


Reconstructing disease transmission dynamics from animal movements and test data

J. A. Enright¹ · A. O'Hare¹ 

Published online: 16 November 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Disease outbreaks are often accompanied by a wealth of data, usually in the form of movements, locations and tests. This data is a valuable resource in which data scientists and epidemiologists can reconstruct the transmission pathways and parameters and thus devise control strategies. However, the spatiotemporal data gathered can be both vast whilst at the same time incomplete or contain errors frustrating the effort to accurately model the transmission processes. Fortunately, several techniques exist that can be used to infer the relevant information to help explain these processes. The aim of this article is to provide the reader with a user friendly introduction to the techniques used in dealing with the large datasets that exists in epidemiological and ecological science and the common pitfalls that are to be avoided as well as an introduction to inference techniques for estimating parameter values for mathematical models from spatiotemporal datasets.

Keywords Epidemiology · Modelling · Bayesian Inference · Simulation · Networks · Spatio-temporal

1 Introduction

Spatiotemporal data are those that contain both spatial (location) and temporal (time) properties. In veterinary epidemiology, this may be the records of tests carried out

at a particular location (e.g. a farm) and time or simply the movement of animals, the recording of which is often mandated by governments and provides researchers with a wealth of data with which to analyse the outbreak and transmission of diseases (Moustakas and Evans 2016; Hong and Paik 2012). In many cases the volume of collected data poses a significant statistical and computational challenge to the understanding of both outbreak patterns, transmission dynamics and thus control of the epidemic.

Most often, the transmission processes of a disease are, at least, partially understood (Lowe et al. 2015; Reiczigel et al. 2010). For example, diseases such as foot and mouth disease (FMD) (Kao et al. 2007; Keeling et al. 2001), avian influenza (AI) (Gumel 2009) and classical swine fever (CSF) (González-Parraa et al. 2011) are spread by close contact with infected individuals with little or no latent stages while diseases such as bovine tuberculosis (bTB) (Moustakas and Evans 2015; Biek et al. 2012) have long latent periods and all contain a temporal element in the form of animal movements. From a phenomenological perspective we can write simple compartmental models for these diseases and solve them on the network of farms (spatially) for a period of time (temporally) incorporating the movements of (potentially infected) animals in the model. This presupposes that we know the transmission parameters which, in reality, are either unknown or estimated.

Several techniques can be used to extract useful information from these datasets or to infer disease transmission parameters. In this article we will review some useful techniques that can be used to obtain pertinent information from a large spatiotemporal dataset and, using generated datasets, provide examples of these techniques.

✉ A. O'Hare
anthony.ohare@stir.ac.uk

J. A. Enright
jae@cs.stir.ac.uk

¹ Computing Science and Mathematics, University of Stirling, Stirling FK9 4LA, UK

2 Network analysis of spatiotemporal data

Purely spatial models are often confounded in modelling disease spread when there is occasional long-distance infectious contacts [e.g. livestock trading over long distances, which played an important role in the 2001 foot-and-mouth outbreak in Britain (Kiss et al. 2006; Kao et al. 2007)]. When either a reasonable model or explicit records are available for these pairwise contacts, it may be appropriate to model them using a contact network. Doing so makes available a variety of network analysis methods, which have become very popular (e.g. Stärk et al. 2006; Dube et al. 2009; Martínez-López et al. 2009) and are often referred to as “social network analysis”.

Traditionally, it has been common to ignore the dynamic nature of many contact networks, and instead process known or modelled contacts into a static network prior to analysis: common approaches have included aggregation of contacts over some appropriate time frame into a “snapshot” network, or taking an average network over a longer time period. More complex adaptations are possible to preserve more information (Holme 2013).

However, due in part to better data resolution and the increase in availability of analytic tools, it is becoming more common to include the dynamic nature of contact networks in epidemiological analyses. There is evidence that ignoring temporal information about contacts can give a deceptive picture. For example, consider two different orderings of infectious contacts: either a contact between A and B followed by a contact between B and C, or the alternate ordering of a contact between B and C followed by a contact between A and B. In the first case, there is potential for pathogen movement from A to C, but in the second, there is no such possibility. A static aggregation of either ordering would result in the same network, and potentially identify pathogen flow from A to C as a possibility in both cases.

