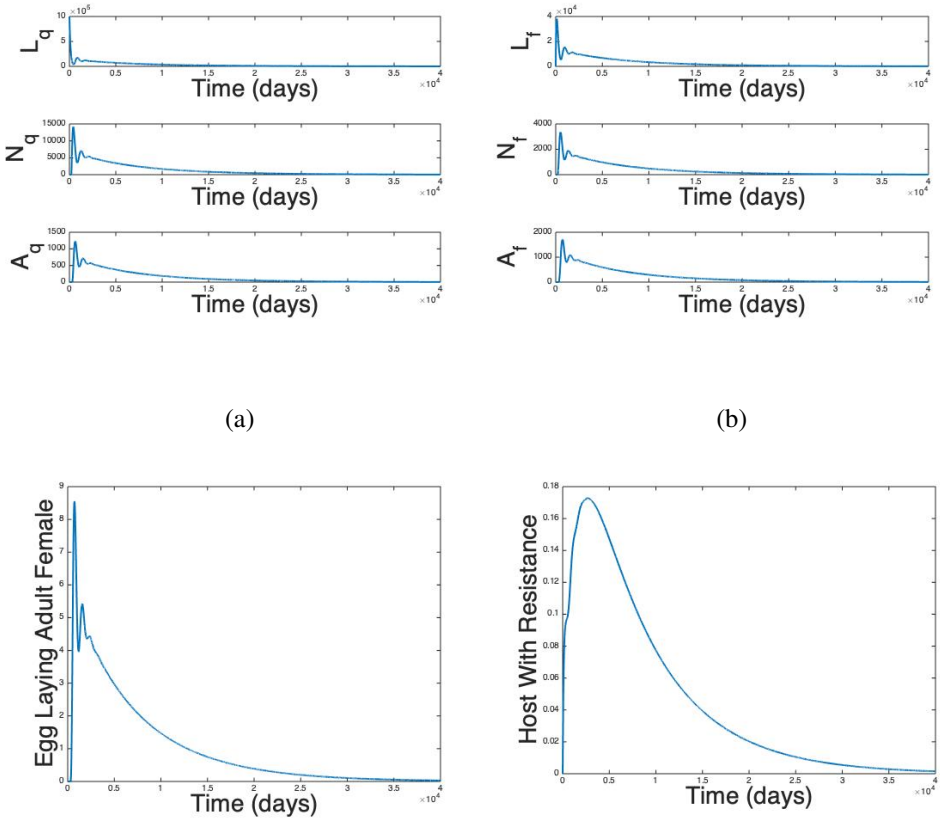


Figure 3.2: Case 1, $\mathcal{R}_0^v < 1$ where $\beta_L = 0.6 \times 10^{-4}$, $\beta_N = 1.8 \times 10^{-4}$ and $p = 200$ yields $\mathcal{R}_0^v = 0.89$.

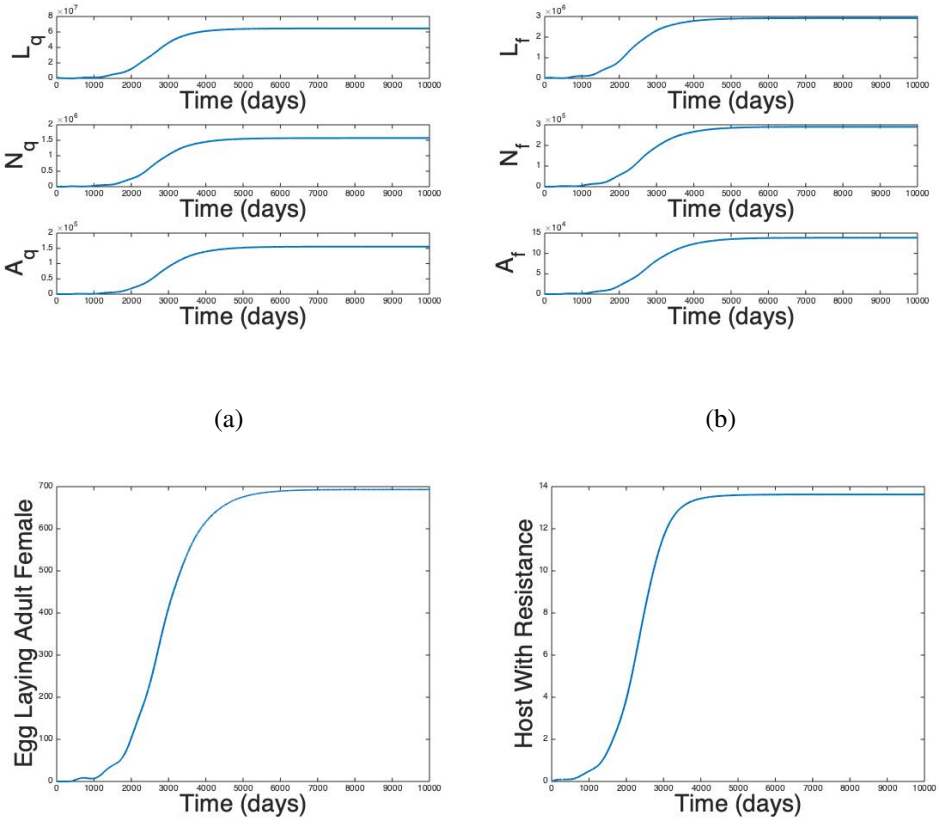


Figure 3.3: In Case 2 the values of p and κ have changed to $p = 1500$ and $\kappa = 0.1 \times 10^{-5}$ and the reproduction number increased to $\mathcal{R}_0^v = 6.71$. The simulations run for a time span of 10000 days. The equilibrium points for each stage of questing, feeding and adult egg laying female tick are as follows: $L_q = 6.5 \times 10^7, N_q = 1.6 \times 10^6, A_q = 1.6 \times 10^5$ $L_f = 2.9 \times 10^6, N_f = 2.9 \times 10^5, A_f = 1.4 \times 10^5, A_{elf} = 693$. In addition, the equilibrium point of the host with resistance is 13.

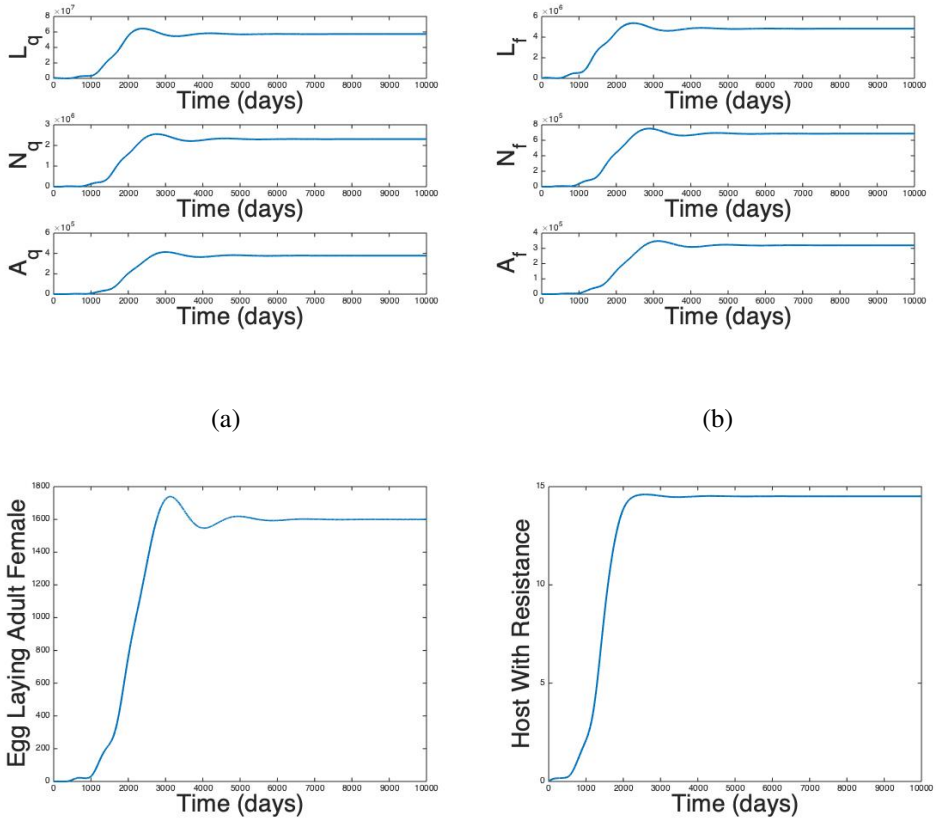


Figure 3.4: In Case 3 the values of β_L and β_N have changed to $\beta_L = 1.2\times10^{-4}, \beta_N = 3\times10^{-4}$ producing a higher reproduction number, $\mathcal{R}_0^v = 16.9$. The simulation are again running for a time span 10000 days. The equilibrium points for each stage of questing, feeding and adult egg laying female tick are as follows: $L_q = 5.7 \times 10^7, N_q = 2.3 \times 10^6, A_q = 3.8 \times 10^5, L_f = 4.8 \times 10^6, N_f = 6.9 \times 10^5, A_f = 3.2 \times 10^5, A_{elf} = 1600$. In addition, the equilibrium point of the host with resistance is 14.

