Research

Author for correspondence:
Sinead E. Morris
e-mail: semorris@princeton.edu

†These authors contributed equally to this study.

Partially observed epidemics in wildlife hosts: modelling an outbreak of dolphin morbillivirus in the northwestern Atlantic, June 2013–2014

Sinead E. Morris[1,†], Jonathan L. Zelner[2,†], Deborah A. Fauquier[3], Teresa K. Rowles[3], Patricia E. Rosel[4], Frances Gulland[5,6] and Bryan T. Grenfell[1,7]

[1]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA
[2]Robert Wood Johnson Health and Society Scholars Program, Columbia University, New York, NY, USA
[3]National Marine Fisheries Service, Marine Mammal Health and Stranding Response Program, Silver Spring, MD, USA
[4]National Marine Fisheries Service, Southeast Fisheries Science Center, Lafayette, LA, USA
[5]The Marine Mammal Centre, Sausalito, CA, USA
[6]US Marine Mammal Commission, 4340 East West Highway, Bethesda, MD, USA
[7]Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

Morbilliviruses cause major mortality in marine mammals, but the dynamics of transmission and persistence are ill understood compared to terrestrial counterparts such as measles; this is especially true for epidemics in cetaceans. However, the recent outbreak of dolphin morbillivirus in the northwestern Atlantic Ocean can provide new insights into the epidemiology and spatio-temporal spread of this pathogen. To deal with uncertainties surrounding the ecology of this system (only stranded animals were observed), we develop a statistical framework that can extract key information about the underlying transmission process given only sparse data. Our self-exciting Poisson process model suggests that individuals are infectious for at most 24 days and can transfer infection up to two latitude degrees (220 km) within this time. In addition, the effective reproduction number is generally below one, but reaches 2.6 during a period of heightened stranding numbers near Virginia Beach, Virginia, in summer 2013. Network analysis suggests local movements dominate spatial spread, with seasonal migration facilitating wider dissemination along the coast. Finally, a low virus transmission rate or high levels of pre-existing immunity can explain the lack of viral spread into the Gulf of Mexico. More generally, our approach illustrates novel methodologies for analysing very indirectly observed epidemics.

1. Introduction

In July 2013, an outbreak of dolphin morbillivirus (DMV) was officially declared along the northwestern (NW) Atlantic coast of the United States and has since been implicated in the stranding of over 1600 common bottlenose dolphins (*Tursiops truncatus*, hereafter referred to as bottlenose dolphin). DMV, a member of the *Morbillivirus* genus that also includes measles, rinderpest, phocine distemper virus (PDV) and canine distemper virus, infects at least 14 odontocetes species worldwide and is a major cause of morbidity and mortality [1–5]. The current outbreak has been declared an unusual mortality event (UME) due to the high number and unexpected nature of the strandings that have occurred [6]. In fact, this outbreak represents the largest number of bottlenose dolphin strandings reported in the NW Atlantic since the last recorded DMV epidemic in

1987–1988, when between 10 and 50% of the coastal population was believed to have died [7–9].

The underlying dynamics and persistence of morbilliviruses in marine mammals are poorly understood, especially for cetaceans. A variety of epidemiological dynamics have been observed in DMV hosts, from endemic transmission among pilot whales (*Globicephala* spp.) in the southwestern Pacific Ocean to recurrent epidemics among dolphins in the Mediterranean Sea, Atlantic Ocean and Gulf of Mexico (GoM) [2], but the mechanisms underlying these different patterns remain elusive. Key questions regarding the epidemiology of the virus in different cetacean species, and transmission of the virus between species, have yet to be answered. However, the current NW Atlantic UME provides a unique opportunity to explore the spatio-temporal spread of DMV within a species, and thus gain new insights into the dynamics of this poorly understood pathogen.

Mathematical and statistical modelling can help gain insight into the important drivers of disease dynamics in marine systems, but one limiting factor is the availability of detailed population and epidemiological data. For example, successful modelling of PDV outbreaks in North Sea harbour seals (*Phoca vitulina*) enabled estimation of key epidemiological parameters such as the basic reproduction number, R_0 (the average number of secondary cases caused by an infected individual in a completely susceptible population) [10–13]. However, such work was aided by the 'haul-out' behaviour of harbour seals: individuals spend a significant amount of time aggregating on land, which enables accurate counts of population densities and individual mortalities to be made [12]. In contrast, much less is known about the coastal bottlenose dolphin populations of the NW Atlantic and no modelling work has been done that explicitly explores the drivers of DMV dynamics in this system. Population estimates are available, but some of the stocks are highly mobile and there is limited information regarding their spatio-temporal distribution and the extent of mixing within and between different stocks.

The natural history of DMV in cetaceans is also a major source of uncertainty. As with other morbilliviruses, immunity following recovery from infection is likely to be long-term [5,14], but other key epidemiological parameters such as R_0, the virulence of the virus, and the duration of infectiousness are largely undocumented. Transmission is mainly thought to occur when susceptible individuals inhale aerosolized virus particles from infected individuals, possibly when close-knit groups form for travelling and feeding [15]. However, vertical transmission and direct transmission through bodily fluids may also be possible [15]. Other important quantities for DMV spread that have still to be determined include the distance an individual is likely to travel while infectious, and the spatio-temporal distribution of the susceptible population.

An additional question regarding the transmission dynamics is whether the risk of infection scales with host density (often referred to as density-dependent transmission), or remains independent of host density (frequency-dependent transmission) [16]. Density-dependent transmission may be favoured by the respiratory means of infection coupled with the gregarious behaviour of dolphins [5,15]. However, if the average number of contacts of an individual remains relatively constant in the face of host density change, then frequency-dependent transmission may be more appropriate [10,16]. Bottlenose populations inhabiting estuarine waters in the NW Atlantic are thought to maintain stable group sizes which may favour frequency-dependent transmission [17], but the degree to which this accurately captures transmission dynamics in the coastal populations is unclear.

The above uncertainties present challenges when designing an appropriate epidemic model, and so we construct a minimal stochastic framework based on a self-exciting Poisson process that requires only limited information to describe the spatio-temporal dynamics of the system. Such models are more commonly used to capture the spatio-temporal clustering of earthquakes [18] and have also been employed to model urban crime distribution [19]. However, there is great potential for application within an infectious disease context as the local nature of transmission naturally leads to clustering of cases in space and time [19–21]. We demonstrate how this method can be used to infer epidemiological parameters such as the range of host movement, the duration of infectiousness and the effective reproduction number, R (the average number of secondary cases caused by an infected individual in a population for which there may be pre-existing immunity). Furthermore, since the spatio-temporal pattern of strandings is likely influenced by the migrating habits of the different coastal stocks and the intensity of transmission within and between these stocks, we reconstruct transmission networks that allow us to explore the importance of host mobility in facilitating spread along the coast. Finally, we simulate an epidemic in the GoM to explore why the current UME has seemingly failed to trigger DMV-related mortality in this neighbouring region [22]. Overall, our findings provide a first step towards better understanding *Morbillivirus* spread in the NW Atlantic.

2. Material and methods

2.1. Data

Level A stranding records from 1989 to 2014 were extracted from the National Oceanic and Atmospheric Association Marine Mammal Health and Stranding Response Program (NOAA MMHSRP) National Database. For each reported stranding, the authorized responding agency is required to complete a Marine Mammal Stranding Report—Level A data form (available at www.nmfs.noaa.gov/pr/pdfs/health/levela.pdf) that includes details such as the species, date and location of each stranding. The final data used in this paper were extracted from the database on 23 May 2014 and 30 June 2014, and included strandings up through June 2014. Due to the ongoing nature of the UME at the time of writing, subsequent records may have been added that are not included in our data. Before 1996, annual observed stranding numbers were low relative to subsequent years, reflecting the absence of recorded data south of North Carolina and high variability in reporting along the remaining coastline (electronic supplementary material, figure S1). We therefore excluded these years from further analysis to minimize historical reporting bias. The subsequent stranding data (aggregated by day and latitude band as in all analyses) are available online (Github repository: http://github.com/SineadMorris/Dolphin-morbillivirus).

