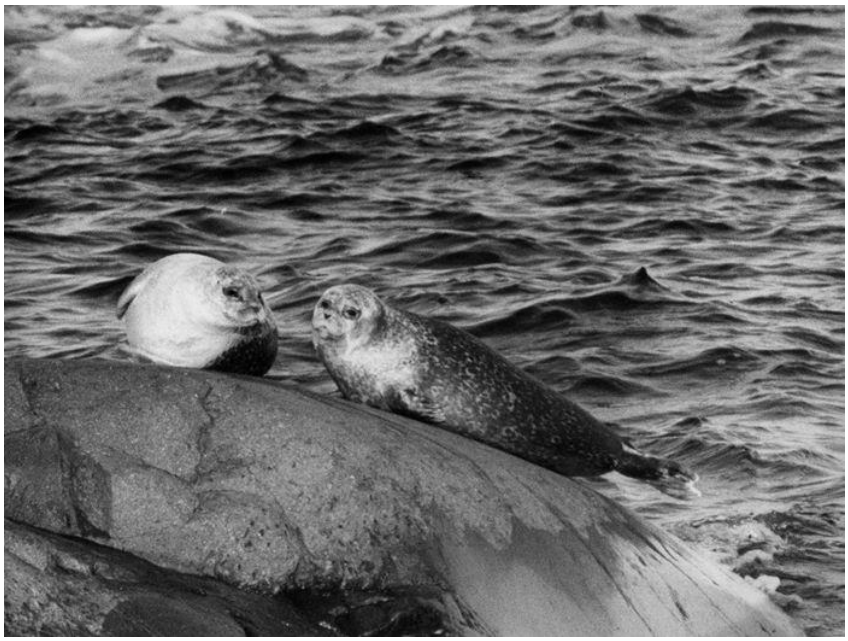
Epidemiological modelling of PDV epizootic in Harbor seals



Mimmi Eriksson

Degree project for Bachelor of Science in Biology

BIO602 Biology: Degree project 15 hec
Spring 2015

Department of Biological and Environmental Sciences
University of Gothenburg

Examiner: Susanne Eriksson
Department of Biological and Environmental Sciences
University of Gothenburg

Supervisor: Karin Hårding
Department of Biological and Environmental Sciences
University of Gothenburg

Faculty of Science

UNIVERSITY OF GOTHENBURG

Front page photo taken by and copyright: Tero Härkönen

Table of Contents

Abstract

English

During the last 30 years a deadly epizootic has affected the Harbour seal colonies in northern Europe twice, including the Swedish oceanic regions Skagerrak and Kattegat. Both outbreaks resulted in mass death in the seal colonies with mortality rates varying between 10% and 60%. The reason behind the epizootics was found to be a virus, later named Phocine distemper virus (PDV).
In this study different factors that may have affected the course of the epizootic are investigated. A stochastic (SEIR) model for epizootic spread is constructed, which simulates how the number of healthy, exposed, infected, immune and dead individuals changes each day during an outbreak. A smaller literature study was performed searching for parameter values on the factors that could influence the epizootic. The different parameters were tested in the stochastic model and compared with empirical data from the PDV outbreak. The empirical data are on the number of dead seals from day-to-day in the first outbreak in Koster (1988).
Two parameters were identified as vital for the outcome of the epizootic, and how well the cumulative curve of dead seals matched empirical data: daily infection rate and recovery rate.
Daily infection rate was found to have a higher influence than the recovery rate on the course of an epizootic, since small changes in the infection value created a larger difference in the outcome.
Future studies could look into the how the virulence may differ between colonies and why. It would also be important to consider that the infection rate can vary at a smaller scale, and differ from day to day during the infectious period. The construction of a detailed SEIR model designed for PDV can generate a better understanding of why the mortality rates varied among colonies.

Swedish

Under de senaste 30 åren har en dödlig epizooti brutit ut vid två separata tillfällen i knubbsäls kolonier i norra Europa, däribland de svenska kustregionerna i Skagerrak och Kattegatt. Båda utbrotten orsakade massdöd i sälkolonierna med en dödlighet som varierade mellan 10 % och 60 %. Orsaken bakom epizootin visade sig vara ett virus, som sedan döptes till Phocine distemper virus (PDV).
I denna studie undersöks olika faktorer som inverkar på hur en epizooti utvecklas. En stokastisk (SEIR) modell över spridning av en epizootisk sjukdom utvecklades, denna simulerar hur antalet friska, exponerade, infekterade, immuna och döda individer förändras dag för dag under ett epizootiutbrott. En mindre litteraturstudie genomfördes för att söka efter parametervärden på faktorer som kan påverka spridningen inom en koloni. Dessa parametrar testades sedan i den stokastiska modellen och jämfördes med empiriska data från PDV utbrottet. Datan bygger på antalet döda sälar dag för dag under det första utbrottet i Koster (1988). Det kumulativa antalet döda sälar jämfördes med prediktioner från modellen.
Två parametrar i modellen identifierades som vitala för utfallet av en epizooti, och för hur väl den kumulativa kurvan över döda sälar matchade empiriska data: daglig infektionshastighet och tillfriskningshastighet. Den dagliga infektionshastigheten (risken att en individ överför smittan) påverkar förloppet av epizootin i högre grad än tillfrisknings-hastigheten av de redan sjuka, då små ändringar i infektionsvärdet skapade större förändring i det slutgiltliga antalet.
Fortsatta studier skulle kunna undersöka den dagliga dödshastigheten (virulensen) och hur den inverkar på sjukdomsförloppet. Det vore även viktigt att överväga att infektionshastigheten kan variera på en finare skala, mellan olika dagar i stadiet som infekterad. Att utveckla en skräddarsydd SEIR modell för PDV kan leda till en bättre förståelse för varför mortaliteten varierar mellan olika sälkolonierna.

Keywords

Phocine distemper virus, PDV, SEIR, Phoca vitulina, European Harbour seal

Introduction

Two epizootics and the virus

Phocine distemper virus (PDV here on) has affected the population of European Harbour seals (*Phoca vitulina*) at two different occasions, 1988 and 2002 (Dietz, Heide-Jorgensen, & Harkonen, 1989; Harding, Harkonen, & Caswell, 2002). It is a virus from the morbillivirus genus and was unknown to science before the outbreak in 1988, it was first considered as a version of Canine distemper virus, CDV, which is common in terrestrial carnivores (Cosby et al., 1988; Mahy, Barrett, Evans, Anderson, & Bostock, 1988; Osterhaus & Vedder, 1988).

Both of these epizootics caused a mass death in the population. During the epizootic in 1988 somewhere around 18 000-23 000 seals died (Dietz et al., 1989; T. Harkonen et al., 2006; ICES, 1990) and mortality in the population was estimated to be 40-60% (T. Harkonen & Heide-Jorgensen, 1990; Heide-Jorgensen, Harkonen, Dietz, & Thompson, 1992; ICES, 1990) everywhere besides in the coastal area of Scotland, were it was estimated to be just about 10-20% (ICES, 1990; P. M. Thompson & Miller, 1992). The outbreak in 2002 killed approximately 30 000 seals, but this time there was a pattern of lower mortality rates than during the 1988 epizootic in most locations. However in the northern Skagerrak it was higher and reached around 66% (T. Harkonen et al., 2006) in contrast to the 56% estimated in 1988 (T. Harkonen & Heide-Jorgensen, 1990).