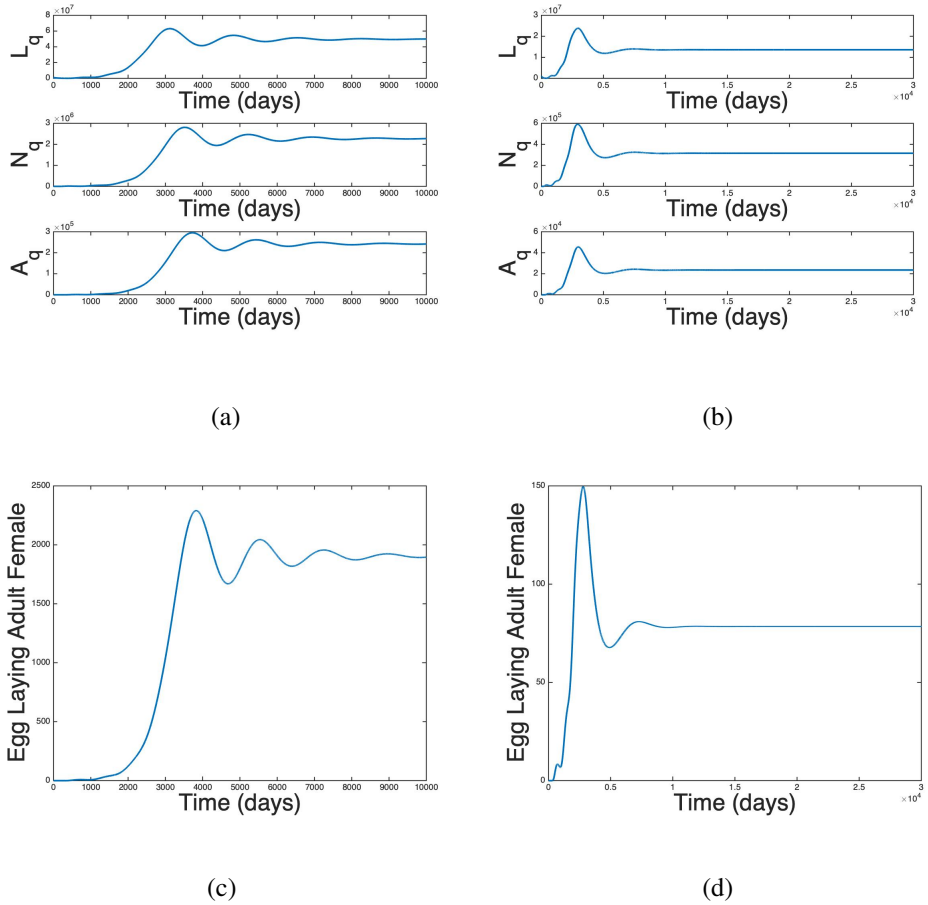


Figure 3.5: The parameter values are the same as in Case 2 except the $\alpha_L = \alpha_N = \alpha_A = 1$ (on the left). The equilibrium points are as follows: $L_q = 5.0 \times 10^7$, $N_q = 2.3 \times 10^6$, $A_q = 2.4 \times 10^5$, $A_{elf} = 1.9 \times 10^3$. There is no resistance and hence $H_r+ = 0$. In case of $\alpha_L = \alpha_N = \alpha_A = 0$ (on the right) the equilibrium points are $L_q = 1.3 \times 10^7$, $N_q = 3.2 \times 10^5$, $A_q = 2.4 \times 10^4$, $A_{elf} = 78$. Since now we introduce resistance, $H_{r+} = 10$.

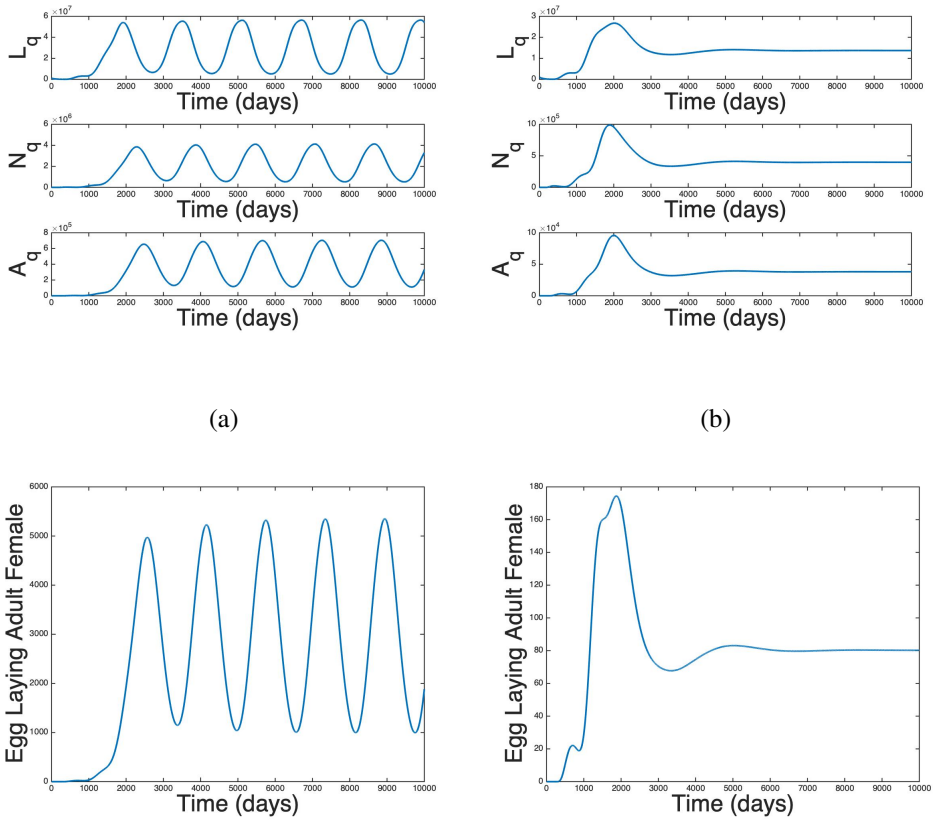


Figure 3.6: The parameter values are the same as in Case 3 except $\alpha_L = \alpha_N = \alpha_A = 1$ (the left). The equilibrium points are as follows: $L_q \approx 2.8\times10^7$, $N_q \approx 2.1\times10^6$, $A_q \approx 3.5\times10^5$, $A_{elf} \approx 2.7\times10^3$. Since resistance factor is not introduced the $H_{r+} = 0$. On the right side the $\alpha_L = \alpha_N = \alpha_A = 0$ and the equilibrium points are as follows: $L_q = 1.4\times10^7$, $N_q = 4.0\times10^5$, $A_q = 3.8\times10^4$, $A_{elf} = 80$. The resistance factor increase the population size from zero to $H_{r+} = 11$