The size of a connected component (a maximal joined-up set of nodes) is often used as an upper bound on the maximum outbreak size in a static network (Dube et al. 2009). In a dynamic network, it is more appropriate to use the idea of an infection chain: the size of a set of nodes that could potentially be infected by temporally possible routes from a single starting point of an outbreak (Dube et al. 2009; Nöremark and Widgren 2014).

There are several pieces of software available to compute such measures on dynamic networks, including EpiContactTrace in R (Nöremark and Widgren 2014), Gephi (a standalone graphical interface for network analysis) (Bastian et al. 2009), ORA-LITE as part of the CASOS project, or the Python module `networkx` (Hagberg et al. 2008). For our example below, we have used `networkx`,¹ but any of the other

available packages would have sufficed: in the subsequent parameter-estimation example we use a different open-source package: Broadwick, written in Java (O’Hare et al. 2016).

To aid in showing the importance of temporal information to understanding a dynamic network’s impact on disease spread, we have generated a simulated dataset of livestock trading amongst farms in a fictional island nation, which we will call Florin. We depict the locations of the fictional farms on a map of Florin in Fig. 1. We provide fictional locations and trades, along with the python code used in this example as supplementary material, in the hope it may serve as a basic tutorial. We will first inspect the network derived from our cattle trades and calculate some summary statistics. We find the data for our network in the cattle trades listed in `movements.csv` (part of S1), where each line contains an ID for a source farm, an ID for a destination farm, and the day number when the movement took place (note that while we have used non-negative integer numbers for the dates, most software is also capable of dealing with string-formatted dates).

In the code in S2, we first load the geographic locations and trade network into a `networkx` directed graph, and then plot it, giving the directed, spatially-embedded network in Fig. 2. The entire network is fairly dense, so we also plot the network composed of only movements on the first day. We then compute the frequencies of out-degrees by node (Fig. 3), aggregating all edges together over time. These sorts of degree distributions are widely used in static networks as an important network characteristic (Dube et al. 2009), but adaptations are increasingly being made to the dynamic setting (Nöremark and Widgren 2014; Holme 2013). One simple adaptation requires us to define a time window size, and calculate degrees of nodes within that size of time window. In the code in S2, we calculate the mean and maximum in-degree by node over a variety of time windows, and plot this in Fig. 4.

Our out-degree distribution in Fig. 3 gives us some information on our network, and allows us to compare it to other well-studied networks: a long-tailed distribution is common in real data-derived networks, and has important implications for spreading processes on the network (Estrada 2010). Our plots of changes in out-degree over differing time windows. Figure 4, show simple linear growth: the expected mean and maximum out-degree are proportional to the time window, with no special important size of time window. Often in real data there is an important time-window: for example, degrees in the Scottish cattle trading network increase dramatically at time windows that are multiples of seven days, due to the weekly timing of British cattle trading markets Cattle Tracing System.

2.1 Importance of temporal information to maximum possible outbreak size

We now turn our attention to the maximum possible outbreak size on our network, furnishing an example of the

¹ The code and data used in this paper are available at <https://github.com/EPICScotland/Broadwick/tree/master/examples/NetworkedSir>.