Both epizootics started at Anholt an island outside Denmark, and continued to spread south towards the Baltic Sea, north along the Swedish coast and towards Norway, south-west past Germany, Holland over to Great Britain, Ireland and last infected was the coast of Scotland (Dietz et al., 1989; T. Harkonen et al., 2006). The whole epizootic lasted for about 9 months during both outbreaks, however each colony harboured the disease for 6 to 13 weeks with the exception of the Moray Firth and Tay regions (with 14-19 weeks longer epizootic period). There was also a difference in the duration of the infection among colonies, with the biggest difference in the Wadden Sea (5 week longer outbreak in the year 1988) (T. Harkonen et al., 2006).

According to T. Harkonen et al. (2006) the outbreak in 2002 started about three weeks later than the one in 1988, 4 May instead of 12 April. They also show that the epizootic during 2002 in general occurred consistently later in the different colonies than during the 1988 epizootic.

Possible reasons behind epizootic development

When looking at accumulative death curves from different colonies it is observed that they differ from each other. There are several different possibilities for this variation.

Firstly there are differences in the behaviour and habits of seals between age groups, sexes and season. P. M. Thompson, Fedak, McConnell, and Nicholas (1989) found it to be a difference between males and females in migratory and haul-out behaviour during the year. In the early summer during the pupping season the females spent most of their time inside the study area and haul-out almost every day. If they left the area they travelled up to 10 km away and never visited a haul-out site outside. This behaviour remained more or less the same even during moulting. The males showed two different patterns, one where they spent short periods in haul-out sites in the study area and then went approximately 15 km outside it for several days. The other pattern showed that they hauled-out most days in the area and made shorter trips outside, less than 24 hours. When the males started the moulting period they spent every day at haul-out sites and the time they spent ashore increased. P. M. Thompson et al. (1989) also showed that the tagged seals spent less time hauled-out inside the study area and took longer trips

outside during winter. But since they only had four tagged seals, one female and three males, the results were not statistically significant. However a study by Cunningham et al. (2009) showed that females spent less time than males hauled-out between October and May and that the females spent more time hauled-out between June and September than the males. A similar result is presented by Dietz, Teilmann, Andersen, Riget, and Olsen (2013) where they saw that Harbour seals that were tagged between September and June had a wider dispersal than the ones tagged between April and June.
Cunningham et al. (2009) also found that individuals often had two main haul-out sites and just briefly visited other sites while travelling between these two.

When it comes to differences between age groups, younger seals swam longer distances in the area, ranging about 85 km in a weekly average distance for a juvenile (less than a year), to an average of about 35 for a sub-adult (between 1-3 years) and about 20 km for an adult (older than 3 years) (Dietz et al., 2013).
In another study made by Klepac et al. (2009) they tested four different models looking at transmission dynamics between and within age groups. The model that fitted their data the best was one that was density dependent and with heterogeneous mixing between groups. When they used this model to calculate transmission rate between these classes they found that the highest rate were between juveniles and sub-adults ($8.99 \cdot 10^{-5}$), followed by juvenile to juvenile ($3.26 \cdot 10^{-5}$) transmission. All classes containing adults had the lowest transmission rate with sub-adult to adult at the bottom ($1.75 \cdot 10^{-6}$). With this and the findings about movement behaviour in mind, one could suspect that the dynamics of this effects the spreading and development of an epizootic. Even though the model made here will (to begin with) look at a colony as a whole with the same probabilities for all individuals to get infected, to recover and to die. The parameters will also have the same values during all days throughout the epizootic.

Besides behavioural differences between age and sex groups there could also be a genetic component behind how the different colonies were effected and the ways the epizootic developed within them. According to Goodman (1998) one can say that the colonies in the area form a metapopulation which can be divided into six populations. And since these populations are more or less separated from each other and are not mixing genetically there could be colonies with a genetic variation that are not as vulnerable to the virus. However Tero Harkonen, Harding, Rasmussen, Teilmann, and Dietz (2007) found that a higher number of adults died in comparison to sub-adults. They state that this shows that the risk of dying within a colony does not originate from genetic differences, but still it may be a genetic difference to look into between populations.
In the same paper Tero Harkonen et al. (2007) showed that there were almost no dead seals found over the age of 14 years old in the 2002 outbreak, probably because they developed immunity during the 1988 epizootic. This might have been a contributing factor in the epizootic differences between colonies in 2002. In addition to immunity another contributing factor can be the annual population growth. The colonies in Scotland had a low yearly growth rate, estimated to 6% (D. Thompson, Lonergan, & Duck, 2005), although the colonies in Skagerrak and Kattegat had a yearly growth rate estimated to be around 12-13% (Harding et al., 2002; T. Harkonen, Harding, & Heide-Jorgensen, 2002)

Another factor to consider is environmental contaminants that could bio-accumulate in the seals and that this might influence the immune system. de Swart, Ross, Vos, and Osterhaus (1996) tested this and found that seals with higher levels of bio-accumulated environmental contaminants had a weaker function of the NK- and T-cell. According to this study NK- and T-cells are important when it comes to fighting off viral infections and this can indeed be a co-

factor responsible for the mass death of seals during the two epizootics. Ross et al. (1995) supports the finding that there is an impaired immune response in seals with accumulated environmental contaminants.

Modelling epizootics and previous work

A common way to better understand how an epizootic operates in a population is to use a SIR model, first formulated by Kermack and McKendrick (1991). The population is divided into groups of susceptible (S), infected (I) and recovered (R) individuals and their dynamics are modelled by three differential equations.

In epidemiological modelling there are two approaches to handle population size, one is density dependent, where the transmission rate are allowed to vary with population size. The other option is a 'frequency dependent' model and they are often used when contact rate between individuals is independent of the population size. (Begon et al., 2002)
A frequency dependent model is the usual approach when it comes to Harbour seals because of their habit of hauling out in aggregated colonies.
In a density dependent model used by Begon et al. (2002), the relationship between susceptible and infected individuals in the populations is represented by the equation:

$$t = \mathrm{SI}$$

While in a frequency dependent model this relationship is represented as:

$$t = {}^{\mathrm{SI}}/_{N}$$

Different approaches with this type of modelling has been done with PDV and the Harbour seal population.

de Koeijer, Diekmann, and Reijnders (1998) studied different colonies and how different parameters varied between them and whether this could be explained by environmental toxins. To do this they modified the model to consider the fact that seals sometimes haul out and gather in big herds. They argue that this should effect the contact rate and make it dependent on the number and density of the herd. They concluded that the most important parameter to determine the course of the epizootic is the survival rate. If it is low the force of infection will be high since the remaining population consists mostly of susceptible individuals, thus if the survival rate is high there will be a high number of recovered individuals in the population and the force of infection will be low. In the end de Koeijer et al. (1998) mention that the survival rate are similar in most areas or in the polluted areas but not in the less polluted Scottish area, where the survival rate is higher. This could be a pointer towards pollutants having a role to play in how the epizootic effects different colonies.