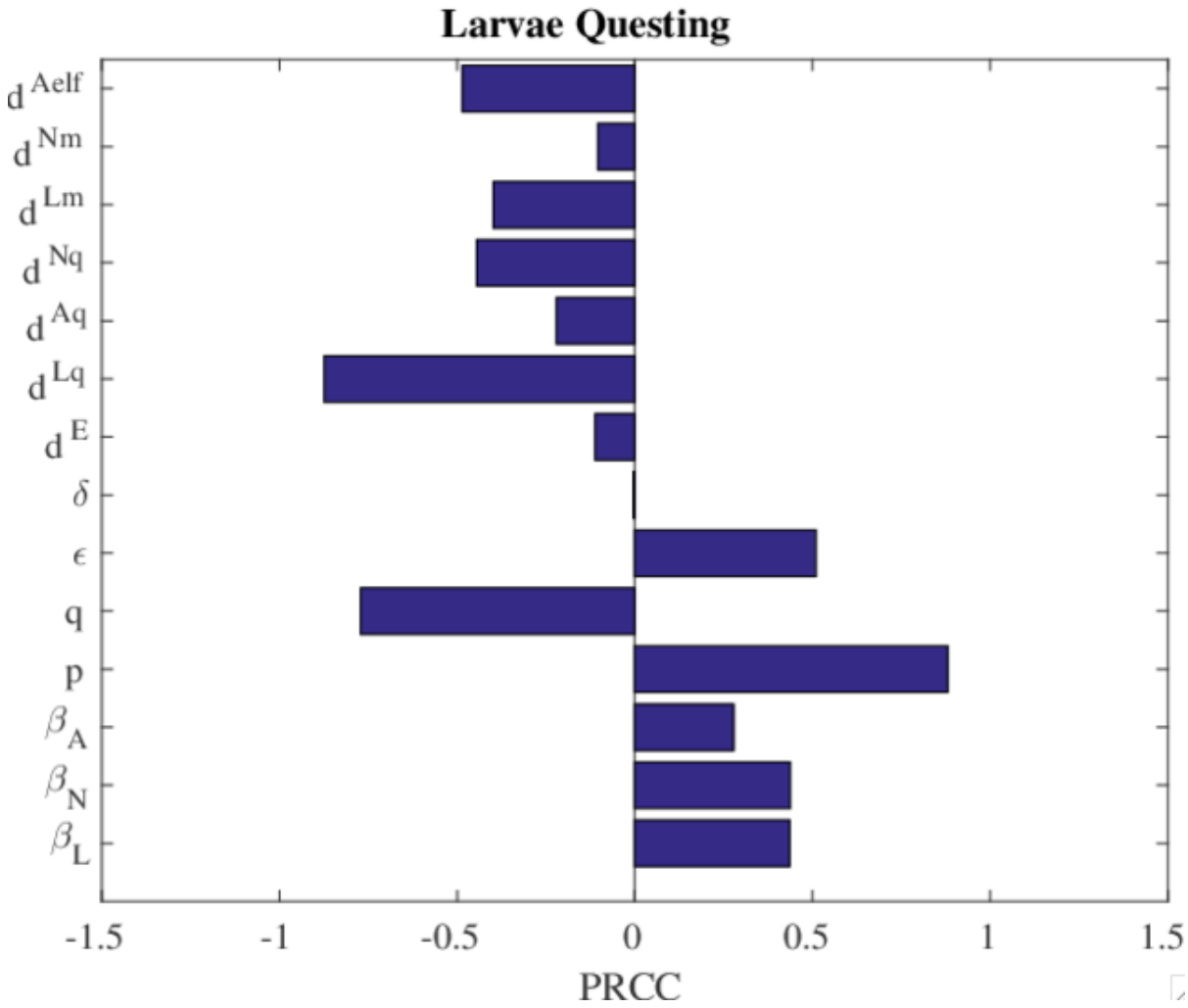


Figure 3.7: PRCC for most of the parameters used in the model at the equilibrium point of L_q. The value of each parameter is taken from 3.1, 3.2 and Case 2 for a range of $(+/-)20\%$

4 Time Series Data Clustering on Bitcoin Stock Market

4.0.1 Motivation: Impact of Temporal Variation of Environmental Conditions on Vector-Borne Disease Spread

Vector-borne diseases spread in nature is highly influenced by the spatiotemporal variation of the environmental conditions where vector-host interaction takes place. This variation can be reflected in the dependence of parameters on spatial and temporal variables. In particular, in a fixed study region, the temporal variation is better captured by the time-dependence of model coefficients including the survival probability, level of questing activities, and developmental delay. This can lead to periodic systems if seasonal temperature variation is the driving force of temporal variation, and this can also lead to much more complicated nonautonomous systems if climate warming and other climate changes can induce irregular change of the tick population ecology and tick-borne disease epidemiology. This requires further extensions of the models we developed in Chapter 2 and Charter 3. In a large geographical area where weather conditions may not have significant changes

from one region to another, but local environment conditions may differ to yield different disease incidences. In these cases, we will have multiple time series of incidence for different areas. This calls for clustering analysis of time series in order to identify key factors contributing to different disease spread patterns. Unfortunately, the time series of Lyme disease incidences in Canada as considered potentially available to this thesis study was not available due to the COVID-19 pandemic. We therefore chose the historical Bitcoin prices that is widely available to test the the clustering analyses.

4.0.2 Background

Bitcoin (BTC) is a virtual currency, part of a larger class of digital currencies known as cryptocurrencies, where payments are pseudo-anonymous and independent of third parties such as banks and governments [70, 71, 72, 73]. Introduced in 2009, BTC is a decentralized cryptocurrency utilizing the blockchain technology to record transactions in a public ledger. Unlike fiat currencies and precious metals, Bitcoin is not a physical good, but instead just a record in a database maintained by a network of users known as ”miners”, public to the world [74, 72, 75]. A study by [76, 77] demonstrated that financial assets display common windows with large price changes that tend to cluster together. In [78], the Bayesian change point model is applied to identify and partition structural breaks in the average return and volatility of the Bitcoin price into segments, which then are merged

into seven different regimes based on similar statistical properties. In addition, [79] concludes that realized volatility is predictable from its past values while the fluctuations in Bitcoin price are unpredictable. Furthermore, we encounter drastic changes in behaviour at the market extremes such as the 2008 global financial crisis where volatility, mean returns and correlation patterns in stock returns changed abruptly. Identifying these regime changes is critical for the analysis of equities, fixed income securities and a great number of macroeconomic variables [80]. Given a firm's trading strategies, a fund manager can utilize the current volatility conditions to adjust which ones to give preference to, or when to step away from the market. Clustering analysis plays a critical role in identifying regime changes, and there have been a number of clustering techniques developed and successfully applied for regime changes in time series data. For example, in [81], the TreeGNG algorithm was used for clustering analysis of a high dimensional data set, full of noise. In this chapter we examine how common clustering techniques perform when used to partition the volatility regimes of historical Bitcoin prices. We also analyze the performance of each algorithm on clustering of the data set by comparing the silhouette value.

Trading strategies utilizing heavy shorting positions perform best in bear markets. In contrast, a high frequency strategy will tend to perform best in a highly liquid and volatile market. It may even be the case that a strategy outperforms expectations about its returns. We call these market conditions, volatility regimes, as volatility is the central factor in

identifying these time periods. For instance, the financial crash of 2008, can be weakly coined as a highly volatile, heavily downward trending market incident.

We refer to this as a weak definition, as there isn't a necessary condition on what is deemed high volatility, and heavily down trending. Furthermore we are conditioned to identify simple linear patterns in our world, extending our assumptions indefinitely, while missing out on much of what only a computer might see, and falling victim to countless biases. For example one might think a low volatile, heavily bullish market might be a market state, but this may not even be possible depending on the volatility measurement chosen, as we will see later on. These intuitive groupings we make, are not a systematic approach to differentiating regions in our state space, and also lack concrete threshold levels as well as a statistical backing.

4.0.3 Clustering

Clustering is a mathematical technique that separates a data set into clusters, where the data points in the same cluster share similarity and data points belonging to different clusters are dissimilar [82, 83, 84]. Clustering helps find structures in a data set and is one of main pattern recognition tasks, and widely used to inform strategy decision making by investors, financial creditors, stock holders, etc [83].

In what follows, we will explore the K-Means [85], Ward's Method [86, 87], Complete-

Linkage [88, 89, 90, 91] and Birch [92] algorithms for the identification of Bitcoin volatility regimes. Our analysis is general enough however, to be extended to other cryptocurrencies and asset groups.

The k-nearest neighbors classifier (know as K-Means) is an unsupervised learning algorithm which begins by randomly selecting K cluster centers [93]. The procedure then uses Euclidean distance to minimize the sum of the square of the distance from each point in order points to clusters, and update cluster centers iteratively [94]. By utilizing Lloyd's Algorithm [85], convergence times can be optimized to make K-Means a very fast classifier. According to [82], K-means is a non-hierarchical clustering method and it is considered to build the most compact clusters for classifying the stock data compared to SOM neural network and Fuzzy C-means. However, [95] points out few major drawbacks to applying K-means algorithm such as the sensitivity of the initial condition and the a-prior assumption about the number of clusters. Finally, the convergence of the algorithm is guaranteed [95]. In [96] K-means is implemented to analyze stock clusters in order to mine stock category clusters for investment.