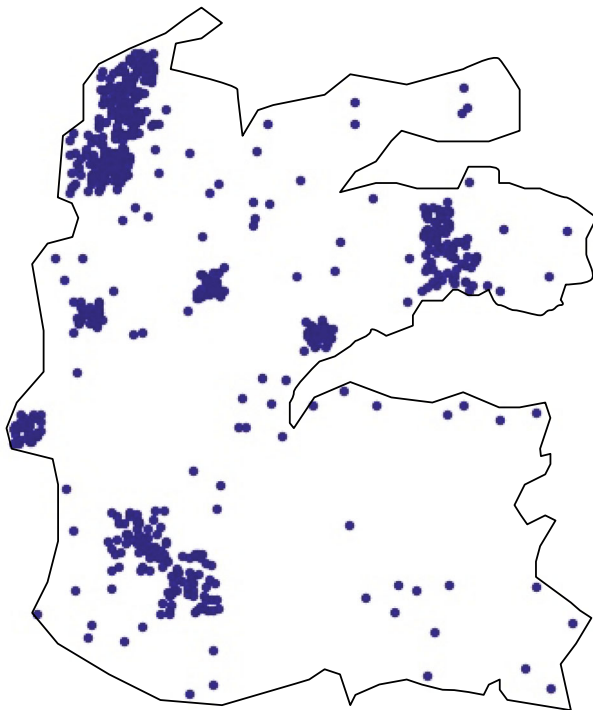


Fig. 1 A map of the fictional island nation Florin, with the locations of its cattle-trading farms shown as dots

importance of temporal information. We take two approaches to measure a possible outbreak size, and find very different answers. In the first approach, we ignore the timing of the movements, and create one aggregated static network with holdings as nodes and a directed link from one to the other if there has been a cattle trade between the holdings in that direction (as in Fig. 2).

We then calculate the number of holdings that are “downstream” of a holding - that is, could be reached by a directed path in the network. We find a directed component

Fig. 2 A spatial embedding of the entire fictional cattle trade network (left) and the fictional cattle trade network of only movements on the first day (right) given in `movements.csv` in `S1`, with edge directions shown by thicker rectangles at the destination of the edge

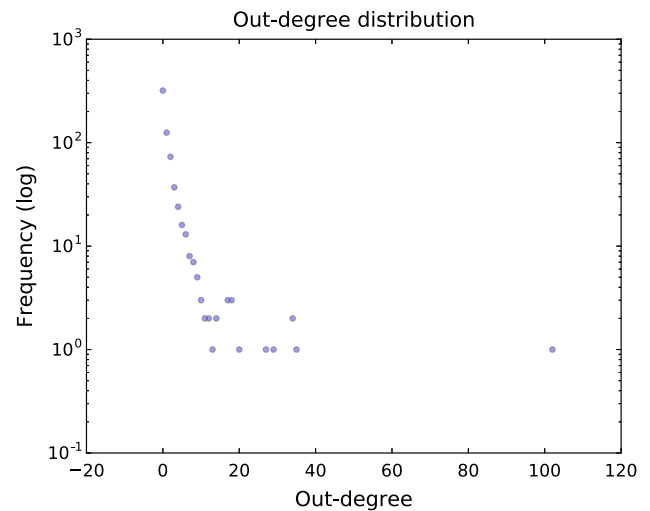
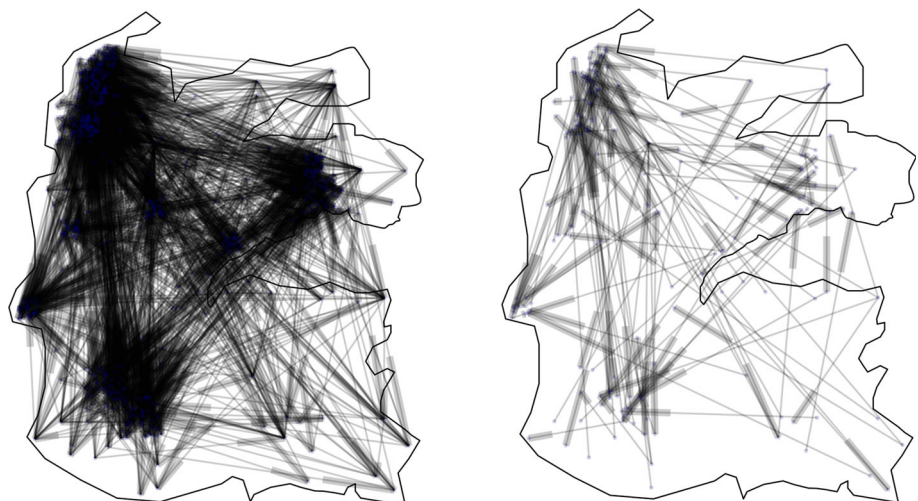


Fig. 3 Frequencies of out-degree by node in the fictional cattle trade network in Florin

of 649 downstream holdings in the network, suggesting that in the worst case scenario in which every contact between an infected farm and a susceptible farm transmits disease, a (fictional) pathogen could spread to up to 649 farms. If this worst-case infection is seeded at random throughout the aggregated network, the mean outbreak size is 17 farms.