Klepac et al. (2009) used a SIR model where they added age groups and tested different transmission rates between the groups, from the same between all groups to differences in transmission rates between all of groups. As mentioned above the transmission rates were highest between sub-adults ($8.99 \cdot 10^{-5}$) and lowest between sub-adult to adult ($1.75 \cdot 10^{-6}$).

In a study by Murray (2009) they used SIR models to make four different models of how different diseases can be transmitted, two similar to terrestrial transmission and two adapted to transmission in water. They used the second terrestrial model, a frequency dependent model to describe how PDV spreads.

Another way of looking at epizootics is with a SEIR model, which basically is similar to a SIR model but also keeps track of the group with individuals in a latency stage. I.e. they are infected but without showing symptoms and they are not yet contagious.

A stochastic version of a SEIR model was used by Heide-Jorgensen and Harkonen (1992). They estimated the probability of infection between different colonies in the eastern North Sea and the intensity of the virus. With this they tested their model and saw that there is no difference in how the epizootic develops between colonies depending on its size.

Swinton, Harwood, Grenfell, and Gilligan (1998) used a SEIR model to investigate whether there is a threshold for the population size where the disease cannot persist and if so, how large it is. They did find that a threshold like this exist, but it is at a level which the seal population will probably never reach (10^8 individuals).

A common parameter to estimate for an epizootic is the basic reproductive number, R_0. This is an estimation of how many new infected individuals one infected individual will produce on average during its infectious stage, in a completely susceptible population.
The R_0-value tells if an epizootic will occur or not, hence if $R_0<1$ an epizootic will not break out, since this means that one infected individual will infected less than one other individual. Which gives that if $R_0>1$ the epizootic will cause an outbreak. (Andersson & Brittin, 2000)
Estimations of R_0 has been done for the different colonies the north Europe region.
de Koeijer et al. (1998) calculated a value of between 2.1 and 3.0, in Swinton et al. (1998) they estimate R_0 to be 2.8 and Klepac et al. (2009) looked at the population in the Wadden sea and found that if one assumes density dependency the R_0 would be approximately 2.03 and for frequency dependency 2.2. From this the R_0 could be said to be somewhere between 2 and 3 newly infected per infected seal.

Purpose

The aim of this study were to describe the PDV outbreak with an epidemiological model that was adopted to the epidemiology of PDV in Harbour seal. A focus was to test the sensitivity of the model towards different factors that can affect the spreading of the virus in the population, e.g. whether the course of the disease differs depending on what time of the year the epizootic starts, whether it is different in a colony depending on its population size and also if the size effects the death rate. Different parameters, virulence, contagiousness and other assumptions like latent period and infection time in the model were to be investigated
A stochastic SEIR model was developed and recommendations are given on how it can be extended.

Method

The data used in this model come from a large data-base covering both PDV epizootics with the cumulative number of dead Harbour seals from all localities along Europe. More specifically the focus was on the data over the outbreak in Koster during 1988. In addition estimations of population sizes from aerial censuses were available (Harding unpublished data).

To simulate the population according to a frequency dependent SEIR model, the groups susceptible (S), exposed (E), infected (I) and recovered (R) are represented in a deterministic way through these equations:

$$S_{t+1} = -\beta \frac{S_t I_t}{N} - mS_t$$

$$E_{t+1} = \beta \frac{S_t I_t}{N} - (m + \sigma)E_t$$

$$I_{t+1} = \sigma E_t - (m + v + \delta)I_t$$

$$R_{t+1} = \delta I_t - mR_t$$

Where β represents the infection rate, σ the rate of which exposed individuals develop the disease and become infected. δ is the recovery rate, m the natural mortality and v the added mortality from the infection. The small subscript t denotes time.

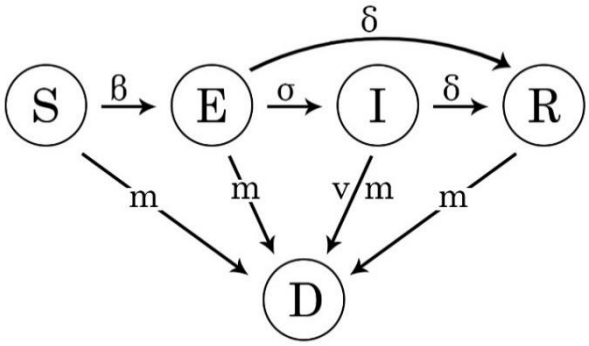


Figure 1. Illustration of the epidemiological model for PDV, circles are states of an individual, arrows indicate the different parameters that determine transitions between these states.

In addition to the basic equations, one for dead seals were added.

$$D_{t+1} = mS_t + mE_t + (m + v)I_t + mR_t$$

All these equations are illustrated in Figure 1 where you can see how individuals move between the different states and which parameters effects this distribution.

The infection rate β and the recovery rate δ in this model are assumed to be the same for all individuals in the population during all stages in the epizootic period. The same assumption is made for the two different mortalities, *m* and *v*.

Instead of using σ as the rate of individuals that gets removed from E into I, it is assumed that an individual will always stay as E for 3 days before moving into I, unless it recovers by chance. Also if an infected individual survives for 12 days it will automatically move to the R category. These are the mean values of latency time (3 days) and maximal time spent infected (12 days)(Heide-Jorgensen & Harkonen, 1992).

To go from the deterministic model towards a stochastic one, the proportions of recovered and dead (both by natural- and virus-induced mortality) in every time step were calculated as a binomial random variable with parameters (S, E, I or R) and (m, v or δ).

Similarly, the number of newly infected individuals were also calculated as a random binominal number from the population S_t and probability of infection p. The probability of getting infected was estimated as a function depending on the number of infected and the infection rate

$$p = 1 - (1 - \beta)^{I_t}$$

To estimate a value of both the natural mortality, *m*, and the mortality due to disease, *v*, they had to be converted from mortality during a period of time to mortality during one time step, this was done as follows.
The virus induced mortality can be described by this equation:

$$v_d + (1 - v_d) \cdot v_d + (1 - v_d)^2 \cdot v_d + (1 - v_d)^3 \cdot v_d \ldots + (1 - v_d)^{time\ steps} = v$$

But to simplify, we calculate probability of survival:

$$(1 - v_d)^{time\ steps} = s$$

To get equation for v_d:

$$s = (1 - v_d)^{time\ steps} \rightarrow s^{1/time\ steps} = 1 - v_d \rightarrow v_d = 1 - s^{1/time\ steps}$$

Solving for v_d:

$$v_d = 1 - 0.4^{1/12} \approx 0.0735 \qquad (1)$$

The value 0.4 comes from 1-0.60, where 0.60 is the estimated mortality during the whole epizootic in Skagerrak during the 1988 outbreak (T. Harkonen & Heide-Jorgensen, 1990; Heide-Jorgensen et al., 1992; ICES, 1990). 12 is used for the number of time steps, since this is the number of days an individual is classified as infectious in the model.