During the current UME, strandings have occurred along the East Coast of the United States, from New York to Florida (27–42° N latitude; electronic supplementary material, figure S3). Individuals are thought to be from the main NW Atlantic coastal bottlenose dolphin stocks: the Northern Migratory Coastal Stock (NMCS), the Southern Migratory Coastal Stock (SMCS), the South Carolina–Georgia Coastal Stock (SCGCS) and the Northern Florida Coastal Stock (NFCS), with population estimates of 11 548, 9173, 4377 and 1219 individuals, respectively [23]. However, there are other resident bay and estuarine stocks in this region,
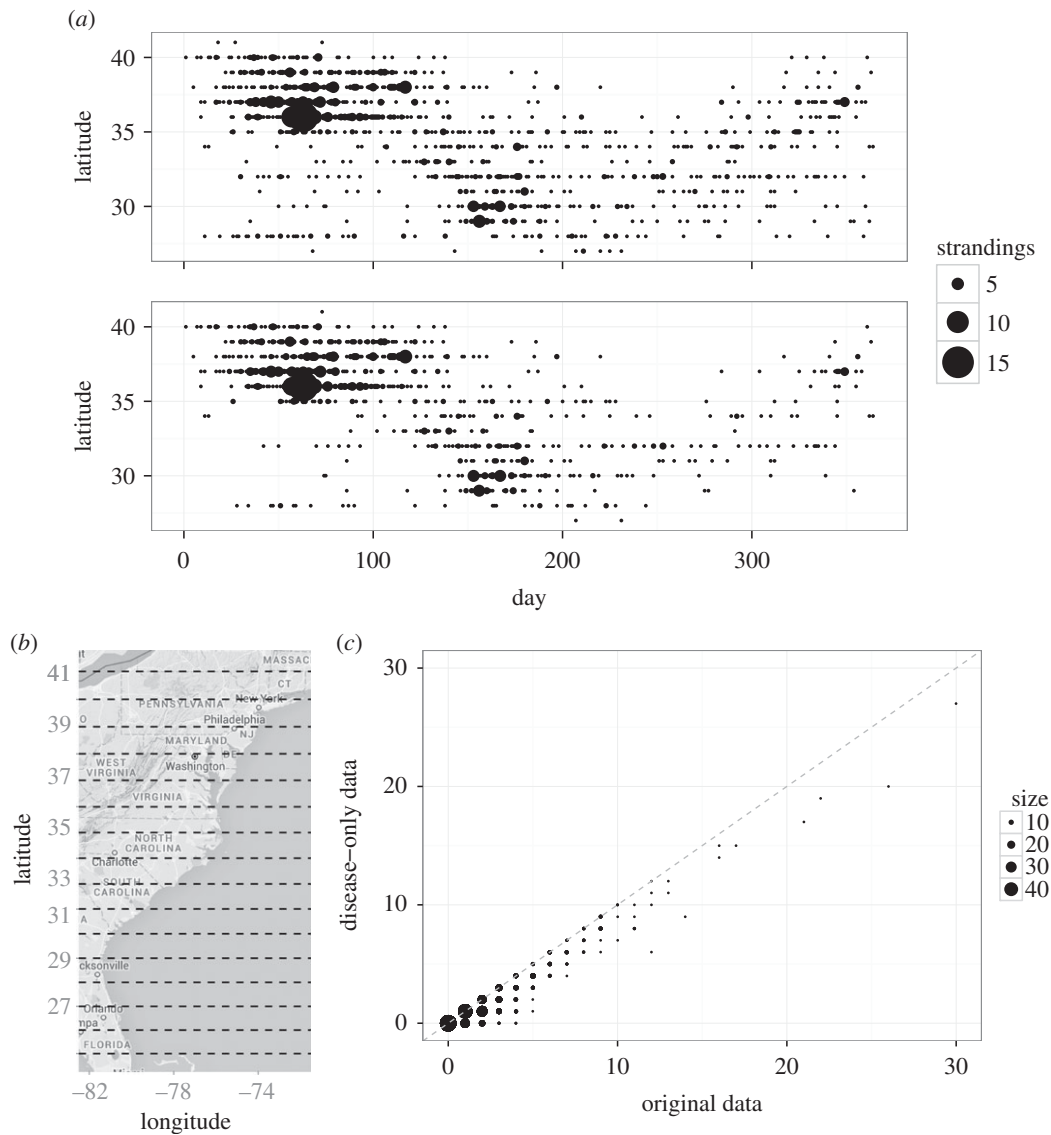
Figure 1. (a) Strandings by latitude and day since the start of the epidemic (from June 2013 to 2014) for the original UME data (top panel) and the new epidemic dataset generated by removing background cases (bottom panel). The size of each point indicates the number of strandings. (b) Map of the United States NW Atlantic coast with latitude bands highlighted by dashed lines (for reference). (c) Comparison of the new epidemic (disease-only data) and original UME datasets. Each point compares the number of strandings on a given day and latitude band between both datasets, with size indicating the number of times that particular coordinate combination occurred. The dashed line indicates when both are equal.

and the true origin of stranded cases has still to be confirmed by genetic analyses [23].

One caveat to these data is that the total number of infected individuals is likely to be underestimated due to inherent difficulties associated with monitoring disease outbreaks among marine mammals: generally only those animals that strand onshore can be reliably identified as infected cases and a larger proportion of animals may die at sea and not be counted in the stranding records. As such, we refer to this as a partially observed epidemic process and apply a minimal stochastic framework (outlined in §2.3) that enables the spatio-temporal dynamics of the system to be modelled despite the limited available data.

2.2. Data preparation

There was no recorded outbreak of morbillivirus infection during 1996–2012 and so these data reflect strandings that would happen in any given year, independent of DMV-related mortality. We refer to these as 'background' strandings and assume that the current UME data include background cases in addition to cases due to DMV. In order to implement our epidemic model, these background cases must first be removed from the UME dataset. To do this, we applied a Poisson generalized linear model to the

1996–2012 data to estimate the average annual rate of background strandings (by season and latitude degree) in non-epidemic years, and then removed the corresponding proportion of strandings from the UME dataset (see electronic supplementary material, S2.1–S2.2, for further details). This procedure preserves the overall spatio-temporal pattern of strandings when compared with the original dataset, but contains fewer sporadic cases and shows a clearer shift in the main stranding density as time progresses (figure 1a). We assume that potential biases in the UME data arising from non-uniform temporal and/or spatial distributions of background strandings (e.g. due to seasonal changes in the spatial distribution of host abundance) are accounted for by removing these from the data. We thus conclude that the resulting dataset sufficiently approximates the distribution of disease-induced strandings only, and it is henceforth referred to as the 'epidemic' dataset.

Implementation of our model also requires estimation of the spatio-temporal distribution of the NW Atlantic coastal bottle-nose dolphin population. The two largest coastal stocks are believed to be migratory: the NMCS occupies waters as far north as New York (approx. 41° N) in summer and as far south as central North Carolina (approx. 34.5° N) in winter, and the SMCS ranges from Virginia (approx. 38° N) in summer to

northern Florida (approx. $28°$ N) in winter [23–25]. However, there is large uncertainty in the precise boundaries of these ranges and how seasonal migration of these stocks affects the spatio-temporal distribution of the entire coastal population. For instance, the smaller coastal stocks (SCGCS and NFCS) are thought to be resident, i.e. non-migratory, but may come into contact with the SMCS at certain times during the year.

Previous work has made progress in this area by using stranding distributions in non-epidemic years to inform hypotheses about the structure and migration patterns of the coastal populations [25]. Guided by this approach, we assumed that the number of background strandings at any location was proportional to the local population density at the time of the strandings. This enabled us to estimate the spatio-temporal distribution of population density along the coast in a typical, non-epidemic year using the seasonal distribution of background strandings and estimates of the total coastal population size (see electronic supplementary material, S2.3, and figures S2 and S5, for further details). Our estimates are in broad agreement with the seasonal migratory ranges described above and are thus used in the epidemic model.

2.3. Epidemic model

We assume that the stranding locations along the coast are uniquely defined by their respective latitude coordinates, so that one spatial dimension, l, is sufficient to model the epidemic process. In addition, we assume that both the duration of time and the distance travelled from each individual's initial infection (unobserved data) to their eventual stranding (observed data) remain constant throughout our dataset (e.g. any pair of individuals that strand 2 days apart are assumed to have been infected 2 days apart). This enables us to approximate the spatio-temporal distribution of cases in the unobserved infection process using the distribution of strandings in our epidemic dataset.