Hierarchical agglomerative clustering methods begin by treating every data point as it own cluster, and iteratively condense "similar" clusters together, until a stopping condition is met, usually the number of clusters desired. The similarity measure can be constructed in many different ways. Ward first describes one such method in 1963 [86], now known as

Ward's Method. The idea is to merge two clusters with the minimum 'cost' as measured by the sum of squared error, effectively minimizing the inter-cluster variances [97]. In [98] Bayesian network has been used to forecast the earnings ratio. When the data is first clustered using the Ward clustering method and then Bayesian network is used for forecasting, the combination increased computational accuracy by 20% for Toyota motor corporation stock price and by 15% NIKKEI stock average.

Complete-linkage is another type of agglomerative hierarchical clustering method [89, 90, 91]. Complete-linkage utilizes a different measure to decide which clusters to combine. It treats the maximum distance that two elements (one in each cluster) within the clusters lie as the distance between clusters [99, 91]. More precisely,

$$D(C_1, C_2) = \max_{a \in C_1, b \in C_2} d(a, b),$$

where C_1, C_2 are the clusters, and $d(a, b)$ is the Euclidean distance between data points a and b.

Birch (Balanced Iterative Reducing and Clustering using Hierarchies) is the first clustering method to handle noise effectively as it uses a CF (clustering feature) tree to compress a large data set[100]. The idea behind Birch is to quickly generate a small approximate subset of clusters in the initial pass of data, and only use those to perform the hierarchical clustering thereon. This allows the algorithm to be orders of magnitude faster in execution than traditional agglomerative methods [101]. With financial data being larger

than ever, it is useful to have such an algorithm that scales with the data. As mentioned in [102], as a hierarchical clustering method that Birch is one downside of this method is the fact of not being able to reverse the merging , the splitting or the exchange between clustering object once the algorithm has been completed.

4.0.4 Resampling and Results

Our training data is taken from `www.BTCe.com`, a Bitcoin exchange which publishes all trading data that passes through their exchange since their inception in 2011. It's important to note that real world trading data has no set sampling rate for transactions. Instead, in one second there may be a large number of orders, while the next may have none. In this study we limit our data sample to BTC data starting from 2015 since the asset was relatively liquid after that. Our goal is to have an hour-by-hour prognosis of the market, hence a minute-by-minute sampling rate is a reasonable choice. Due to the nature of the financial markets, when re-sampling such unstructured data in a given 60 second window, there is the possibility of having more than one measurement, or no measurements at all. In case of more than one measurement, we take the mean of all data points within that time interval in order to avoid approximating the market price as the initial or final price alone. If we have no data over a 60 second window, we proceed by filling the data point as the last entry, known as a forward-fill. This again is intuitive for a price measurement, as the best

approximation of an asset's current value is it's last traded price. One of the measuring tools used in financial instruments is the implied volatility which is calculated by reverse engineering options prices in the market. However, in this study this is not currently viable with the immaturity of the Bitcoin derivatives market. To demonstrate the market direction we calculate short term measurement of the change in price from one minute to the next by the following equation

$$\Delta_t = P_t - P_{t-1}$$

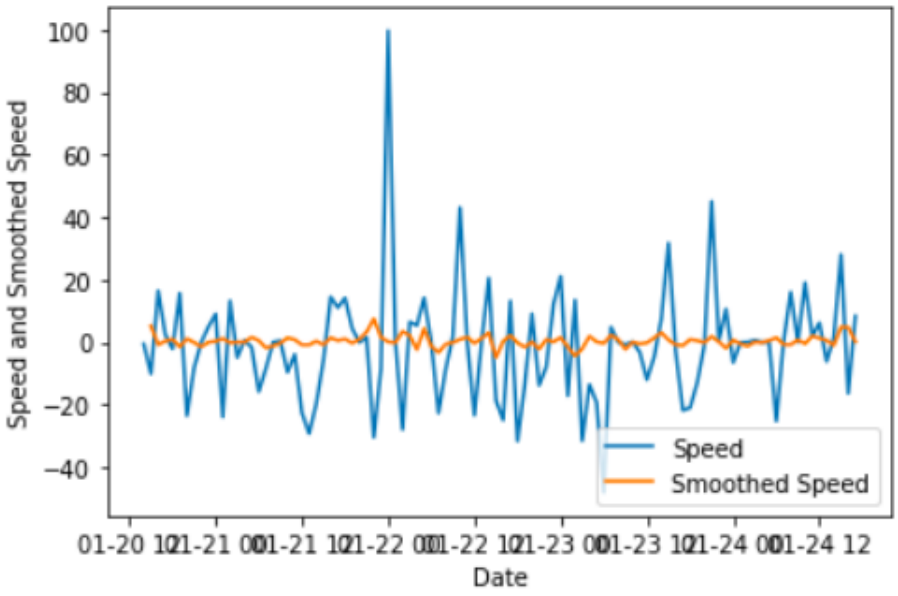


Figure 4.1: The speed data now seems less volatile

The last thing we want to is standardize our different measurements. We do this using

a simple z-score transformation define by

$$z_i = \frac{X_x - \bar{X}}{\sigma_X}$$

with $\bar{X}$ and σ_X being the sample mean and standard deviation respectively, for each measure. Since we may utilize some algorithms which rely on Euclidean distance measurements, it's important that we perform this step in order to assure our measures are of the same scale. Furthermore, since some of the algorithms employ an optimization step, it is ideal to have our data standardized or normalized to assure faster convergence [103]. We choose standardization over normalization in order to keep open boundaries in our domain. In our analysis we use hourly observations but to construct them, we use minute data. This is to assure we are not feeding our algorithms highly correlated data points, since the volatility over the last hour from one minute to the next is nearly identical. We choose to formulate our final data set with one data point per hour, so as to have 8760 points per year's worth of data. The data does display a common shape known as a Markowitz Bullet [104] in finance, which shares similar measurements. There are no obvious disjoint clusters but that is not too troubling as there are specific regions and symmetries present that should partition the data reasonably.

A common heuristic at this point is to use the Elbow Method [105], which looks at the scree plot, and tries to identify where the elbow joint lies on the graph. This however is limited by your scale of the plot, and a complete solution can rely on many different factors