The same can be done for estimating the probability of dying because of natural mortality during one day, giving:

$$m_d = 1 - 0.94^{1/365} \approx 0.000195 \approx 1.695 \cdot 10^{-4}$$

The survival rate of 0.94 during one year is used since T. Harkonen et al. (2002) claims that the survival rate for the population in the Skagerrak region is 0.92-096, however this is the survival rate of adults.

A simulation of the model described above was performed in MatLab R2015a.

Results

A test of how the natural mortality effects the simulation was made and is shown in Table 1. This shows that the end result of the simulation is similar with and without natural mortality. The biggest difference is that with natural mortality the simulation continues to grow with a mean value of 9 individuals during 100 time steps.

Table 1. Shows the final number of dead by the simulation with and without natural mortality m. 20 simulations were made, the mean value and standard deviation were calculated.

Parameters	Without m	With m
Time step 110	851 SD±15.54	853 SD±16.36
Time step 210	851 SD±15.54	862 SD±15.99
Average growth	0	9 SD±3.27

Different values for both infection rate β and recovery rate δ were tested in the model to see how these changes affect the outcome of the simulated cumulative death curve. The values of the parameters used for the simulations are shown in Table 2 and the different β and δ values tested are shown in Table 3.

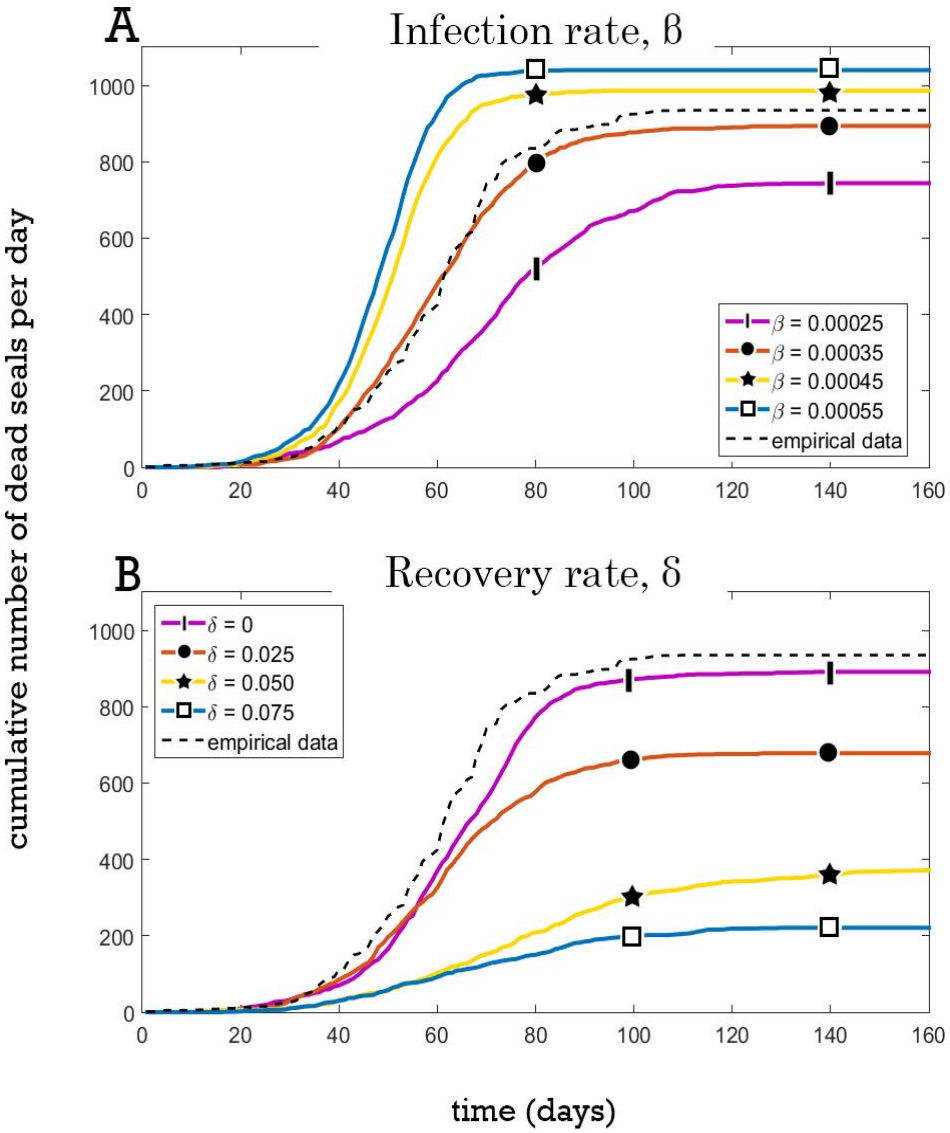


Figure 2. The simulations made by the model and how they vary from the empirical data depending on changes in the two parameters, recovery rate and infection rate. The x-axis shows time in days and the y-axis shows cumulative number of dead individuals per day. ***2A,*** *the different scenarios by change in infection rate but with a fixed recovery rate (set to 0).* ***2B,*** *the different scenarios by changing the recovery rate with a fixed infection rate (set to 0.00035).*

How the variation in recovery rate values affects the form of the curve are shown in Figure 2B, where it can be seen that the higher the recovery rate are, the slower the curves grow and fewer individuals die.

When it comes to the variations in infection rate, that are illustrated in Figure 2A, it can be seen that it affects the curves in roughly the same way, how many that die and how fast the dead accumulate, although with a smaller difference between the values used (see Table 3).

Table 2. The known values used in the simulations of infection rate β and recovery rate δ.

Parameter	Infection rate, β	Recovery rate, δ
Population size, N	1360	1360
Susceptible, S	1359	1359
Exposed, E	1	1
Infected, I	0	0
Recovered, R	0	0
Mortality infection, v_d	$(1\text{-}0.4)^{1/12}$	$(1\text{-}0.4)^{1/12}$
Natural mortality, m_d	0	0
Infection rate, β	See Table 3	0.00035
Recovery rate, δ	0	See Table 3

Table 3. The different values tested for both infection rate β and recovery rate δ.