In contrast, in our second approach, we include the timing of the cattle trades and use a dynamic network for our analysis, and we find a maximum infection chain size of 314, and a mean outbreak size of 5 farms.

Ignoring the timing of movements can also give a deceptive measure of infectious distance between two nodes (here farms), where infectious distance is the number of links in a network that an infectious would have to travel over to move from one node to the other Fig. 5.

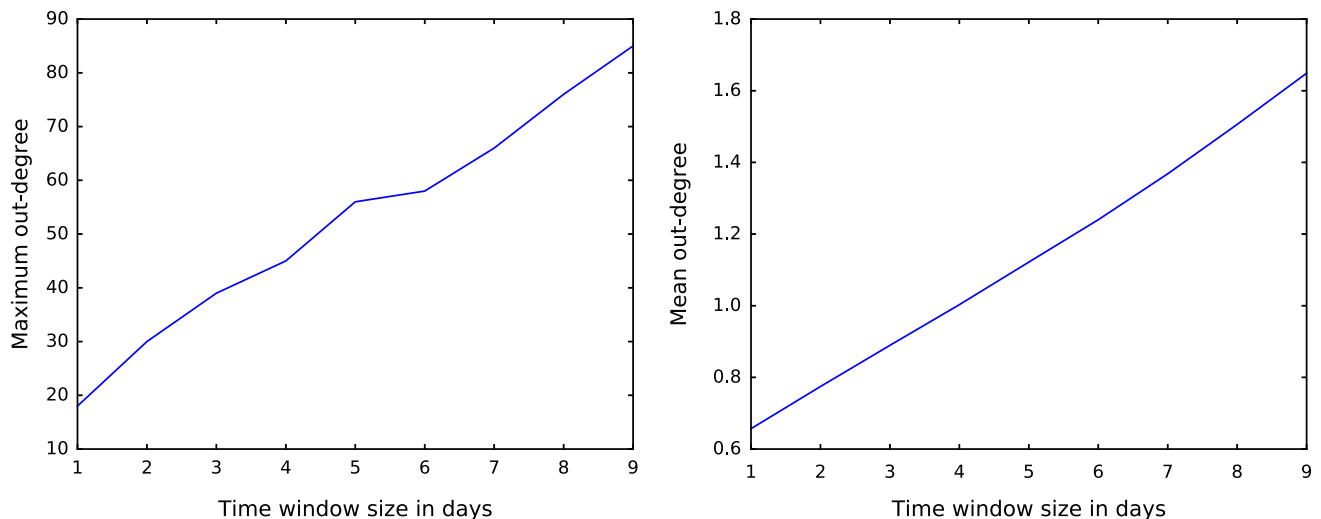
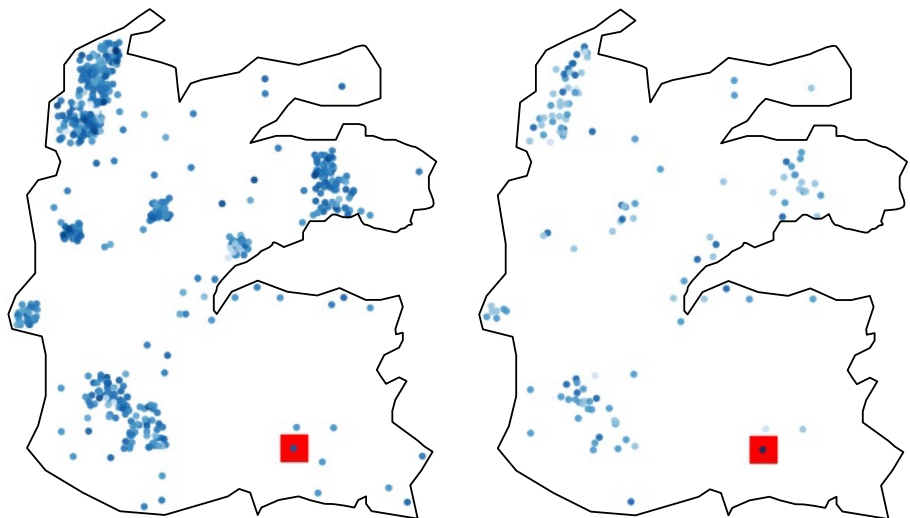