If the full temporal and spatial ranges covered in our epidemic data are defined by $t \in [0, T]$ and $l \in [L_1, L_2]$, respectively, then the distribution of cases, $S_{t,l}$, is modelled as an inhomogeneous (or 'non-stationary') Poisson process [26], for which the probability of observing N total cases is given by

$$\Pr(N) = \frac{\Lambda^N e^{-\Lambda}}{N!}, \quad \text{where } \Lambda = \int_0^T \int_{L_1}^{L_2} \lambda(t, l) \mathrm{d}t \mathrm{d}l.$$

The hazard function (or conditional intensity) at time t and latitude l, $\lambda(t, l)$, represents the rate at which new cases arise at each new point in the epidemic and is analogous to the force of infection in standard epidemiological models [19].

The history of all cases i, in the epidemic process up to time t, is defined by the set of all observation times $\{t_i\}$ that occur before t and the corresponding latitude, l_i, at which the observation was recorded, i.e. $H(t) = \{(t_i, l_i) | t_i < t\}$ [18]. Given this history, λ can be expressed as

$$\lambda(t, l|H(t)) = R \lim_{\Delta_t, \Delta_l \to 0} \left[\frac{P_{\Delta_t, \Delta_l}(t, l|H(t))}{\Delta_t \Delta_l} \right],$$

where $P_{\Delta_t, \Delta_l}(t, l|H(t))$ is the combined probability density that each previous case generates a new case in the time interval $(t, t + \Delta_t)$ and latitude region $(l, l + \Delta_l)$ (determined by the current infectiousness of each case), and R is the effective reproduction number (that may also vary as a function of time and latitude) [18,19,26]. Note that R represents the average number of new cases generated by each case over the entire epidemic, and so multiplying the probability density by R determines how these cases are distributed across space and time. Such frameworks, termed 'self-exciting' Poisson processes due to the influence of previous events on the distribution of future events [27], are often used to model processes exhibiting significant spatio-temporal clustering [19]. We use this framework to model the spatio-temporal distribution of stranding

events in the NW Atlantic system within an epidemiological context, but other applications include modelling of earthquakes and urban crime [18,19].

Similar to previous formulations proposed for earthquake and epidemic modelling [18,28], we define probability distributions, $g(t)$ and $f(l)$, to describe how an individual's infectiousness decays in time and space, respectively. The hazard can then be expressed as

$$\lambda(t, l|H(t)) = R \sum_{i, t_i < t} g(t - t_i) f(l - l_i). \tag{2.1}$$

Since t_i and l_i are the initial observation time and location for individual i, respectively, it follows that $(t - t_i)$ is the current length of time (in days) since the beginning of that individual's infectious period and $(l - l_i)$ is the current distance (in latitude degrees) from the individual's initial location [18,19]. Thus, $g(t - t_i)f(l - l_i)$ represents the current proportional contribution of the individual's overall infectiousness to the hazard at time t and location l, and the total hazard is obtained by taking the sum of the current contributions of all previously infected individuals [19,20].

Many functional forms describing the change in infectiousness are possible, but here we assume an individual's infectiousness decays exponentially in space and time from the initial infection event [20,28]. Specifically, we model the temporal decay of infectiousness using an exponential distribution,

$$g(t - t_i) = \alpha \exp(-\alpha(t - t_i)),$$

and the spatial decay with a normal distribution,

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

The parameters α and σ control the rate of decay of infectiousness in space and time, respectively, and are estimated during the fitting process. Although their biological interpretation is limited by the sparseness of our data, i.e. they cannot incorporate the influence of unobserved cases, they are valuable in that they allow us to characterize the dynamics of the observed infection process within an epidemiological context.

The effective reproduction number is expressed as $R = \beta D_{t,l}^\psi$, where β is the pathogen transmission rate, $D_{t,l}$ is the population size at time t and latitude l, and $\psi \in \{0, 1\}$ is a scaling parameter that determines whether transmission is frequency- or density-dependent. When transmission is frequency-dependent, $\psi = 0$ and R is constant throughout the epidemic, whereas when transmission is density-dependent, $\psi = 1$ and R changes in proportion to the local population distribution [16]. We investigate both scenarios in our analyses to determine which most accurately captures the underlying transmission process. We use R instead of R_0 since we do not have sufficient evidence to assume the entire population is susceptible. Instead, we allow for the possibility that some individuals may be immune, and note that the formulation for the density-dependent model may be viewed as an upper bound for R as it incorporates the total population size rather than the unknown susceptible proportion.

This framework can also be adapted to capture periods where the transmission intensity may significantly fluctuate above or below the baseline value, R. For example, an additional reproduction number, $\hat{R} = \hat{\beta} D_{t,l}^\psi$, can be defined for $t \in [t_1, t_2]$ and $l \in [l_1, l_2]$ such that

$$\lambda(t, l) = \begin{cases} \hat{R} \sum_{i, t_i < t} g(t - t_i) f(l - l_i) & \text{if } l_1 \le l \le l_2 \text{ and } t_1 \le t \le t_2 \\ R \sum_{i, t_i < t} g(t - t_i) f(l - l_i) & \text{otherwise}. \end{cases}$$

One final point is that the self-exciting Poisson framework described here does not account for the dynamic depletion of susceptible individuals over space and time as the epidemic

Table 1. Model parameters.

parameter	description
β	transmission rate
D	population density
R	average number of secondary infections caused by one individual
σ	rate of spatial decay of infectiousness
α	rate of temporal decay of infectiousness
ψ	transmission scaling factor (0 for frequency- and 1 for density-dependence)
λ	hazard function

progresses. The number of strandings on each day is small relative to the total population size and thus it is assumed that susceptible depletion will not play a significant role in the transmission dynamics. However, we also develop an analogous binomial chain model that accounts for susceptible depletion to test this assumption (see electronic supplementary material, S6, for further details).

2.3.1. Parameter estimation

All parameters incorporated in the model are defined in table 1, and the log-likelihood of observing the entire time series is [19,26]

$$L = \left[\sum_{i=1}^{N} \log(\lambda(t_i, l_i)) \right] - \int_0^T \int_{L_1}^{L_2} \lambda(t, l)\mathrm{d}t\mathrm{d}l$$

(see electronic supplementary material, S3, for a derivation of this formula). To estimate the unknown parameters, α, β and σ, Bayesian inference was conducted via Hamiltonian Markov chain Monte Carlo (MCMC) sampling using uniform prior distributions in the R package RStan [29,30]. The effective reproduction number was calculated using estimates from the inferred parameter posterior distributions and model predictions were compared with the observed data via self-exciting Poisson simulations of the form: $S_{t,l} = \mathrm{Pois}(\lambda_{t,l})$.

Model fits were also compared using approximate Akaike information criterion (AIC) and Watanabe−Akaike information criterion (WAIC) methods, which compare the predictive accuracy of different models while penalizing those with more parameters [31–33]. Since AIC calculations rely on a point value for the maximum log-likelihood (that must be estimated from our posterior distributions), they are not entirely suited to a Bayesian framework and so we include these only for approximate comparisons. In contrast, WAIC calculations average across the posterior distribution rather than requiring a point estimate to be made and so may be more appropriate in this setting [32,33]. We thus use WAIC in conjunction with AIC to compare our different models. Further details of these information criterion calculations are given in electronic supplementary material, S4.

2.4. Transmission networks

Using the distribution of parameter estimates from the best-fitting model, transmission networks were reconstructed that connected stranded individuals with the source of their stranding (i.e. the previously stranded individual that most likely caused their infection). Briefly, for each *Tursiops* that stranded at time t_i and latitude l_i, we compared the contributions of all previously stranded individuals to the hazard at t_i and l_i and used the resulting probability distribution to determine the

most likely source of infection (see electronic supplementary material, S7, for further details). Directed networks were then created by grouping individuals according to the latitude where they stranded. For example, a connection from node 36° N to node 38° N in the network would represent a transmission event originating from an individual that stranded at 36° N to an individual that consequently stranded at 38° N.

A different transmission network was generated for each value in a randomly sampled subset of the estimated parameter distributions, and these transmission networks then served as a tool for visualizing and interpreting the results of the model inference. In particular, network properties were evaluated using betweenness measures with the R package igraph [34]. The betweenness of a node j is the number of shortest paths between any pair of network nodes that pass through j, and thus indicates the influence or centrality of j within the network [35]. This measure enabled quantification of the temporal change in the spatial structure of the inferred transmission networks over the course of the epidemic.