Parameter	Tested values			
Infection rate, β	0.00025	0.00035	0.00045	0.00055
Recovery rate, δ	0	0.025	0.050	0.075

Another observation is that these simulations are closer to the empirical data from Koster in 1988. In Figure 3 it is shown that the red line with circles where β is set to 0.00035 and δ is set to 0 is the closest of the simulations to the empirical data curve from Koster, 1988. This is why the values behind this simulation is the ones chosen for the final simulation, as is illustrated in

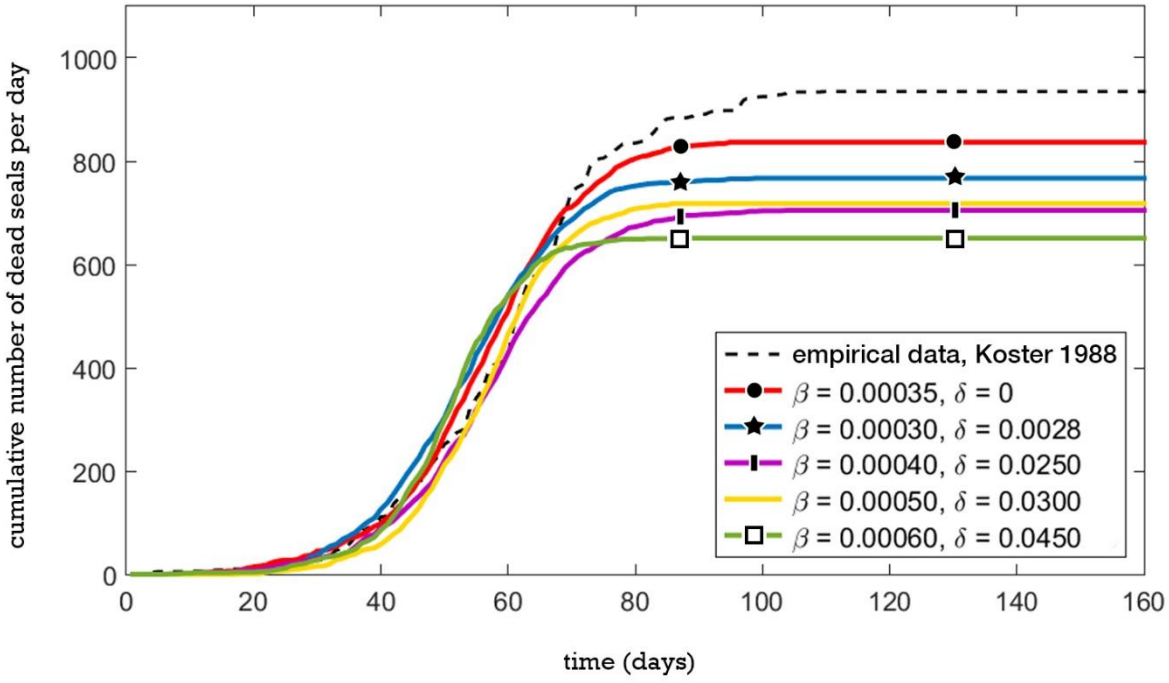


Figure 3. Illustration of different combinations of values for the infection rate, β and the recovery rate, δ that gives the simulation a similar shape as the empirical data curve from the outbreak in Koster 1988. The x-axis shows time in days and the y-axis shows cumulative number of dead individuals per day.

Figure 4.

It is also noticeable here as well as in Figure 2B that the recovery rate, δ affects the final number of total dead, since this is becoming lower and lower with increasing values for δ.

In the final simulation (see Figure 4A) it is noticeable that the simulated curve is not completely

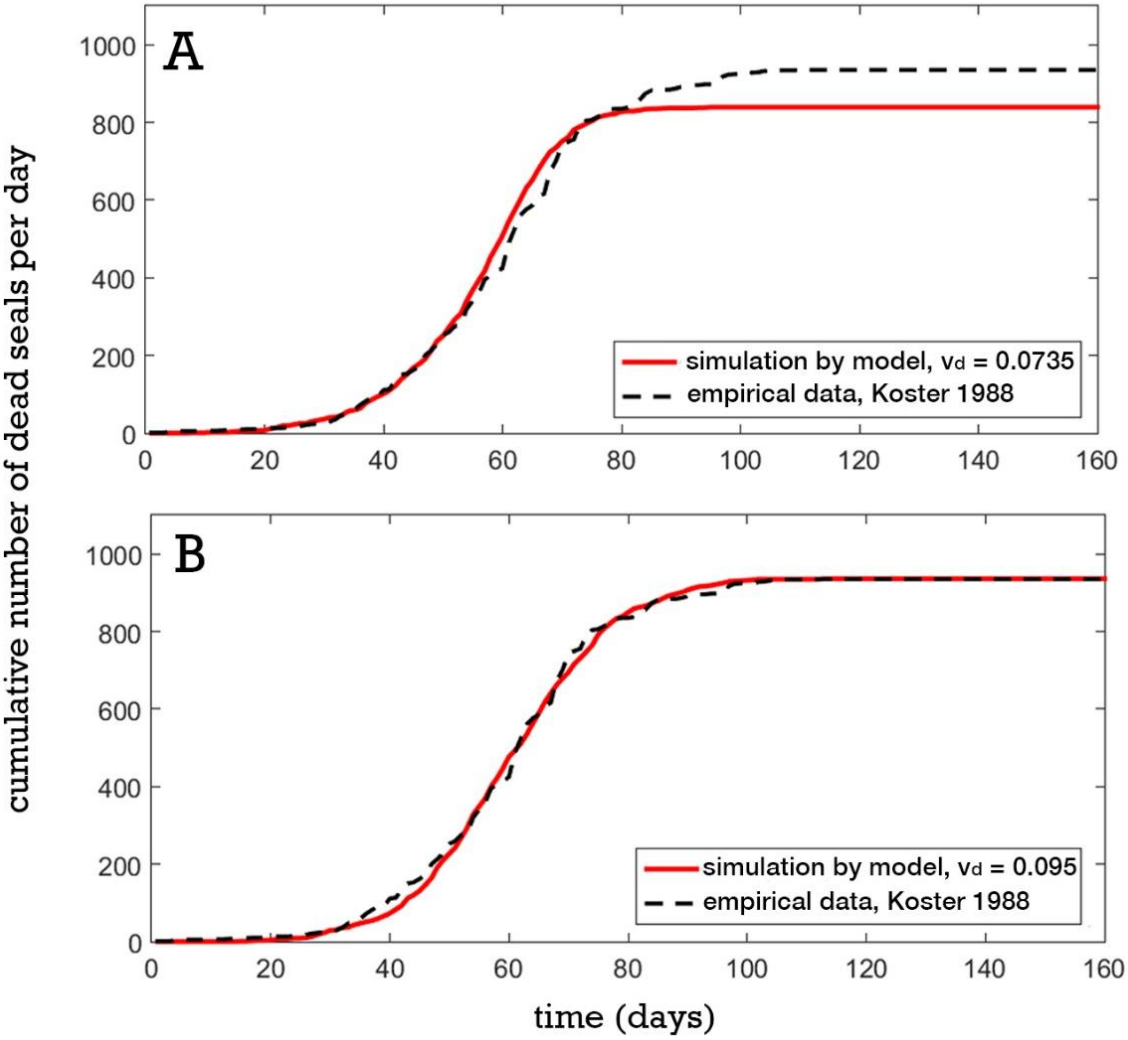


Figure 4 Illustration of the simulation with the final values, the recovery rate (δ) set to 0 and the infection rate (β) set to 0.00035. The x-axis shows time in days and the y-axis shows cumulative number of dead individuals per day. ***4A*** *shows the simulation with the value 0.0735 on the daily mortality added by the infection (v) calculated above in equation 1. In* ***4B*** *the daily mortality added by the infection (v) is changed to 0.095.*

similar to the empirical data curve. The empirical data curve has a slightly higher steady state than the one simulated based on the given parameter values (see Figure 4A). Because of this difference between them, new simulations with varying mortality from the virus *v* were made. Out of those simulations the one that best resembled the curve from Koster 1988 was a simulation with *v = 0.095* instead of the calculated *0.0735*, this is illustrated in Figure 4B.