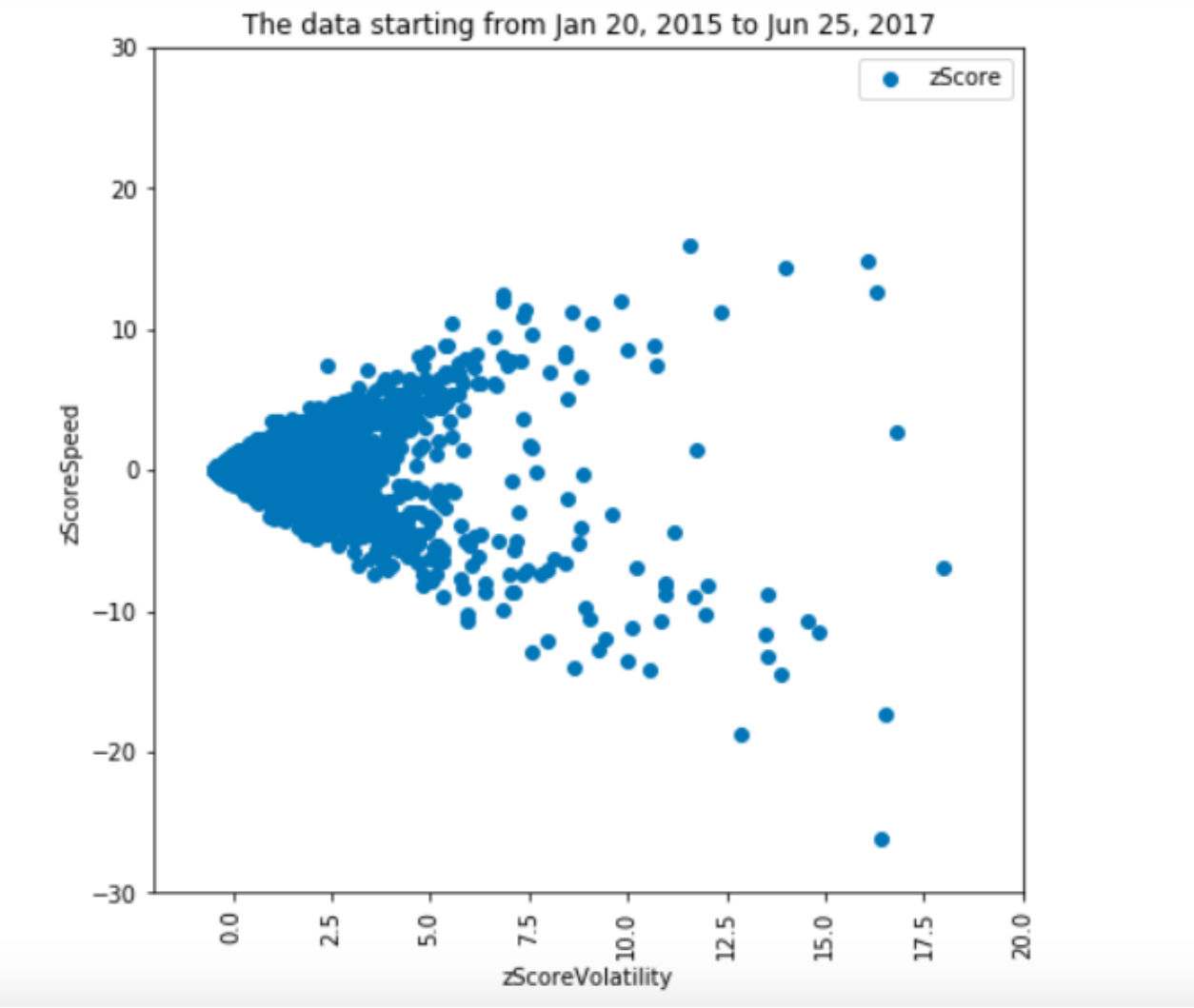


Figure 4.2: The snap of a the data set from January 2015 to Jun 2017

and parameters. Other choices for our clusters is described in greater detail in [106] and more specifically for the k-means algorithm in [107]. We decide to stop adding clusters at the point where any new cluster reduced our total inertia by less than 20%. Since this was at 4 clusters, we decided to continue our analysis with 5 to assure the symmetry criteria mentioned above.

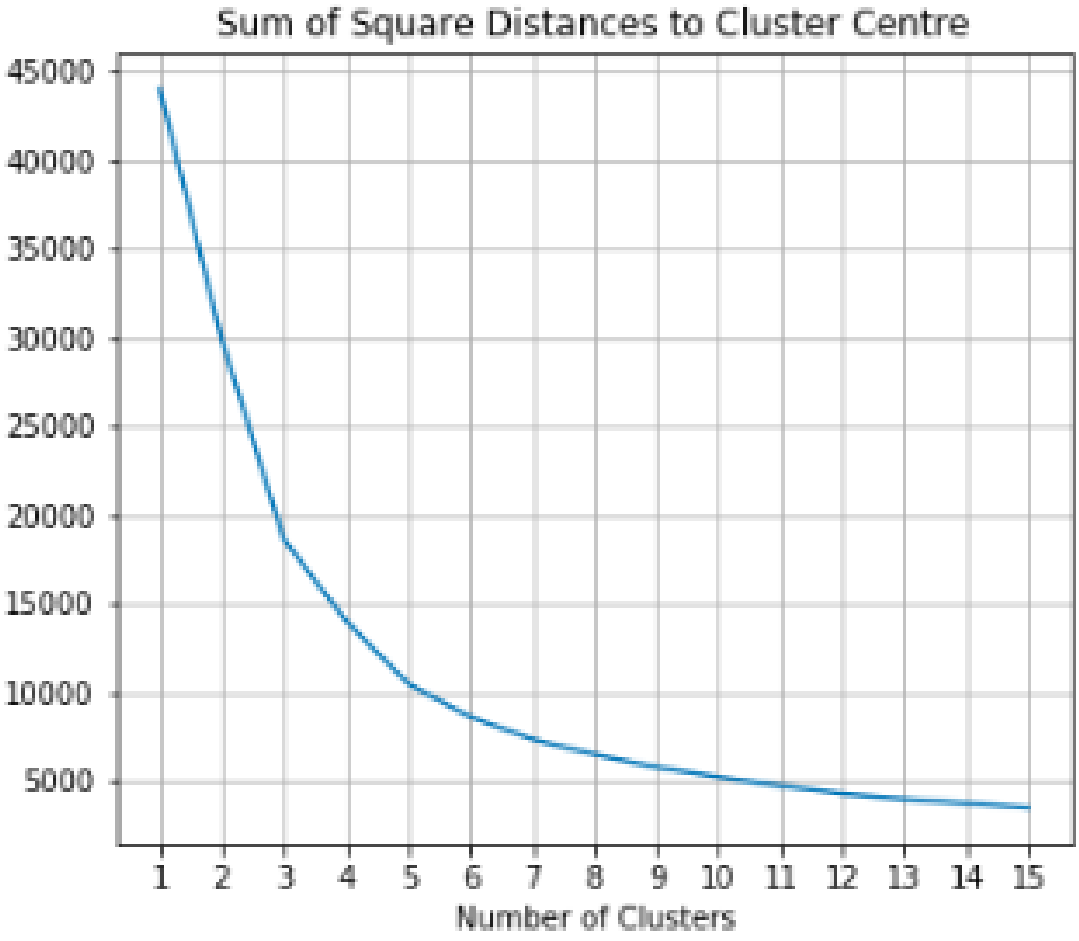
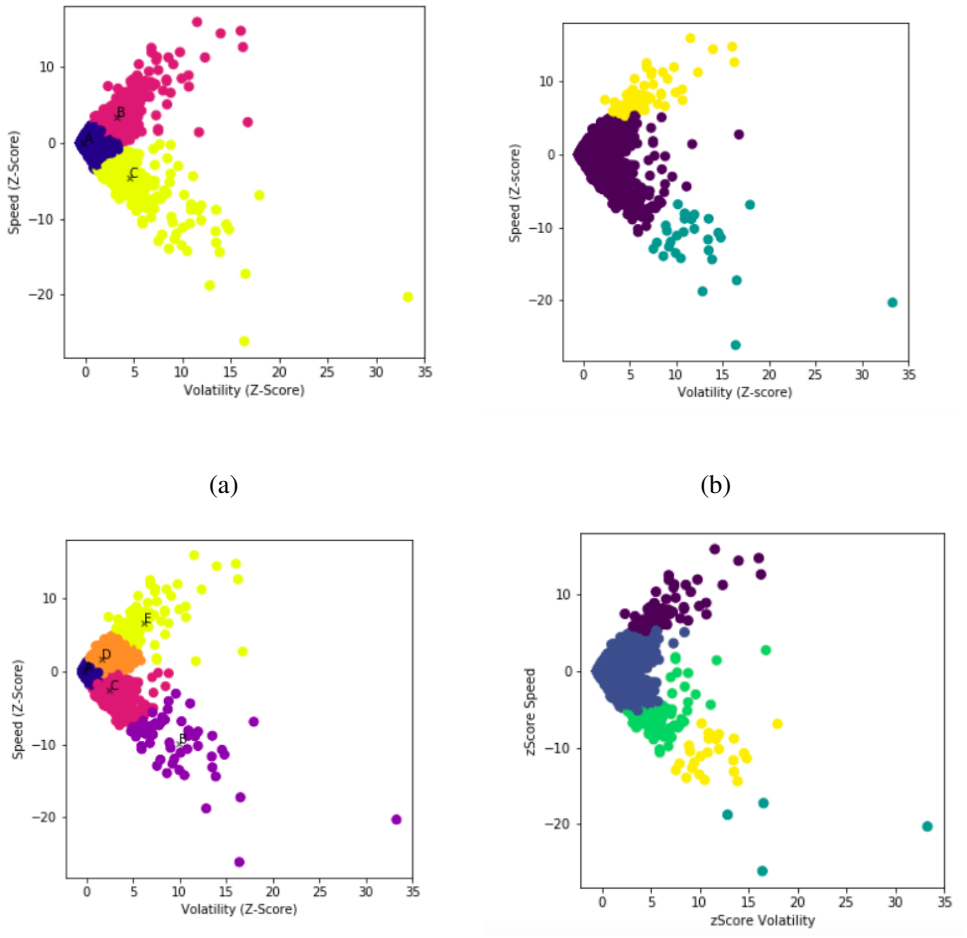


Figure 4.3: Elbow Method of deciding the appropriate number of clusters.



(a) (b)

(c) (d)

Figure 4.4: In Figure a and c the data is clustered using K-MEANS for 3 and 5 clusters respectively. In Figure b and d Birch method is clustering there data into 3 and 5 clusters respectively.

Next, following [108] we calculate silhoutte value which will then validate the performance of each the clustering technique. When the value of silhoutte is close to 1, the data points are well classified. From Table. 4.1 the `Birch` method has a higher silhoutte value than `K-Means` and therefore it is a better model for clustering the data points.

Number of Clusters	`K-Means`	`Birch`
3	0.86	0.91
5	0.8	0.9

Table 4.1: The value of silhoutte for `K-Means` and `Birch` clustering method for different number of clusters.