2.5. Simulating an outbreak in the Gulf of Mexico

To investigate the apparent lack of viral spread to bottlenose dolphin populations in the GoM, we simulated an outbreak in this region by introducing one infected case into an otherwise susceptible population. Cases were generated according to the self-exciting Poisson process defined above, with the constraint that the cumulative number of cases could not exceed the total estimated population size. We considered only the stock that would most likely be expected to come in first contact with the currently affected NW Atlantic stocks, namely the GoM Eastern Coastal Stock that inhabits waters along the west coast of Florida from 25 to 31° N. In addition, we assumed that the total population size (7702) was equally distributed along this stretch of coastline throughout the simulation [36]. Median parameter values from the posterior distributions of the best-fitting model were used as simulation inputs.

3. Results

To assess whether frequency- ($\psi = 0$) or density-dependent ($\psi = 1$) transmission best characterized the underlying epidemic dynamics, both models were fitted to the data using MCMC methods as described above. There is some variation between the parameter estimates (primarily in the estimates for R), but the main distinction is that the frequency-dependent model returns lower AIC and WAIC values than the density-dependent model and thus represents a better fit to the entire dataset (table 2). However, the strongest epidemic signal in the data occurs during a significant peak in stranding numbers in August 2013, between southeast Virginia and northeast North Carolina (36–37° N, figure 1*a,b*), that is better captured by the density-dependent model (electronic supplementary material, figures S6 and S7). Thus, although frequency-dependent transmission appears to dominate in general, it is likely that there are other factors contributing to an increased transmission intensity during this peak period that the model does not account for.

In order to better capture the key dynamics of the epidemic, we defined a new model that incorporated the frequency-dependent nature of transmission while also allowing the significant peak in August 2013 to be captured. Given the isolated nature of this signal, in space and in time, we adapted the frequency-dependent model to incorporate an additional transmission rate around the occurrence of
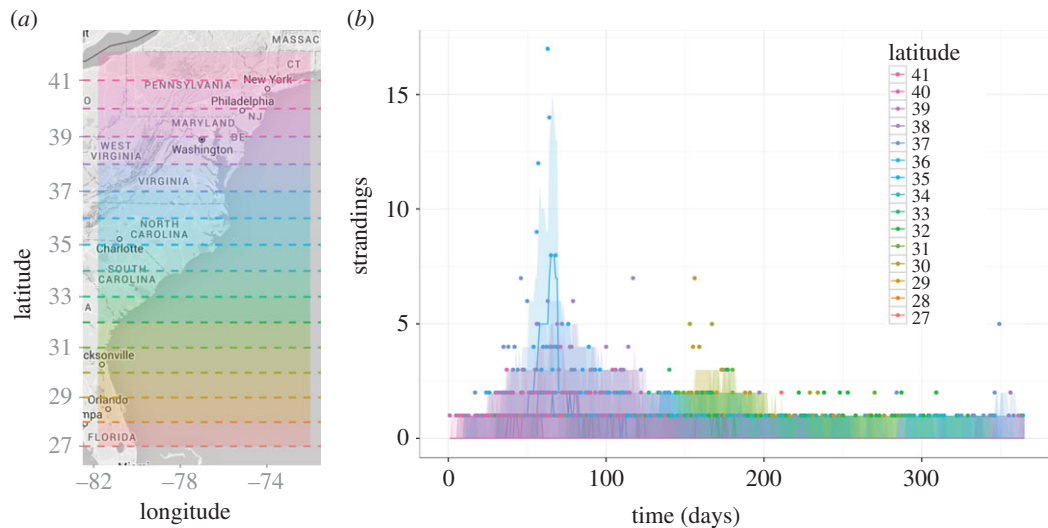
Figure 2. (a) Map of the United States NW Atlantic coast with latitude bands highlighted by dashed lines and shaded regions (for reference). (b) Simulated predictions of the frequency-dependent model with location-specific transmission rate. Cases were simulated across latitudes, l, and time, t, as $Pois(\lambda_{l,t})$ using the hazard function calculated during parameter estimation. Lines represent median values from 2000 simulations in RStan, shaded regions represent the 2.5th–97.5th quantile range and points represent actual data. In both panels, latitude bands are highlighted by the same colour.

Table 2. Model comparisons. Values indicate median parameter estimates (2.5th–97.5th quantiles) and information criterion from 2000 RStan simulations (not including warm-up sampling) of each model.

model	density-dependent	frequency-dependent	frequency-dependent with location-specific transmission rate
R	min: 0.045 (0.043–0.048)[a]	1.01 (0.95–1.08)	0.95 (0.89–1.02)
	max: 2.68 (2.53–2.85)		
$\hat{R}$	—	—	2.58 (2.08–3.17)
σ	1.05 (0.90–1.22)	0.78 (0.68–0.91)	0.89 (0.76–1.07)
α	0.11 (0.09–0.14)	0.13 (0.11–0.15)	0.12 (0.1–0.14)
number of parameters	3	3	4
approximate AIC	3821.11	3557.39	3498.09
WAIC	7358.52	7235.64	6528.21

[a]Reproductive values are calculated using the inferred posterior distribution of β (results not shown). The minimum value refers to R at the lowest population estimate for any one latitude degree and time point, and the maximum value refers to R at the largest population estimate.

the peak, between latitudes 36–37° N and days 50–70 (as outlined in §2.3). This modified model provided the best overall fit to the data (see AIC and WAIC values, table 2) and was also able to capture the isolated peak better than either the frequency- or density-dependent models alone (figure 2b). We thus conclude that this model represents the closest approximation to the underlying dynamical process and use this for all further analyses.

We also investigated frequency- and density-dependent iterations of a binomial chain model incorporating susceptible depletion, but in each case the model was a poorer fit (both in terms of the information criterion values and in capturing the significant peak) than the self-exciting models without susceptible depletion (see electronic supplementary material, S6, for details). This result is discussed further below.

The baseline estimate for the effective reproduction number, R, from the best-fitting model is below one (table 2), reflecting the pattern of sporadic cases that characterizes the majority of the epidemic data. The additional estimate is substantially higher ($\hat{R} = 2.58$), capturing an increase in

transmission intensity around the main epidemic signal. The mean infectious period of an individual ($1/\alpha$) is 8.3 days, and 95% of an individual's infectiousness is estimated to occur within 24 days and 2 latitude degrees (roughly 220 km) of the initial infection event (figure 3a–d). These values represent upper bounds for the likely infectious period and spatial range of an infected bottlenose dolphin. In particular, the estimated infectious period is reasonable for the generation time of a morbillivirus [10].