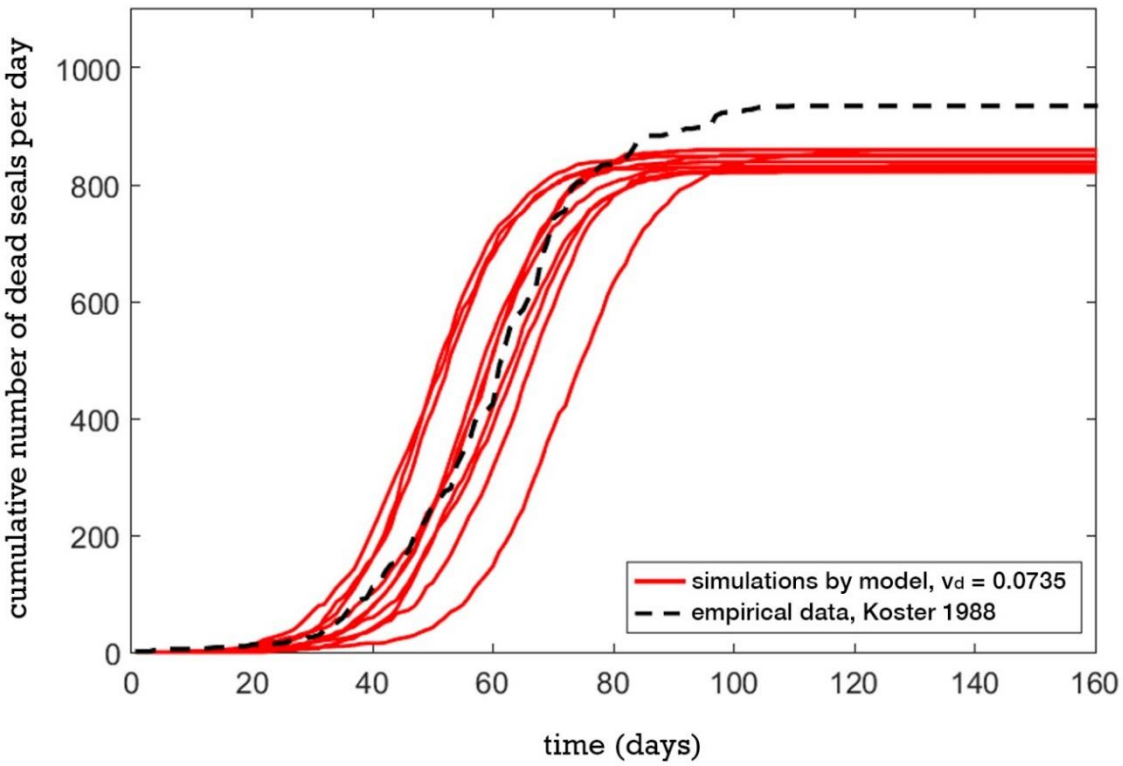


Figure 5. A series of 10 simulations by the model with the finally assumed values in comparison to the empirical data from Koster in 1988. The x-axis shows time in days and the y-axis shows cumulative number of dead individuals per day.

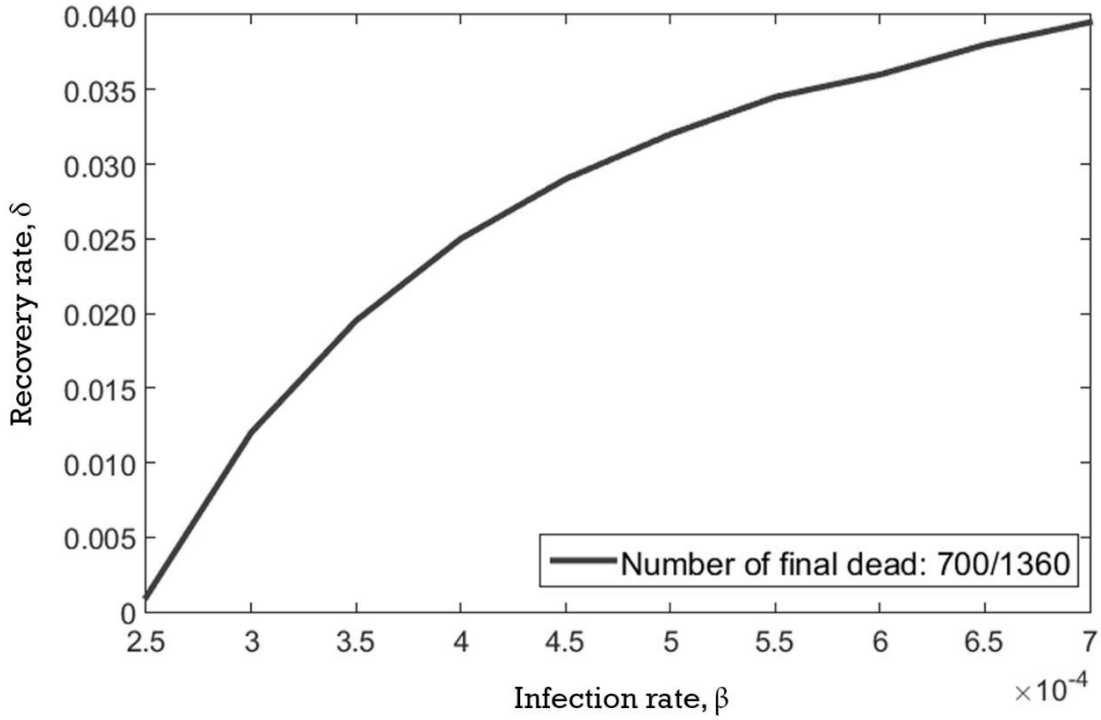


Figure 6.The relationship between the recovery rate, δ (y-axis) and the infection rate, β (x-axis) given a set number of final dead, 700 ±10 out of a population size of 1360 individuals.

In Figure 5 the variations between simulations with fixed parameter values, same values as in the simulation in Figure 4A, are illustrated. There is a slight variation between the simulations, which is expected since it is a stochastic model. However the shape of the curves are more or less the same for all cases.

Lastly in Figure 6 it is illustrated which combinations of parameter values of infection rate β and recovery rate δ that can produce the found final size of an epizootic (in this case 700 dead out of a population with the size 1360). The effect over the final number of dead comparing recovery rate with the infection rate, it is visible that the recovery rate has a slightly heavier influence compared to when the steady state occurs since the curve bends upwards towards the y-axis. If the two parameters would influence the outcome equally there would be a straight line.

Discussion

The repeated large scale epizootics in the Harbour seal populations in Northern Europe raises a number of questions. Firstly, why do the development of the epizootics differ between colonies and especially why do the final number of dead vary? What is the underlying difference between colonies? Is it that the disease hit the colony during different times within the year in combination with the seasonal differences in behaviours of the seals? Could it be because of immunosuppression caused by PCB or other environmental contaminants? The present study aims at gaining insights to this by analysing the epidemiological curves from the outbreak and thereby to get a better understanding of the processes behind the differences in regional mortality rates. This was done by making simulations testing different parameter values to investigate which parameters could potentially determine the shape of an epizootic curve, based on cumulative number of dead per day during the outbreak.