5 Conclusion and Future Work

In this thesis, we have formulated a compartmental model for lone star ticks and white-tailed deers, where the Ricker function is adopted for birth rate and the migration effect is emphasized. The positivity of solution and the stability of the unique positive equilibrium are proved. Numerical results confirm the theoretical statements, and the sensitivity analysis shows the correlation between equilibrium populations and the parameters. Finally the effects of migration rates and death rates of hosts and ticks are explored, and migration rate has a strong positive correlation with tick and host populations.

Our model can be extended to more complicated situations. For example, the parameter $\hat{q}$ can depend on the population of host, and seasonality can be included in the birth and death rates. Comparison with real data from Canada (such as British Columbia) and USA is ongoing, which could provide more insights into the effects of various parameters. We are also planing to include the life stages, age structure and multiple patches in the model in a future study. The transmission rate could also be represented by a period function depending on time, where seasonality will also be reflected.

In this thesis, we have also formulated a delay differential model for black leg ticks, stratified based on stage and activity, with a particular focus on the host grooming behaviour. The basic reproduction number was calculated and the condition for local stability of tick-free equilibrium, for which the tick population go extinct, and also for existence and uniqueness of a positive equilibrium was given. Model parameterization and numerical simulations were carried out to demonstrate the dynamics of tick and host population with and without the grooming behaviour and the effect of the resistance factor on the value of equilibrium points are studied. Parameters related to the grooming and resistance factors, α_L, α_N, α_A, and κ have no effect on the initial growth rate of ticks since these parameters do not change the value of $\mathcal{R}_0^v$. However, with an increase of the intensity of the grooming behaviour from no resistance to a high level of resistance, where either the hosts show intensified grooming behaviour or ticks are withdrawn from feeding or dead, the values of equilibrium points of all tick stages decrease. From the numerical simulations we observed structural changes of the dynamical behaviour of the tick population by changing the parameter values reflecting the effect of the host resistance. Also, the intensified resistance results in higher equilibrium values for H_{r+}.

A sensitivity analysis of the positive equilibrium value to the parameters was carried out by performing LHS and PRCC. From PRCC we observed high positive correlation between the maximum number of eggs per female adult tick (p) and larvae questing; as

more eggs are produced the higher the number of larvae questing. The female proportion parameter (ϵ) is also positively correlated to larvae questing. As the female rate proportion increases the higher number of egg production and therefore increasing the value of larvae questing. In contrast, the value of the strength of density dependence (q) and death rate of larvae questing (d^{Lq}) are negatively correlated with the population of larvae questing. As the death rate increases there will be a lower population size of larvae questing. Lastly, as the number of larvae questing increases there will be harder to find resources to survive, hence as q increases the number the L_q decreases.

The study represented in this part of the thesis has some limitations. The death rates are assumed to be constants for each stage of the tick and we have ignored the possibility of death during the feeding process resulting from serous exudes which could engulf the tick. Also, interpreting the host resistance as a kind of immunity to ticks we can consider the situation where the host resistance decreases in time the hosts lose immunity to ticks. The molting process is demonstrated by constant delay functions. Future work could incorporate the temperature and humidity on molting process by incorporating a time dependent parameter.

We also conducted time series clustering analysis. Using a group of common clustering algorithms, we identified market volatility regimes using common market measures. Patterns arose that gave way to predictive models not otherwise attainable by common means

of the human eye. More importantly, these cluster divide the market into well defined regions which we can easily check and optimize for given a series of trading algorithms simple back-testing procedures. We aim to perform such back-testing in the future for a more in depth analysis of the strength of our results and predictors.

It was shown that both the K-Means and Birch algorithms were reasonable choices for the division of the data. Birch algorithm performed better than K-Means on clustering our data set. Choices of 3 or 5 clusters were reasonable choices for our total market regimes. Lastly, an improvement to the model was explored as to increase the accuracy of our predictions. We aim to add volume and news sentiment measures to hopefully increase our predictive power. Also, we wish to apply a Markov chain model to each cluster and approximate a market in the time period ahead in a future study.

Vector-borne disease spread patterns in Canada and globally are characterized by their dependence on seasonal variations of the weather conditions and other environmental conditions which vary from one geographical location to another. Understanding these patterns from mathematical modelling point of views requires certain homogeneity in these spatiotemporally varying environments so we can fit the model to the time series of disease incidences. In a future study when some high quality data of time series of Lyme disease incidences become available, we should perform the time series clustering analysis to identify clusters and parametrize the models accordingly.

Bibliography

[1] C. J. Holderman and P. Kaufman. “Lone star tick Amblyomma americanum (Linnaeus)(Acari: Ixodidae)”. In: *EENY580, Entomol. Nematol. Dept. UF/IFAS Extension, Gainesville, FL* (2014).

[2] C. D. Patrick and J. A. Hair. “Oviposition behavior and larval longevity of the lone star tick, Amblyomma americanum (Acarina: Ixodidae), in Different Habitats”. In: *Annals of the Entomological Society of America* 72.2 (1979), pp. 308–312.

[3] J. K. Weeks. “Modeling nymph lone star ticks in the changing landscape of the Virginian Peninsula: a tool for urban planning in the context of human health”. In: *Undergraduate Honors Theses. W&M Pbulish* (2013).

[4] B. Fishcer and S. Myron. “The pricing of options and corporate liabilities”. In: *Journal of Political Economy* 81.3 (1973), pp. 637–654. DOI: 10.1086/260062.

[5] M. P. Nelder et al. “Occurrence and distribution of Ambylomma americanum as determined by passive surveillance in Ontario, Canada (1999–2016)”. In: *Ticks and Tick-borne Diseases* 10.1 (2019), pp. 146–155.

[6] R. Rosa and A. Pugliese. “Effects of tick population dynamics and host densities on the persistence of tick-borne infections”. In: *Mathematical biosciences* 208.1 (2007), pp. 216–240.

[7] J. Feder et al. “Southern tick-associated rash illness (STARI) in the north: STARI following a tick bite in Long Island, New York”. In: *Clinical Infectious Diseases* 53.10 (2011), e142–e146.

[8] J. Goddard and A. S. Varela-Stokes. “Role of the lone star tick, Amblyomma americanum (L.), in Human and Animal Diseases”. In: *Veterinary Parasitology* 160.1-2 (2009), pp. 1–12.

[9] B. E. Anderson et al. “Amblyomma americanum: a potential vector of human ehrlichiosis”. In: *The American Journal of Tropical Medicine and Hygiene* 49.2 (1993), pp. 239–244.

[10] Canada.ca. “CDC”. In: (2018). URL: `https://www.cdc.gov/ticks/diseases/index.htmlCDCwebsite`.

[11] J. K. Khoury et al. “A tick-acquired red meat allergy: A case series”. In: *The American Journal of Emergency Medicine* (2017).

[12] S. E. Wolver et al. “A peculiar cause of anaphylaxis: no more steak?” In: *Journal of General Internal Medicine* 28.2 (2013), pp. 322–325.

[13] S. P. Commins et al. “Delayed anaphylaxis, angioedema, or urticaria after consumption of red meat in patients with IgE antibodies specific for galactose-α-1, 3-galactose”. In: *Journal of Allergy and Clinical Immunology* 123.2 (2009), pp. 426–433.

[14] S. P. Commins and T. A. Platts-Mills. “Delayed anaphylaxis to red meat in patients with IgE specific for galactose alpha-1, 3-galactose (alpha-gal)”. In: *Current Allergy and Asthma Reports* 13.1 (2013), pp. 72–77.

[15] J. L. Kennedy et al. “Galactose-α-1, 3-galactose and delayed anaphylaxis, angioedema, and urticaria in children”. In: *Pediatrics* 131.5 (2013), e1545.

[16] R. K. Raghavan et al. “Current and Future Distribution of the Lone Star Tick, Amblyomma americanum (L.)(Acari: Ixodidae) in North America”. In: *PloS One* 14.1 (2019), e0209082.

[17] D. Chen et al. “Analyzing the correlation between deer habitat and the component of the risk for lyme disease in Eastern Ontario, Canada: A GIS-based approach”. In: *ISPRS International Journal of Geo-Information* 4.1 (2015), pp. 105–123.

[18] N.H. Ogden et al. “Climate change and the potential for range expansion of the Lyme disease vector Ixodes scapularis in Canada”. In: *International Journal for Parasitology* 36.1 (2006), pp. 63–70.

[19] T. R. Hofmeester et al. “Deer presence rather than abundance determines the population density of the sheep tick, Ixodes ricinus, in Dutch forests”. In: *Parasites & Vectors* 10.1 (2017), p. 433.