The marginal distributions for the hazard function in space and time highlight hotspots of transmission risk (figure 4a,b). Across space, there is a clear global maximum around 36–37° N, with a local plateau around 30–32° N, and across time there is one global peak towards the end of summer (at around 70 days), and another local peak near the beginning of winter (around 160 days). This suggests these particular regions and time periods are important for propagating epidemic spread and reinforces the idea that the region between southeast Virginia and northeast North Carolina may be driving an increase in transmission intensity towards the end of summer.
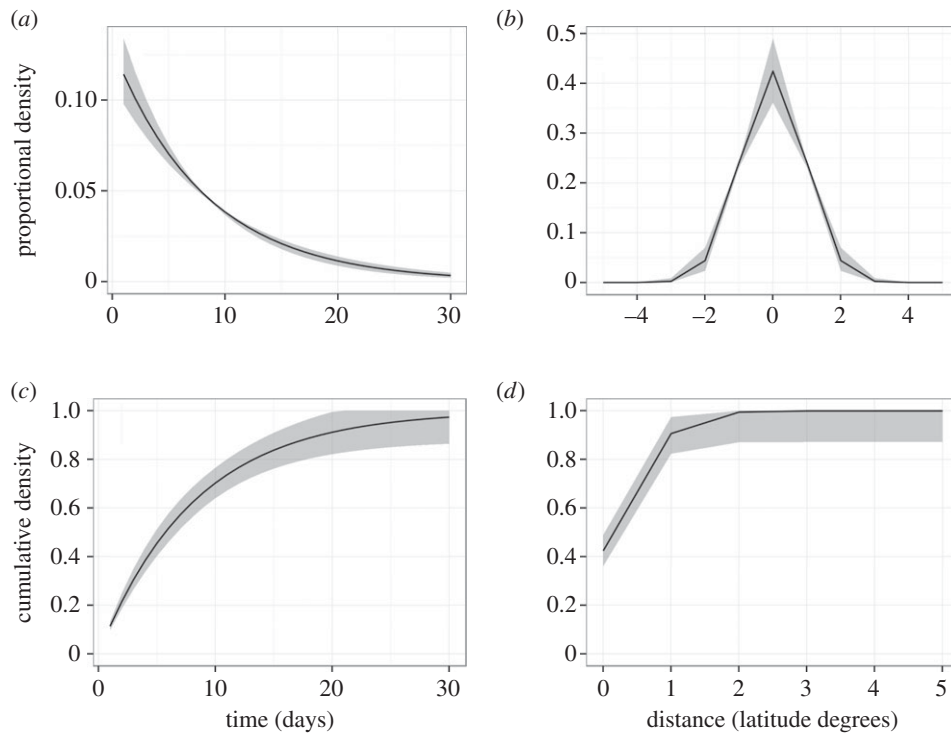
Figure 3. Individual infectiousness functions from the frequency-dependent model with location-specific transmission rate. Probability distributions for the temporal (a) and spatial decay functions (b). Cumulative distributions for the temporal (c) and spatial decay functions (d). Solid black lines are median values of 2000 simulations in RStan and grey shaded regions represent the 2.5th–97.5th quantile range.
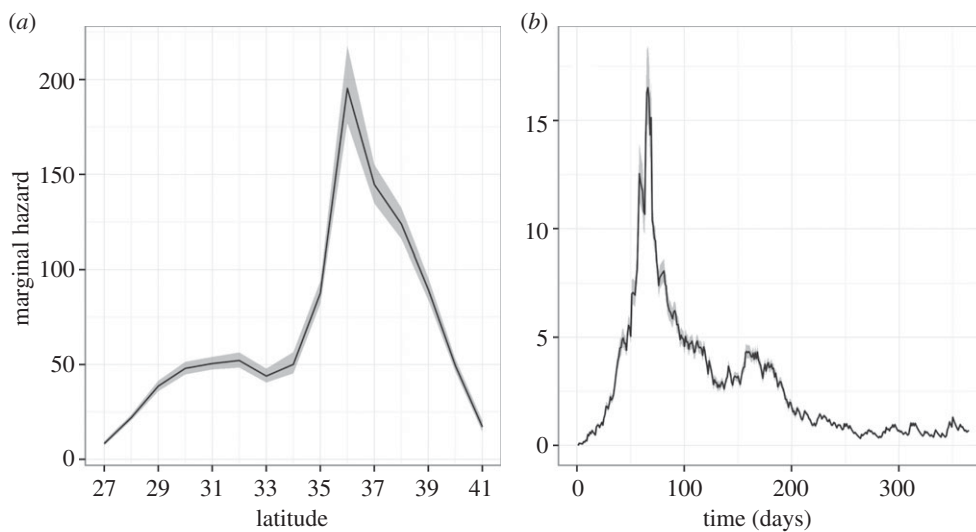

Figure 4. Marginal hazard distributions in space (a) and in time (b) for the frequency-dependent model with location-specific transmission rate. Solid black lines indicate median values of 2000 simulations in RStan and grey shaded regions represent the 2.5th–97.5th quantile range.

3.1. Transmission networks

Transmission networks were generated as outlined in §2.4. Transmission events resulting from movements of the order of 1–2 latitude degrees are most common (in accordance with our spatial range estimates), although long-range movements do occur and are important for seeding events in regions where there are no (or relatively few) infections (e.g. electronic supplementary material, figure S10). Therefore, high mobility of some individuals (e.g. during migration) may explain the wide dissemination of the virus along the coast.

To explore this further, we measured temporal changes in the betweenness of the network as an indication of the relative importance of each latitude degree in propagating the underlying transmission process. In general, the most influential nodes at any time point are clustered around the regions where transmission is greatest (figure 5). As would be expected, the nodes with the highest initial betweenness are those clustered around the northern latitudes where the epidemic is believed to have begun, in particular the 36–37° N region. Following the most influential nodes throughout the year then signals how the centre of mass of the epidemic changes: the highest betweenness values gradually shift to the more southerly regions during winter and autumn, before moving northwards again during the warmer months. This again suggests that seasonal migration is likely to play a key role in the spatial spread of the epidemic.

3.2. Gulf of Mexico

We simulated an epidemic in the GoM assuming frequency-dependence and using parameter estimates obtained from the
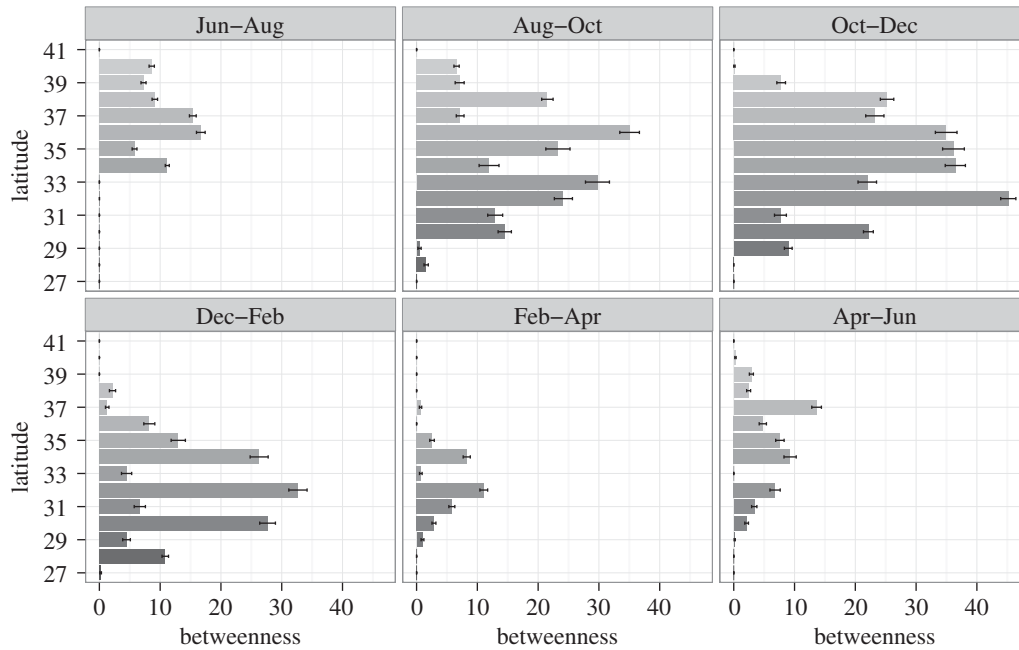
Figure 5. Betweenness of each latitude band in the reconstructed transmission networks across bimonthly intervals, starting 26 June 2013 (darker colours represent more southerly latitude points). Betweenness measures the number of shortest paths running through a particular node for each of the sampled networks. Mean values (grey shaded bars) and standard error bars (black lines) are calculated from 100 sampled network trees.
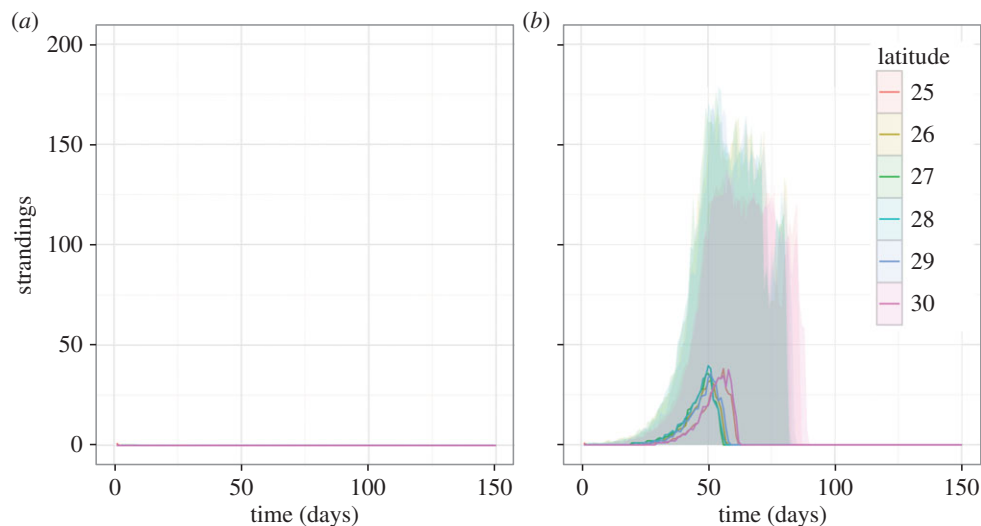