The simulation shown in Figure 4A (which illustrates the simulated curve with the estimated parameter values, the curve formed by the empirical data from the outbreak in Koster 1988 and how these two differ from each other) raises the question, what is the reason behind the difference in the final size of the number dead. In finding an answer for this the model and the estimated parameter values have to be looked into.

Natural mortality was found to be a low number ($1.7 \cdot 10^{-4}$) and because the epizootic period are short, the natural mortality was excluded as an important factor on the outcome of the cumulated death curves. This was also tested as is shown in Table 1, where the simulations with the natural mortality set to $1.7 \cdot 10^{-4}$ gave an average of 853 individuals at the steady state (time step 110) and then grew with an average of 9 over the next 100 time steps. When the simulation with natural mortality was set to 0 it resulted in an average of 851 at the steady state (time step 110). I conclude that the choice to set natural mortality to 0 does not affect the epidemiology of PDV much. However it is worth mentioning that for longer epizootics and other host parasite systems this might not be the case and natural mortality shall not be disregarded routinely.

Next, I look into the length of the latency (E) and infection period (I) which are both set as a fixed number of time steps in this model: 3 and 12 days each. Since these are derived from an average value across many individuals it implies that both these time periods can be either longer or shorter for a particular individual which add another stochastic component to the system. However, Heide-Jorgensen and Harkonen (1992) suggest that as long as the average length of E and I is set to a fixed number, in this study 15 days, any combination of the length

of E and I do not affect the over all probability of infection, however they do not say anything about how this would affect the R_0 value. Maybe this variation between how many days an individual is E or I affects the R_0 value, this can be something to bear in in mind and investigate further.
With this in mind I believe that these fixed values are not the reason behind the final dead variation. However, Swinton et al. (1998) report that this period, as E and I combined, can vary from 11-18 days, this variation can be a factor to bear in mind when developing the model further. There are no reason to think that the lengths of latency and infection periods would differ among Harbour seal colonies, the length of E and I are probably linked to the epidemiology of PDV and are specific for each host species

Even if I do not believe that the number of days in the latency or infectious states as they are formulated in the model right now affects the outcome of the final size of the simulation there are still improvements that can be done. When it comes to the latency period I think that a good way to organise it a next version of the model would be to make σ, the rate of removal, to a daily rate from E into I, but also to have a fixed number of maximum days, similar to how the infection periods are formulated in this model now. In addition to this, variation in the value of removal during the latency period. I suggest making it bigger with time since the longer an individual has been infected the more the virus should accumulate inside the individual, hence giving it a bigger risk of developing the disease.
When it comes to infection periods I think that the way it is structured here is a good way to start, it would be interesting to look into whether there can be a variation in the infection rate, β, the recovery rate, δ, and maybe even in the mortality added by the infection between the days an individual are classified as I. It is not unlikely that the infectiousness are influenced by the behaviour (contact rate) or the immune system of an individual (body condition, pollutants etc.).

When it comes to ν, the mortality from the disease for the infected, it can easily be stated that the estimated value (0.0735) can have a large influence on the difference in the final number of dead, especially when the simulation with a higher value for the virus induced mortality (0.095) results in an almost identical curve as the empirical data (Figure 4B). However, by using the same equation used when ν_d were estimated (Equation 1), this would raise the mortality by the whole epizootic up to 70% instead of the previously estimated 60% (T. Harkonen & Heide-Jorgensen, 1990; Heide-Jorgensen et al., 1992; ICES, 1990). Even if 70% is a similar number as the epizootic mortality of 2002 in Koster, 66% (T. Harkonen et al., 2006), I do not think that the variation between the simulation and the empirical data can be explained by a higher mortality caused by the infection. Although I think it is safe to say that a contributing factor to the difference between curves from different locations and epizootic outbreak have a difference in virus induced mortality. However this raises the question of why there is a difference in this mortality, what factors play a part in a colony's resistance against PDV? A future tool would be to estimate R_0 for this new SEIR model.

To look into the estimations of infection rate (0.00035) and the recovery rate (0) in the final simulation, it is strange and highly unlikely that the recovery rate would have been 0 in the real epizootic outbreak. This is however nothing I can prove and would therefor suggest to make a tool to systematically test all possible combinations of values on the different parameters to get estimations for the values that best explain the curve. With these estimations it is possible to see if these numbers seem reasonable from an ecological perspective or if there in fact can be other factors to take into consideration to get a better and more accurate model.

The model could be improved by the addition of age groups, this is because of the difference in behaviour of the different ages and sexes of seals that I mentioned in the introduction. Younger seals tend to travel longer distances and adult seals travel less (Dietz et al., 2013), they also have different contact rates on land during the lactation, breeding and moulting periods therefore giving them different probabilities to infect others or to get infected themselves seems appropriate.
The emigrational behaviour also varies with season (Cunningham et al., 2009; Dietz et al., 2013; P. M. Thompson et al., 1989) and since both the epizootics started in April-May and continued through-out the year hitting the last colony around August (Swinton et al., 1998), this can be a contributing factor to why some of the colonies that got infected later in the year also had different curves with a lower slope and a smaller number of dead at the end of the epizootic.
Lastly I want to mention the environmental contaminants as a possible factor to consider as well since seals with higher levels of bio-accumulated environmental contaminants have a reduced immune response (de Swart et al., 1996; Ross et al., 1995). Levels of PCB in seal tissue at the time of the outbreak was higher in Kattegat, Skagerrak compared to the Scottish colonies for example (Aguilar, Borrell, & Reijnders, 2002).

Conclusion

One conclusion form this study is that out of the two parameters infection rate and recovery rate, recovery rate is the one with more influence over the final number of dead individuals.
The Model developed here seems to be a good basic model that easily could be developed further. The first thing that should be done is to make a system to estimate the parameter values given a certain curve shape.
Also I strongly suggest that an age structure in the model would be preferred, secondly seasonal differences in contact rate and thirdly, environmental contaminants (affecting virulence) can be considered when comparing different locations.
Lastly looking into differences in parameter values between days during the infection period when it comes to infection rate, recovery rate and maybe even infection induced mortality can provide valuable insights into PDV epidemiology.

Acknowledgements

Thanks to Álvaro Pérez who helped me to get the MatLab code to finally work and to Albert Monteith for correctional reading.

References

Aguilar, A., Borrell, A., & Reijnders, P. J. H. (2002). Geographical and temporal variation in levels of organochlorine contaminants in marine mammals. *Marine Environmental Research, 53*(5), 425-452. doi: 10.1016/s0141-1136(01)00128-3

Andersson, H., & Brittin, T. (2000). *Stochastic Epidemic Models and Their Statistical Analysis*. New York: Springer-Verlag.