[20] X. Wu et al. “Developing a temperature-driven map of the basic reproductive number of the emerging tick vector of Lyme disease Ixodes scapularis in Canada”. In: *Journal of Theoretical Biology* 319 (2013), pp. 50–61.

[21] K. Liu, Y. Lou, and J. Wu. “Analysis of an age structured model for tick populations subject to seasonal effects”. In: *Journal of Differential Equations* 263.4 (2017), pp. 2078–2112.

[22] D. Haile and G. Mount. “Computer simulation of population dynamics of the lone star tick, Amblyomma americanum (Acari: Ixodidae)”. In: *Journal of Medical Entomology* 24.3 (1987), pp. 356–369.

[23] G. Mount et al. “New version of LSTSIM for computer simulation of Amblyomma americanum (Acari: Ixodidae) population dynamics”. In: *Journal of Medical Entomology* 30.5 (1993), pp. 843–857.

[24] H. Wang et al. “Simulation of climate-tick-host-landscape interactions: Effects of shifts in the seasonality of host population fluctuations on tick densities”. In: *Journal of Vector Ecology* 40.2 (2015), pp. 247–255.

[25] A. Kaizer et al. “Modeling the biotic and abiotic factors that describe the number of active off-host Amblyomma americanum larvae”. In: *Journal of Vector Ecology* 40.1 (2015), pp. 1–10.

[26] H. D. Gaff and L. J. Gross. “Modeling tick-borne disease: a metapopulation model”. In: *Bulletin of Mathematical Biology* 69.1 (2007), pp. 265–288.

[27] H. Gaff and E. Schaefer. “Metapopulation models in tick-borne disease transmission modelling”. In: *Modelling Parasite Transmission and Control*. Springer, 2010, pp. 51–65.

[28] T. Caraco et al. “Stage-structured infection transmission and a spatial epidemic: a model for Lyme disease”. In: *The American Naturalist* 160.3 (2002), pp. 348–359.

[29] J. M. Heffernan, Y. Lou, and J. Wu. “Range expansion of Ixodes scapularis ticks and of Borrelia burgdorferi by migratory birds”. In: *Discrete Contin Dyn Syst Ser B* 19 (2014), pp. 3147–3167.

[30] G. Baygents and M. Bani-Yaghoub. “A mathematical model to analyze spread of hemorrhagic disease in white-tailed deer population”. In: *Journal of Applied Mathematics and Physics* 5.11 (2017), pp. 2262–2282.

[31] X. Wu and et al. “Stage-structured population systems with temporally periodic delay”. In: *Math. Meth. Appl. Sci* 38 (2015), pp. 3464–3481.

[32] T. M. Kollars et al. “Host associations and seasonal activity of Amblyomma americanum (Acari: Ixodidae) in Missouri”. In: *Journal of Parasitology* 86.5 (2000), pp. 1156–1160.

[33] Xiunan W. and Xiaolin Z. “Dynamics of a Time-Delayed Lyme Disease Model with Seasonality”. In: *SIAM J. Applied Dynamical Systems* 16 (2017), pp. 853–881.

[34] M. Labuda et al. “Amplification of tick-borne encephalitis virus infection during co-feeding of ticks”. In: *Medical and Veterinary Entomology* 7.4 (1993), pp. 339–342.

[35] X. Zhang, X. Wu, and J. Wu. “Critical contact rate for vector host pathogen oscillation involving co-feeding and diapasue”. In: *Journal of Biological Systems* 25.04 (2017), pp. 657–675.

[36] D. Earn et al. *Mathematical epidemiology*. Springer Berlin, 2008.

[37] B. Gomero. “Latin Hypercube Sampling and Partial Rank Correlation Coefficient analysis applied to an optimal control problem”. In: (2012).

[38] S. Marino et al. “A methodology for performing global uncertainty and sensitivity analysis in systems biology”. In: *Journal of Theoretical Biology* 254.1 (2008), pp. 178–196.

[39] S. M. Blower and H. Dowlatabadi. “Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example”. In: *International Statistical Review/Revue Internationale de Statistique* (1994), pp. 229–243.

[40] L Robbin Lindsay et al. “Survival and development of Ixodes scapularis (Acari: Ixodidae) under various climatic conditions in Ontario, Canada”. In: *Journal of Medical Entomology* 32.2 (1995), pp. 143–152.

[41] Nicholas H Ogden et al. “Risk maps for range expansion of the Lyme disease vector, Ixodes scapularis, in Canada now and with climate change”. In: *International Journal of Health Geographics* 7.1 (2008), pp. 1–15.

[42] J. M. Pound et al. “Systemic treatment of white-tailed deer with ivermectin-medicated bait to control free-living populations of lone star ticks (Acari: Ixodidae)”. In: *Journal of Medical Entomology* 33.3 (1996), pp. 385–394.

[43] X. Wu. “Modeling the Impact of Climate Change on Tick Population Dynamics and Tick-Borne Disease Spread.” In: *PhD thesis* (2013).

[44] A.L. Hoch, R.W. Barker, and Jakie A Hair. “Measurement of physical parameters to determine the suitability of modified woodlots as lone star tick habitat”. In: *Journal of Medical Entomology* 8.6 (1971), pp. 725–730.

[45] A. C. Steere, J. Coburn, and L. Glickstein. “The emergence of Lyme disease”. In: *The Journal of Clinical investigation* 113.8 (2004), pp. 1093–1101.

[46] J. Piesman et al. “Duration of tick attachment and Borrelia burgdorferi transmission.” In: *Journal of Clinical Microbiology* 25.3 (1987), pp. 557–558.

[47] A. C. Steere et al. “The spirochetal etiology of Lyme disease”. In: *New England Journal of Medicine* 308.13 (1983), pp. 733–740.

[48] R. C. Johnson et al. “Borrelia burgdorferi sp. nov.: etiologic agent of Lyme disease”. In: *International Journal of Systematic and Evolutionary Microbiology* 34.4 (1984), pp. 496–497.

[49] S. J. Dumler. “Molecular diagnosis of Lyme disease: review and meta-analysis”. In: *Molecular Diagnosis* 6.1 (2001), pp. 1–11.

[50] Canada.ca. “Surveillance of Lyme disease”. In: (2018). URL: https://www.canada.ca/en/public-health/services/diseases/lyme-disease/surveillance-lyme-disease.html.

[51] N. K. Madhav et al. “A dispersal model for the range expansion of blacklegged tick (Acari: Ixodidae)”. In: *Journal of Medical Entomology* 41.5 (2004), pp. 842–852.

[52] S. Randolph. “Tick ecology: processes and patterns behind the epidemiological risk posed by ixodid ticks as vectors”. In: *Parasitology* 129.S1 (2004), S37–S65.

[53] W. Lumsden. “Advances in Parasitology”. In: *Academic Press* 293-309 (1980), pp. 395–400.

[54] B. M. Wagland. “Host resistance to cattle tick (Boophilus microplus) in Brahman (Bos indicus) cattle. II. The dynamics of resistance in previously unexposed and exposed cattle.” In: *Australian Journal of Agricultural Research* 29 (1978), pp. 395–400.

[55] J. Bowessidjaou. “Effects and duration of resistance acquired by rabbits on feeding and egg laying in Ixodes Ricinius.” In: *L. Experientia* 33 (1977), pp. 528–530.

[56] S. J. Brown. “Highlights of Contemporary Research on Host Immune Response to Ticks”. In: *Elsevier Science Poblishers* 28 (1988), pp. 321–334.

[57] S. J. Brown and P. W. Askenase. “Rejection of ticks from guinea pigs by anti-hapten-antibody-mediated degranulation of basophils at cutaneous basophil hypersensitivity sites: role of mediators other than histamine”. In: *The Journal of Immunology* 134.2 (1985), pp. 1160–1165.

[58] G. Fan, H. R. Thieme, and H. Zhu. “Delay differential systems for tick population dynamics”. In: *Journal of Mathematical Biology* 71.5 (2015), pp. 1017–1048.

[59] X. Wang and X. Zhao. “Dynamics of a Time-Delayed Lyme Disease Model with Seasonality”. In: *SIAM. Applied Dynamical Systems* 16 (2017), pp. 853–881.

[60] R. Jennings et al. “How ticks keep ticking in the adversity of host immune reactions”. In: *Journal of Mathematical Biology* 78.5 (2019), pp. 1331–1364.

[61] N. H. Ogden et al. “A dynamic population model to investigate effects of climate on geographic range and seasonality of the tick Ixodes scapularis”. In: *International Journal for Parasitology* 35.4 (2005), pp. 375–389.