Fig. 4 Node-wise maximum (*left*) and mean (*right*) out-degree over a variety of sizes of time windows in the fictional cattle trading network of Florin

Fig. 5 Holdings that are reachable from the farm shown with a *red square* in networks that aggregate and ignore time (on the *left*), and fully consider time (on the *right*). Farms shown by *darker dots* are farther in network distance. The fictional farm locations and movements available as S1 were used to create these figures, along with python code in S2



2.2 Real-data example

We give a real-data example of a similar difference using Scottish data in Example 1 and Fig. 6.

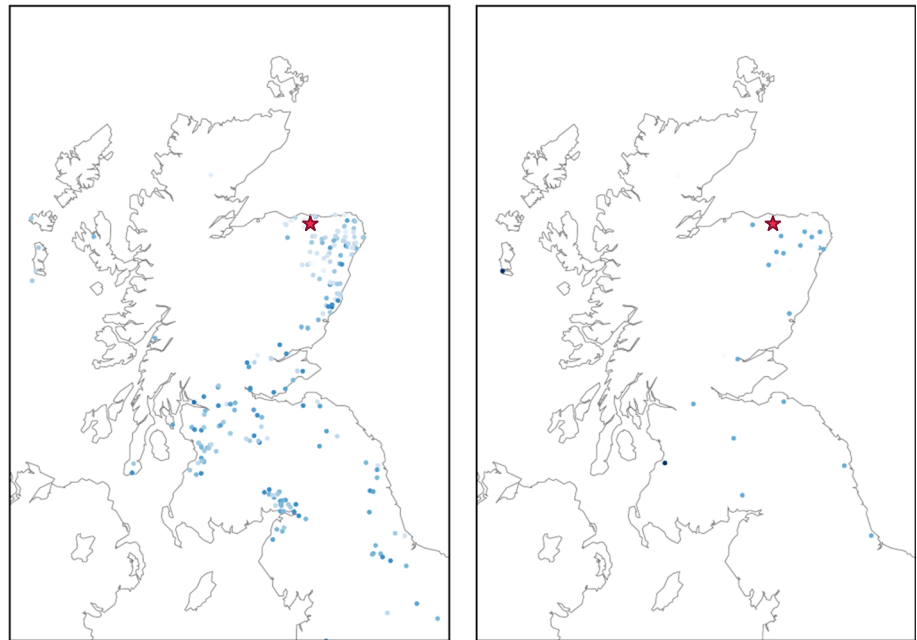
Example 1: In Fig. 6 a picture of a single holding in Scotland and the network distance between it and other holdings in the aggregate network and the fully dynamic network ScotEID—scottish EID livestock traceability research. We can see that many holdings that are reachable in the time aggregating network are not reachable in the network that fully considers time, and that the distances for holdings that are reachable in both are not preserved. If we ignored the timing of cattle movements, we might think that all the holdings shown in blue dots on the lefthand side of the below figure are close in the network to the starred holding, and therefore at risk of a disease in an outbreak involving the starred holdings, but we can see from the right hand side of the figure that this isn't actually the case.

Thus far we have restricted ourselves to considering a known network without any spatial interactions or stochastic disease processes: our focus has been on examining the network itself. We now turn our attention to a realistic example of a disease and the Bayesian methods that can help us infer parameters required in modelling it.