Begon, M., Bennett, M., Bowers, R. G., French, N. P., Hazel, S. M., & Turner, J. (2002). A clarification of transmission terms in host-microparasite models: numbers, densities and areas. *Epidemiology and Infection, 129*(1), 147-153. doi: 10.1017/s0950268802007148

Cosby, S. L., McQuaid, S., Duffy, N., Lyons, C., Rima, B. K., Allan, G. M., . . . Orveli, C. (1988). Characterization of a seal morbillivirus. *Nature, 336*(6195), 115-116. doi: 10.1038/336115b0

Cunningham, L., Baxter, J. M., Boyd, I. L., Duck, C. D., Lonergan, M., Moss, S. E., & McConnell, B. (2009). Harbour seal movements and haul-out patterns: implications for monitoring and management. *Aquatic Conservation-Marine and Freshwater Ecosystems, 19*(4), 398-407. doi: 10.1002/aqc.983

de Koeijer, A., Diekmann, O., & Reijnders, P. (1998). Modelling the spread of phocine distemper virus among harbour seals. *Bulletin of Mathematical Biology, 60*(3), 585-596. doi: 10.1006/bulm.1997.0030

de Swart, R. L., Ross, P. S., Vos, J. G., & Osterhaus, A. (1996). Impaired immunity in harbour seals (Phoca vitulina) exposed to bioaccumulated environmental contaminants: Review of a long-term feeding study. *Environmental Health Perspectives, 104*, 823-828. doi: 10.2307/3432713

Dietz, R., Heide-Jorgensen, M. P., & Harkonen, T. (1989). Mass Deaths of Harbor seals (Phoca vitulina) in Europe. *Ambio, 18*(5), 258-264.

Dietz, R., Teilmann, J., Andersen, S. M., Riget, F., & Olsen, M. T. (2013). Movements and site fidelity of harbour seals (Phoca vitulina) in Kattegat, Denmark, with implications for the epidemiology of the phocine distemper virus. *ICES Journal of Marine Science, 70*(1), 186-195. doi: 10.1093/icesjms/fss144

Goodman, S. J. (1998). Patterns of extensive genetic differentiation and variation among European harbor seals (Phoca vitulina vitulina) revealed using microsatellite DNA polymorphisms. *Molecular Biology and Evolution, 15*(2), 104-118.

Harding, K. C., Harkonen, T., & Caswell, H. (2002). The 2002 European seal plague: epidemiology and population consequences. *Ecology Letters, 5*(6), 727-732.

Harkonen, T., Dietz, R., Reijnders, P., Teilmann, J., Harding, K., Hall, A., . . . Thompson, P. (2006). A review of the 1988 and 2002 phocine distemper virus epidemics in European harbour seals. *Diseases of aquatic organisms, 68*(1), 115-130.

Harkonen, T., Harding, K., Rasmussen, T. D., Teilmann, J., & Dietz, R. (2007). Age- and Sex-Specific Mortality Patterns in an Emerging Wildlife Epidemic: The Phocine Distemper in European Harbour Seals. *PLoS One, 2*(9). doi: 10.1371/journal.pone.0000887

Harkonen, T., Harding, K. C., & Heide-Jorgensen, M. P. (2002). Rates of increase in age-structured populations: a lesson from the European harbour seals. *Canadian Journal of Zoology-Revue Canadienne De Zoologie, 80*(9), 1498-1510. doi: 10.1139/z02-141

Harkonen, T., & Heide-Jorgensen, M. P. (1990). Short-term effects of the mass dying of Harbor seals in the Kattegat-Skagerrak area during 1988. *Zeitschrift Fur Saugetierkunde-International Journal of Mammalian Biology, 55*(4), 233-238.

Heide-Jorgensen, M. P., & Harkonen, T. (1992). Epizootiology of the Seal Disease in the Eastern North Sea. *Journal of Applied Ecology, 29*(1), 99-107. doi: 10.2307/2404353

Heide-Jorgensen, M. P., Harkonen, T., Dietz, R., & Thompson, P. M. (1992). Retrospective of the 1988 European seal epizootic. *Diseases of aquatic organisms, 13*(1), 37-62. doi: 10.3354/dao013037

ICES. (1990). Report of the joint meeting of the working group on Baltic seals and the study group on the effects of contaminants on marine mammals. CM Documents - ICES, 1990(N:14): Stockholm, 14-17 May 1990.

Kermack, W. O., & McKendrick, A. G. (1991). Contributions to the mathematical theory of epidemics 1. (Reprinted from proceedings of the royal society, Vol 115A, p. 700-721, 1927). *Bulletin of Mathematical Biology, 53*(1-2), 33-55. doi: 10.1007/bf02464423

Klepac, P., Pomeroy, L. W., Bjornstad, O. N., Kuiken, T., Osterhaus, A. D. M. E., & Rijks, J. M. (2009). Stage-structured transmission of phocine distemper virus in the Dutch 2002 outbreak. *Proceedings of the Royal Society B-Biological Sciences, 276*(1666), 2469-2476. doi: 10.1098/rspb.2009.0175

Mahy, B. W. J., Barrett, T., Evans, S., Anderson, E. C., & Bostock, C. J. (1988). Characterization of a seal morbillivirus. *Nature, 336*(6195), 115-115. doi: 10.1038/336115a0

Murray, A. G. (2009). Using simple models to review the application and implications of different approaches used to simulate transmission of pathogens among aquatic animals. *Preventive Veterinary Medicine, 88*(3), 167-177. doi: 10.1016/j.prevetmed.2008.09.006

Osterhaus, A., & Vedder, E. J. (1988). Identification of virus causing recent seal deaths. *Nature, 335*(6185), 20-20. doi: 10.1038/335020a0

Ross, P. S., Deswart, R. L., Reijnders, P. J. H., Vanloveren, H., Vos, J. G., & Osterhaus, A. (1995). Contaminant-Related Suppression of Delayed-Type Hypersensitivity and Antibody Responses in Harbor Seals Fed Herring from the Baltic Sea. *Environmental Health Perspectives, 103*(2), 162-167. doi: 10.2307/3432272

Swinton, J., Harwood, J., Grenfell, B. T., & Gilligan, C. A. (1998). Persistence thresholds for phocine distemper virus infection in harbour seal Phoca vitulina metapopulations. *Journal of Animal Ecology, 67*(1), 54-68.

Thompson, D., Lonergan, M., & Duck, C. (2005). Population dynamics of harbour seals Phoca vitulina in England: monitoring growth and catastrophic declines. *Journal of Applied Ecology, 42*(4), 638-648. doi: 10.1111/j.1365-2664.2005.01025.x

Thompson, P. M., Fedak, M. A., McConnell, B. J., & Nicholas, K. S. (1989). Seasonal and Sex-Related Variation in the Activity Patterns of Common Seals (Phoca vitulina). *Journal of Applied Ecology, 26*(2), 521-535. doi: 10.2307/2404078

Thompson, P. M., & Miller, D. (1992). Phocine distemper virus outbreak in the Moray Firth common seal population: an estimate of mortality. *Science of the Total Environment, 115*(1-2), 57-65. doi: 10.1016/0048-9697(92)90032-n