Figure 6. Simulations of an outbreak initiated in the GoM using parameters obtained from the best-fitting model. (a) Best-case scenario using the baseline reproduction number ($R_0 = 0.95$). (b) Worst-case scenario using the increased reproduction number ($\hat{R}_0 = 2.58$). Solid lines represent median values of 100 simulations and shaded regions represent the 2.5th – 97.5th quantile range. Latitude bands are distinguished by colour.

best-fitting model (as described in §2.5) to explore the likelihood of an outbreak in this region. We additionally assumed the population was entirely susceptible and so our NW Atlantic effective reproduction number estimates, $R = 0.95$ and $\hat{R} = 2.58$, can be viewed as lower bounds for R_0 and $\hat{R}_0$ in the GoM, representing the 'best' and 'worst' case scenarios, respectively. In the best-case scenario, the outbreak always fails to take off (figure 6a), whereas in the worst-case scenario, although not all outbreaks take off, those that do result in rapid viral spread until all susceptibles have been infected (figure 6b). Although all epidemics are initiated in the lowest latitude band (where dolphins from the GoM and NW Atlantic are most likely to come in contact), peaks tend to occur first in the intermediate latitudes as a result of the uniform distribution of individuals

and the additional force of infection gained from being surrounded by infected neighbours.

To investigate potential levels of pre-existing immunity in the GoM, we use standard epidemiological theory which states that R must be less than one to avoid an outbreak (where $R = R_0 \times s$ and s is the proportion of the population that are susceptible) [16]. In the best-case scenario, $R_0 < 1$ and so an outbreak will never take off. In the worst-case scenario, $\hat{R}_0 = 2.58$ and so

$$\hat{R} = 2.58 \times s < 1 \quad \Leftrightarrow \quad s < \frac{1}{2.58} = 0.388 \sim 38.8\%.$$

This suggests at least 61.2% of the population should be immune to avoid an outbreak.

4. Discussion

The aim of this paper was to gain insight into the spatio-temporal spread of DMV in the NW Atlantic using data that may represent only partially recovered *Tursiops* mortality. With a self-exciting Poisson framework, we were able to estimate previously unknown epidemiological parameters and reconstruct networks to visualize key aspects of the underlying transmission process. This has provided a greater understanding of the ecology of disease spread among the dolphin stocks in this region.

4.1. Model fits and parameter estimates

The best-fitting model assumed frequency-dependent transmission, which agrees with previous findings for PDV and measles [10,11,37]. Frequency-dependence suggests that transmission is largely independent of population density, and typically results from populations with stable social structures in which individuals generally have a fixed number of contacts (where 'contact' refers to interactions that may lead to disease transmission) [10,16]. However, at this stage, we cannot rule out density-dependent transmission fully, as the relatively poor overall fit of that model could be due to local inaccuracies in our population distribution estimates (which rely on historical stranding data and the assumption that stranding numbers in a particular area are proportional to the local population density). In particular, the density-dependent model captured the outlying peak in cases better than the original frequency-dependent model (with constant transmission intensity). This is most likely due to the large population sizes predicted in this region that increase the local hazard of the density-dependent model and thus allow more cases to be generated. More detailed data on stock movement, social structures and virus transmission pathways are thus needed to further understand the scaling of transmission in this system.

The estimate for the baseline effective reproduction number ($R = 0.95$) suggests low-intensity transmission and is consistent with the occurrence of sporadic cases and the absence of a characteristic epidemic growth phase. Allowing an additional transmission rate around the outlying peak (36–37° N, August 2013) increased the model fit and resulted in a substantial increase in transmission intensity ($\hat{R} = 2.58$). One driver of this increase may be seasonal migration. Around Virginia Beach, VA (a section of coastline within this 36–37° N region), Barco *et al.* [38] found substantially higher numbers of *Tursiops* in August than at any other time of year, potentially due to the overlapping of the NMCS and SMCS at the extreme boundaries of their migratory ranges. This aggregation of individuals at higher densities could cause nonlinear increases in the frequency of contact between hosts that would explain the subsequent jumps in transmission intensity that were not captured by the model with constant transmission parameter. Moreover, our additional estimate falls within the range of reproduction number estimates for PDV (2–3), for which high aggregation of seals at haul-out sites was believed to facilitate rapid virus transmission [11].

The infectious period of an individual had an upper bound of 24 days, with a mean of 8 days. This is similar to other morbillivirus infections, including measles (6–9 day latent period and 6–7 day infectious period) and PDV (11–18 day combined latent and infectious period) [10,11,39]. Two latitude degrees (approx. 220 km) was the upper bound for the

expected travel distance of an individual during their infectious period; further distances were possible, but rare. The coastal stocks exhibit both migratory and sedentary behaviour (depending on the time of year) and thus movement patterns are highly variable and difficult to quantify. However, our estimated ranges are supported by movement data from tagged NMCS individuals that show weekly travel distances of up to 347 km (mean: 114 km; 2.5th–97.5th quantiles: 2–299 km; $N = 240$) (A. Hohn 1998–2000, unpublished data). We note that the spatial and temporal decay parameters (σ and α) may be overestimated as our model does not account for the influence of unobserved cases. However, it is encouraging that our estimates do not vary substantially between the different models tested and that they are within such close ranges of the corresponding quantities discussed above from related viruses (measles and PDV) and the population movement data.

The inference process was also applied to frequency- and density-dependent models incorporating susceptible depletion, but the overall fits to the data were worse than both the frequency- and density-dependent Poisson processes. There are two likely reasons for this. First, since only reported strandings are included in the analysis and the total number of these observed cases is small relative to the total size of the NW Atlantic population, new infections are unlikely to significantly alter the available pool of susceptible individuals in our models. Second, our models do not explicitly account for the extent and distribution of pre-existing immunity to DMV, or the role of host migration in shaping transmission dynamics and the subsequent spatio-temporal distribution of susceptibility. These factors should strongly influence the importance of susceptible depletion in this system, but due to lack of information could not be included in the current framework without compromising model parsimony. A greater understanding of the underlying spatio-temporal distribution of susceptibility to DMV would enable future modelling to incorporate susceptible depletion in a more realistic framework.

4.2. Transmission networks

Analysis of the reconstructed transmission networks indicated that certain regions may disproportionately promote virus transmission at different times of the year, and long-range movements of dolphins are important for seeding the epidemic in new, unaffected regions. Measuring betweenness at discrete intervals suggested a seasonal shift in the centre of gravity of the underlying transmission process, further supporting the idea that stock migration is a key factor in propagating virus spread along the coast [24]. In particular, Rosel *et al.* [24] proposed that the seasonal spread of the 1987–1988 epidemic was a result of the virus being transferred from the NMCS to the SMCS following interaction between the two stocks at times of population overlap. More detailed information on the seasonal movements of the coastal stocks is needed to investigate this further.

4.3. Gulf of Mexico

Assuming the virus was introduced to bottlenose dolphins in the GoM (e.g. through contact with the NW Atlantic stocks) and the transmission intensity was equal to our increased estimate, $\hat{R}$, we found that an outbreak should occur unless there were high levels of pre-existing immunity in the population. The last confirmed DMV outbreak in the GoM occurred in 1994, with suspected outbreaks also occurring

in 1990 and 1992 [1,40]. Given that bottlenose dolphins in this region can live up to 40–50 years, there may be survivors in the current population that have maintained immunity from these previous outbreaks [41,42]. Although Rowles *et al.* [41] found seropositivity to DMV ranging from 0 to 18% between sampling regions in the GoM, the extent and spatial distribution of evidence of previous exposure across the wider population remains unclear, and further information is needed to compare this with our 61% estimate derived for the Eastern Coastal Stock.

On the other hand, we found that an outbreak would fail to take off at the baseline transmission rate, even in an entirely susceptible population. This could also sufficiently explain the lack of substantial DMV spread since only sporadic cases would be expected in this scenario. Given the association between $\hat{R}$ and large seasonal aggregations of individuals in the NW Atlantic, it is possible that the smaller size of the GoM Eastern Coastal Stock is insufficient to generate the frequency of contacts needed to support a transmission intensity of comparable levels. Further modelling work could help to identify a threshold population size above which density-dependent effects may become important in this region.