[62] X. Wu, V. RSK Duvvuri, and J. Wu. “Modeling dynamical temperature influence on tick Ixodes scapularis population”. In: *International Congress on Environmental Modelling and Software* 223 (2010).

[63] Y. Lou and J. Wu. “Tick seeking assumptions and their implications for Lyme disease predictions”. In: *Ecological Complexity* 17 (2014), pp. 99–106.

[64] R. Rosà et al. “Thresholds for disease persistence in models for tick-borne infections including non-viraemic transmission, extended feeding and tick aggregation”. In: *Journal of Theoretical Biology* 224.3 (2003), pp. 359–376.

[65] H. D. Gaff and L. J. Gross. “Modeling Tick-Borne Disease: A Metapopulation Model”. In: *Bulletin of Mathematical Biology* 69 (2007), pp. 265–288.

[66] R. Rosà and A. Pugliese. “Effects of tick population dynamics and host densities on the persistence of tick-borne infections”. In: *Mathematical Biosciences* 208.1 (2007), pp. 216–240.

[67] H. L. Smith. “Monotone Semiflows Generated by Functional Differential Equations .” In: *Journal Of Differential Equations* 66 (1987), pp. 420–442.

[68] S. M. Blower and H. Dowlatabadi. “Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example”. In: *International Statistical Review/Revue Internationale de Statistique* (1994), pp. 229–243.

[69] S. Marino et al. “A methodology for performing global uncertainty and sensitivity analysis in systems biology”. In: *Journal of Theoretical Biology* 254.1 (2008), pp. 178–196.

[70] B. Segendorf. “What is bitcoin”. In: *Sveriges Riksbank Economic Review* 2 (2014), pp. 71–87.

[71] S. Nakamoto. “Bitcoin: A peer-to-peer electronic cash system”. In: (2008).

[72] A. M. Antonopoulos. *Mastering Bitcoin: unlocking digital cryptocurrencies*. ” O’Reilly Media, Inc.”, 2014.

[73] M. Crosby et al. “Blockchain technology: Beyond bitcoin”. In: *Applied Innovation* 2.6-10 (2016), p. 71.

[74] A. Salman and M. G. Razzaq. “Bitcoin and the world of digital currencies”. In: *Financial Management from an Emerging Market Perspective* (2018), pp. 271–281.

[75] L. Cocco and M. Marchesi. “Modeling and simulation of the economics of mining in the Bitcoin market”. In: *Plos One* 11.10 (2016), e0164603.

[76] J. Fan and I. Kleshchelski. “Trend following and volatility regimes”. In: (2017).

[77] R. Cont. “Volatility clustering in financial markets: empirical facts and agent-based models”. In: *Long Memory in Economics*. Springer, 2007, pp. 289–309.

[78] S. Thies and P. Molnár. “Bayesian change point analysis of Bitcoin returns”. In: *Finance Research Letters* 27 (2018), pp. 223–227.

[79] H. A. Aalborg, P. Molnár, and J. E. Vries. “What can explain the price, volatility and trading volume of Bitcoin”. In: *Finance Research Letters* 29 (2019), pp. 255–265.

[80] A. Ang and A. Timmermann. “Regime changes and financial markets”. In: *Annu. Rev. Financ. Econ.* 4.1 (2012), pp. 313–337.

[81] K. A. Doherty et al. “Hierarchical topological clustering learns stock market sectors”. In: *2005 ICSC Congress on Computational Intelligence Methods and Applications*. IEEE. 2005, 6–pp.

[82] S. Nanda, B. Mahanty, and M. Tiwari. “Clustering Indian stock market data for portfolio management”. In: *Expert Systems with Applications* 37.12 (2010), pp. 8793–8798.

[83] M. Momeni, M. Mohseni, and M. Soofi. “Clustering stock market companies via k-means algorithm”. In: *Kuwait Chapter of Arabian Journal of Business and Management Review* 33.2578 (2015), pp. 1–10.

[84] G. Guojun, M. Chaoqun, and J. Wu. *Data Clustering: Theory, Algorithms, and Applications*. ”ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadephia”, 2007.

[85] S. Lloyd. “Least squares quantization in PCM”. In: *IEEE Transactions on Information Theory* 28 (2 1982), pp. 129–137.

[86] F. Murtagh and P. Legendre. “Ward’s hierarchical clustering method: Clustering criterion and agglomerative algorithm”. In: *arXiv preprint arXiv:1111.6285* (2011).

[87] J. Ward. “Hierarchical Grouping to Optimize an Objective Function”. In: *Journal of the American Statistical Association* 58 (1963), pp. 236–244.

[88] D. Defays. “An efficient algorithm for a complete link method”. In: *The Computer Journal. British Computer Society* 20 (1977), pp. 364–366.

[89] W. H. Day and H. Edelsbrunner. “Efficient algorithms for agglomerative hierarchical clustering methods”. In: *Journal of Classification* 1.1 (1984), pp. 7–24.

[90] I. Davidson and SS. Ravi. “Agglomerative hierarchical clustering with constraints: Theoretical and empirical results”. In: *European Conference on Principles of Data Mining and Knowledge Discovery*. Springer. 2005, pp. 59–70.

[91] K. Sasirekha and P. Baby. “Agglomerative hierarchical clustering algorithm”. In: *International Journal of Scientific and Research Publications* 83 (2013), p. 83.

[92] T. Zhang, R. Ramakrishnan, and M. Livny. “BIRCH: an efficient data clustering method for very large databases”. In: *Proceedings of the 1996 ACM SIGMOD*

International Conference on Management of Data - SIGMOD '96 (1996), pp. 103–114.

[93] K. Wu, Y. Wu, and H. Lee. "Stock trend prediction by using K-Means and Apriori all algorithm for sequential chart pattern mining." In: *J. Inf. Sci. Eng.* 30.3 (2014), pp. 669–686.

[94] M. Mahajan, P. Nimbhorkar, and K. Varadarajan. "The planar k-means problem is NP-hard". In: *International Workshop on Algorithms and Computation*. Springer. 2009, pp. 274–285.

[95] J. M. Pena, J. A. Lozano, and P. Larranaga. "An empirical comparison of four initialization methods for the k-means algorithm". In: *Pattern Recognition Letters* 20.10 (1999), pp. 1027–1040.

[96] S. Liao, H. Ho, and H. Lin. "Mining stock category association and cluster on Taiwan stock market". In: *Expert Systems with Applications* 35.1-2 (2008), pp. 19–29.

[97] R. C. de Amorim. "Feature relevance in wards hierarchical clustering using the L p norm". In: *Journal of Classification* 32.1 (2015), pp. 46–62.

[98] Y. Zuo and E. Kita. "Stock price forecast using Bayesian network". In: *Expert Systems with Applications* 39.8 (2012), pp. 6729–6737.

[99] R. L. de Mantaras and L. Saitia. “Comparing conceptual, divisive and agglomerative clustering for learning taxonomies from text”. In: *16th European Conference on Artificial Intelligence Conference proceedings*. Vol. 110. 2004, p. 435.

[100] L. O’callaghan et al. “Streaming-data algorithms for high-quality clustering”. In: *Proceedings 18th International Conference on Data Engineering*. IEEE. 2002, pp. 685–694.

[101] D. Shalini, M. Shashi, and A. Sowjanya. “Mining frequent patterns of stock data using hybrid clustering”. In: *2011 Annual IEEE India Conference*. IEEE. 2011, pp. 1–4.

[102] B. Bini and T. Mathew. “Clustering and regression techniques for stock prediction”. In: *Procedia Technology* 24 (2016), pp. 1248–1255.

[103] I. Mohamad and D. Usman. “Standardization and Its effects on K-Means clustering algorithm”. In: *Research Journal of Applied Sciences, Engineering and Technology* 6 (Sept. 2013), pp. 3299–3303.

[104] H. Markowitz. “Portfolio Selection”. In: *The Journal of Finance* 7.1 (1952), pp. 77–91. ISSN: 00221082, 15406261. URL: http://www.jstor.org/stable/2975974.

[105] D. J. Ketchen and C. L. Shook. “The application of cluster analysis in strategic management research: an analysis and critique”. In: *Strategic Management Journal* 17.6 (1996), pp. 441–458.

[106] Y. Jung et al. “A decision criterion for the optimal number of clusters in hierarchical clustering”. In: (2002).

[107] D.T. Pham, S.S. Dimov, and C.D. Nguyen. “Selection of K in K-means clustering”. In: *J. Mechanical Engineering Science* 219 (2004), pp. 103–119.

[108] P. J. Rousseeuw. “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis”. In: *Journal of Computational and Applied Mathematics* 20 (1987), pp. 53–65.