Seasonal migration and contact between different stocks are also key factors in promoting viral dissemination along the coast [1]. Therefore, another reason for limited spread could be that there is simply not enough contact between the GoM and NW Atlantic stocks to provide sufficient transmission opportunities. Rosel *et al.* [24] found substantial genetic differences between NW Atlantic and GoM bottlenose dolphins that was consistent with limited mixing between the two populations. However, given that an infected individual stranded as far south as the Florida Keys in November 2014 [43], there is clear potential for viral spread into the neighbouring GoM. More information on the levels of contact between these two regions would enable this to be explored further.

4.4. Caveats and future directions

Marine mammals, and especially cetaceans, represent challenging study subjects as it is common to have only limited data regarding population sizes, individual dispersal patterns and disease prevalence [44]. We used historical stranding distributions to infer the spatio-temporal distribution of population density and assumed that the observed epidemic strandings could be used as a proxy for the underlying infection process. However, there are a number of caveats to these assumptions. Firstly, the stranding data do not include those individuals that recover from infection or those that die from the disease but do not strand onshore, and so we have only a partial representation of the infection process. The self-exciting Poisson framework provides a means of identifying key epidemiological parameters from the outbreak despite these sparse data; however, there may be important information embedded in the unobserved infection process that is not accounted for. Unfortunately, integrating an additional unobserved process into the model is challenged by identifiability issues arising from the sparseness of the data and the computational intensity of performing inference on a model with more unknown variables.

Secondly, wind, oceanic currents and the length of coastline in a particular area can affect the likelihood of an individual stranding onshore, and the location and timing of the stranding should it occur [25,45]. Using daily wind force data, Rijks *et al.* [45] found that wind had a confounding effect on the distribution of seal strandings during the 2002 PDV outbreak in the Netherlands. Similar data from the NW Atlantic could be used to assess the influence of wind and other weather variables on stranding distributions. Future work could account for the length of available coastline when calculating the distribution of population density from historical strandings and could also integrate weather components into the model as an additional term affecting the location and timing of strandings. This would enable a more realistic representation of the stranding process and could also provide greater insight into the unobserved infection process.

Thirdly, temporal and spatial heterogeneity in reporting effort can influence the likelihood and promptness of carcass identification [25]. Although this may be a significant factor in the historical stranding data, we expect reporting to be uniformly high across all locations during the current UME due to the public attention and rapid response of NOAA. Again, more detailed information on the distribution of the stocks would better inform our population estimates and decrease the likelihood of historical reporting bias influencing our findings.

In addition, we assume the population is closed and do not include the reintroduction of DMV from other marine mammal populations that have been implicated as reservoir hosts, such as pilot whales and offshore bottlenose dolphins [1,46]. Such spillover could be incorporated into the model via an additional hazard function term representing some baseline risk of external infection [20,21]. However, further information on the prevalence of DMV in other Atlantic cetaceans, and the frequency of contact between these populations and the coastal bottlenose stocks, would be needed to assess the likelihood of these events and guide the parametrization and functional form of such a term.

Our hazard function also assumes an exponential form for the spatial and temporal decay of infectiousness and does not explicitly account for individual migration. There have been more complex models of dolphin movement but these have been based on explicit tracking of individuals [47,48]. Since our stranding data contain no information on how animals move during the epidemic, we have adopted the most parsimonious framework to describe dispersal that can provide biologically interpretable results in broad agreement with existing knowledge and data. More information on how virus infectivity changes over time within a given individual and on the complex movements, migration patterns, and interactions between the different stocks in the NW Atlantic is needed to guide alternative hazard functions. In particular, tagging individuals during disease outbreaks is an important area for future work.

Finally, age-structured models have provided key insights into the influence of social structure on the spread of other morbilliviruses such as PDV and measles [11,49]. In particular, subadult harbour seals were disproportionately important in transmitting PDV during the 2002 Dutch outbreak [11]. Juvenile bottlenose dolphins tend to have a larger number of contacts than individuals in any other age class [50], and thus may similarly be important vectors driving the spread of DMV in the NW Atlantic. In addition, there may be differences in stranding times between age classes that could only be accounted for in an age-structured model. Estimates of age for stranded *Tursiops* (measured by length) are available from the current UME dataset. This would therefore be another interesting covariate for future analyses.

In conclusion, despite many unknowns surrounding the spread of DMV in cetaceans, we have gained important insights into the epidemiology of this pathogen by using a simple statistical framework to model the current NW Atlantic UME. Furthermore, there is clear potential to extend this methodology beyond the current framework (e.g. [20,21,28,51–54]), although requirements for future applications will depend on the system under consideration, in particular on the resolution and spatio-temporal patterns of the observed data, and the complexity of the desired hazard function for capturing underlying disease dynamics. Overall, however, we believe this work illustrates the general utility of self-exciting Poisson models to analyse partially observed spatial epidemics. Future serological studies, migration data and population census information would enable more rigorous testing of the key findings outlined in this work and provide a platform for the development of more sophisticated frameworks to model the underlying system dynamics.

References

1. Duignan PJ, House C, Odell DK, Wells RS, Hansen LJ, Walsh MT, St Aubin DJ, Rima BK, Geraci JR. 1996 Morbillivirus infection in bottlenose dolphins: evidence for recurrent epizootics in the western Atlantic and Gulf of Mexico. *Mar. Mamm. Sci.* **12**, 499–515. (doi:10.1111/j.1748-7692.1996.tb00063.x)

2. Van Bressem M-F *et al.* 2001 An insight into the epidemiology of dolphin morbillivirus worldwide. *Vet. Microbiol.* **81**, 287–304. (doi:10.1016/S0378-1135(01)00368-6)

3. Domingo M *et al.* 1990 Morbillivirus in dolphins. *Nature* **348**, 21. (doi:10.1038/348021a0)

4. Barrett T, Visser IKG, Mamaev L, Goatley L, Van Bressem M-F, Osterhaus ADME. 1993 Dolphin and porpoise morbilliviruses are genetically distinct from phocine distemper virus. *Virology* **193**, 1010–1012. (doi:10.1006/viro.1993.1217)

5. Van Bressem M-F, Van Waerebeek K, Raga JA. 1999 A review of virus infections of cetaceans and the potential impact of morbilliviruses, poxviruses and papillo-maviruses on host population dynamics. *Dis. Aquat. Organ.* **38**, 53–65. (doi:10.3354/dao038053)

6. NOAA Fisheries. 2015 FAQs on Morbillivirus and the Atlantic UME. See http://www.nmfs.noaa.gov/pr/health/mmume/faqs_morbillivirus_ume.html.

7. Lipscomb TP, Schulman FY, Moffett D, Kennedy S. 1994 Morbilliviral disease in Atlantic bottlenose dolphins (*Tursiops truncatus*) from the 1987–1988 epizootic. *J. Wildl. Dis.* **30**, 567–571. (doi:10.7589/0090-3558-30.4.567)

8. Scott GP, Burn DM, Hansen LJ. 1988 The dolphin dieoff: long-term effects and recovery of the population. In *Proc. OCEANS'88. A Partnership of Marine Interests*, pp. 819–823. (doi:10.1109/OCEANS.1988.794905)

9. Eguchi T. 2002 A method for calculating the effect of a die-off from stranding data. *Mar. Mamm. Sci.* **18**, 698–709. (doi:10.1111/j.1748-7692.2002.tb01067.x)

10. Swinton J, Harwood J, Grenfell BT, Gilligan CA. 1998 Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. *J. Anim. Ecol.* **67**, 54–68. (doi:10.1046/j.1365-2656.1998.00176.x)

11. Klepac P, Pomeroy LW, Bjørnstad ON, Kuiken T, Osterhaus ADME, Rijks JM. 2009 Stage-structured transmission of phocine distemper virus in the Dutch 2002 outbreak. *Proc. R. Soc. B* **276**, 2469–2476. (doi:10.1098/rspb.2009.0175)

12. Grenfell BT, Lonergan ME, Harwood J. 1992 Quantitative investigations of the epidemiology of phocine distemper virus (PDV) in European common seal populations. *Sci. Total Environ.* **115**, 15–29. (doi:10.1016/0048-9697(92)90029-R)

13. Härkönen T *et al.* 2006 A review of the 1988 and 2002 phocine distemper virus epidemics in European harbour seals. *Dis. Aquat. Organ.* **68**, 115–130. (doi:10.3354/dao068115)

14. Black FL. 1991 Epidemiology of paramyxoviridae. In *The paramyxoviruses* (ed. DW Kingsbury), pp. 509–536. New York, NY: Springer.

15. Van Bressem M-F *et al.* 2014 Cetacean morbillivirus: current knowledge and future directions. *Viruses* **6**, 5145–5181. (doi:10.3390/v6125145)

16. Keeling MJ, Rohani P. 2008 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.

17. Fertl D. 1994 Occurrence patterns and behavior of bottlenose dolphins (*Tursiops truncatus*) in the Galveston ship channel, Texas. *Texas J. Sci.* **46**, 299–318.

18. Ogata Y. 1998 Space-time point-process models for earthquake occurrences. *Ann. Inst. Stat. Math.* **50**, 379–402. (doi:10.1023/A:1003403601725)

19. Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, Tita GE. 2011 Self-exciting point process modeling of crime. *J. Am. Stat. Assoc.* **106**, 100–108. (doi:10.1198/jasa.2011.ap09546)

20. Meyer S, Elias J, Höhle M. 2012 A space–time conditional intensity model for invasive meningococcal disease occurrence. *Biometrics* **68**, 607–616. (doi:10.1111/j.1541-0420.2011.01684.x)

21. Höhle M. 2009 Additive-multiplicative regression models for spatio-temporal epidemics. *Biom. J.* **51**, 961–978. (doi:10.1002/bimj.200900050)

22. NOAA Fisheries. 2014 FAQs on the ongoing Gulf of Mexico dolphin die-off. See http://www.nmfs.noaa.gov/pr/health/mmume/cetacean_gulfofmexico_faq.htm.

23. Waring GT, Josephson E, Maze-Foley K, Rosel PE. 2014 U.S. Atlantic and Gulf of Mexico marine mammal stock assessments—2013. *NOAA Tech. Memo.* **228**, 1–464.

24. Rosel PE, Hansen L, Hohn AA. 2009 Restricted dispersal in a continuously distributed marine species: common bottlenose dolphins *Tursiops truncatus* in coastal waters of the western North Atlantic. *Mol. Ecol.* **18**, 5030–5045. (doi:10.1111/j.1365-294X.2009.04413.x)

25. McLellan WA, Friedlaender AS, Mead JG, Potter CW, Pabst DA. 2002 Analysing 25 years of bottlenose (*Tursiops truncatas*) strandings along the Atlantic coast of the USA: do historic records support the migratory stock hypothesis? *J. Cetacean Res. Manage.* **4**, 297–304.

26. Daley DJ, Vere-Jones D. 2003 *An introduction to the theory of point processes. Volume I: elementary theory and methods*, 2nd edn. New York, NY: Springer.

27. Hawkes AG. 1971 Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90. (doi:10.1093/biomet/58.1.83)

28. Diggle PJ. 2007 Spatio-temporal point processes: methods and applications. In *Statistical methods for spatio-temporal systems* (eds B Finkenstädt, L Held, V Isham), pp. 1–45. Boca Raton, FL: CRC.

29. Stan Development Team. 2014 Stan: a C++ library for probability and sampling, v. 2.2. See http://mc-stan.org/.

30. R Core Team. 2014 R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

31. Akaike H. 1974 A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723. (doi:10.1109/TAC.1974.1100705)

32. Watanabe S. 2010 Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11, 3571–3594.

33. Gelman A, Hwang J, Vehtari A. 2014 Understanding predictive information criteria for Bayesian models. Stat. Comput. 24, 997–1016. (doi:10.1007/s11222-013-9416-2)

34. Csardi G. 2015 Package 'igraph'. The Comprehensive R Archive Network. See http://cran.r-project.org/web/packages/igraph/igraph.pdf.

35. Girvan M, Newman MEJ. 2002 Community structure in social and biological networks. Proc. Natl Acad. Sci. USA 99, 7821–7826. (doi:10.1073/pnas.122653799)

36. Waring GT, Josephson E, Maze-Foley K, Rosel PE. 2013 U.S. Atlantic and Gulf of Mexico marine mammal stock assessments—2012. NOAA Tech. Memo. 223, 1–419.

37. Bjørnstad ON, Finkenstädt BF, Grenfell BT. 2002 Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. Ecol. Monogr. 72, 169–184. (doi:10.1890/0012-9615(2002)072[0169:DOMEES]2.0.CO;2)

38. Barco SG, Swingle WM, McLellan WA, Harris RN, Pabst DA. 1999 Local abundance and distribution of bottlenose dolphins (Tursiops truncatus) in the nearshore waters of Virginia Beach, Virginia. Mar. Mamm. Sci. 15, 394–408. (doi:10.1111/j.1748-7692.1999.tb00809.x)

39. Anderson RM, May RM. 1992 Infectious diseases of humans: dynamics and control. Oxford, UK: Oxford Science Publications.

40. Litz JA et al. 2014 Review of historical unusual mortality events (UMEs) in the Gulf of Mexico (1990–2009): providing context for the multi-year northern Gulf of Mexico Cetacean UME declared in 2010. Dis. Aquat. Organ. 112, 161–175. (doi:10.3354/dao02807)

41. Rowles TK, Schwacke LS, Wells RS, Saliki JT, Hansen L, Hohn AA, Townsend F, Sayre RA, Hall AJ. 2011 Evidence of susceptibility to morbillivirus infection in cetaceans from the United States. Mar. Mamm. Sci. 27, 1–19. (doi:10.1111/j.1748-7692.2010.00393.x)

42. Wells RS, Scott MD. 1999 Bottlenose dolphin Tursiops truncatus (Montagu, 1821). Handb. Mar. Mamm. 6, 137–182.

43. NOAA Fisheries. 2015 2013–2015 Bottlenose dolphin unusual mortality event in the mid-Atlantic. See http://www.nmfs.noaa.gov/pr/health/mmume/midatldolphins2013.html.

44. Gulland FMD, Hall AJ. 2007 Is marine mammal health deteriorating? Trends in the global reporting of marine mammal disease. EcoHealth 4, 135–150. (doi:10.1007/s10393-007-0097-1)

45. Rijks JM, Van de Bildt MWG, Jensen T, Philippa JDW, Osterhaus ADME, Kuiken T. 2005 Phocine distemper outbreak, The Netherlands, 2002. Emerg. Infect. Dis. 11, 1945–1948. (doi:10.3201/eid1112.050596)

46. Duignan PJ et al. 1995 Morbillivirus infection in cetaceans of the western Atlantic. Vet. Microbiol. 44, 241–249. (doi:10.1016/0378-1135(95)00017-5)

47. Viddi CFA, Harcourt RG, Hucke-Gaete R, Field IC. 2011 Fine-scale movement patterns of the sympatric Chilean and Peale's dolphins in the Northern Patagonian Fjords. Mar. Ecol. Prog. Ser. 436, 245–256. (doi:10.3354/meps09251)

48. Bailey H, Thompson P. 2006 Quantitative analysis of bottlenose dolphin movement patterns and their relationship with foraging. J. Anim. Ecol. 75, 456–465. (doi:10.1111/j.1365-2656.2006.01066.x)

49. Schenzle D. 1984 An age-structured model of pre- and post-vaccination measles transmission. Math. Med. Biol. 1, 169–191. (doi:10.1093/imammb/1.2.169)

50. Wells RS. 2014 Social structure and life history of bottlenose dolphins near Sarasota Bay, Florida: insights from four decades and five generations. In Primates and cetaceans: field research and conservation of complex mammalian societies (eds J Yamagiwa, L Karczmarski), pp. 149–172. Tokyo, Japan: Springer.

51. Cauchemez S, Ferguson NM. 2012 Methods to infer transmission risk factors in complex outbreak data. J. R. Soc. Interface 9, 456–469. (doi:10.1098/rsif.2011.0379)

52. Cauchemez S et al. 2011 Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. Proc. Natl Acad. Sci. USA 108, 2825–2830. (doi:10.1073/pnas.1008895108)

53. Scheel I, Aldrin M, Frigessi A, Jansen PA. 2007 A stochastic model for infectious salmon anemia (ISA) in Atlantic salmon farming. J. R. Soc. Interface 4, 699–706. (doi:10.1098/rsif.2007.0217)

54. Aldrin M, Storvik B, Frigessi A, Viljugrein H, Jansen PA. 2010 A stochastic model for the assessment of the transmission pathways of heart and skeleton muscle inflammation, pancreas disease and infectious salmon anaemia in marine fish farms in Norway. Prev. Vet. Med. 93, 51–61. (doi:10.1016/j.prevetmed.2009.09.010)