3 Parameter inference techniques

For disease outbreaks that are accompanied by spatiotemporal data such as the case of an outbreak of a disease on a single farm with known cattle movements as in Example 1 we can construct a relatively simple agent based model where each agent has a disease state and a location. An important consideration in developing such models is how

Fig. 6 Holdings that are reachable from the farm shown with a red star in networks that aggregate and ignore time (on the left), and fully consider time (on the right). Farms shown by darker dots are farther in network distance. Scottish cattle movements in January of 2010 have been used to create these networks



to obtain meaningful or realistic values for the parameters therein.

In this section we will summarise some techniques for estimating the transmission parameters for a disease that has recorded spatiotemporal data. Broadly speaking, these techniques fall into two categories depending on whether or not we can write out a likelihood function or if the calculation of this function is computationally unfeasible. These methods have the aim of finding those parameters for a particular model that best describe an observed set of data, such as test results, by exploring the space of all parameters through the use of a random walk through this space. The area of space that provides a best fit for the model to the data is recorded and the distribution of each parameter value in this region of space (often referred to as the *posterior* distribution) provides an estimate of the parameters.

We will illustrate one of these techniques using an SIR model on the same data as was used in Sect. 2.

In the following sections we will adopt the following notation, consistent with current literature: θ is a vector of unknown parameters that we wish to infer given some set of observations, \mathbf{D} . We will denote $\eta(\cdot)$ as the computational/mathematical model which will produce a range of possible outcomes that we will write as $X \sim \eta(\theta)$ when run repeatedly for the same set of inputs.² Using these

notations we can write the likelihood of the data under the model given the parameters θ as $\pi(\mathbf{D}|\theta)$.

The Bayesian approach is to find the posterior distribution of θ given \mathbf{D} as

$$\pi(\theta|\mathbf{D}) = \frac{\pi(\mathbf{D}|\theta)\pi(\theta)}{\pi(\mathbf{D})}$$

where $\pi(\theta)$ as the *prior* distribution and reflects the assumptions about the parameters in the model and $\pi(\mathbf{D})$ is the observed data.

3.1 Likelihood-free methods

Approximate Bayesian computation (ABC) are a collection of methods for performing Bayesian inference without the calculation of a likelihood function and are sometimes referred to as likelihood-free algorithms. Recently, these methods have become very popular in biological sciences, most notably genetics (Tanaka et al. 2006; Beaumont et al. 2002) and population biology (Lopes and Beaumont 2010) due to the fact the likelihood function can be difficult or impossible to compute for some models. In this section we will summarise how the method is used in practise, a fuller description of the technique is given in Csilléry et al. (2010).

The most basic form of the ABC algorithm is based on a rejection algorithm and given as:

- (0) Calculate a measure that characterises the system for the observed data \mathbf{D} .
- (1) Draw θ from $\pi(\theta)$.
- (2) Simulate $X \sim \eta(\theta)$.

² This model may either be stochastic in nature, incorporating random mutations in some disease transmitting pathogen, random movements in a network or simply solved using a Gillespie-type algorithm.

- (3) Calculate the *distance measure*, $\rho(\mathbf{X}, \mathbf{D})$, and accept θ if $\rho(\mathbf{X}, \mathbf{D}) \leq \delta$ where δ is the tolerance (accuracy) of the estimation method.
- (4) Repeat these steps until a sufficient number of accepted θ s are drawn.

The accepted values of θ are not drawn from the posterior distribution but an approximation to it (written as $\pi(\theta | \rho(\mathbf{D}, \mathbf{X}) \leq \delta)$). When $\delta = 0$ this algorithm draws from the posterior distribution $\pi(\theta | \mathbf{D})$. The smaller the value of δ the more accurate the approximation to the posterior distribution but this comes with added computational cost. The *distance measure*, $\rho(\mathbf{X}, \mathbf{D})$, is usually taken as the euclidean distance $\|\mathbf{X} - \mathbf{D}\|$. If θ is large (i.e. the data are high dimensional) is common to use a summary statistic to summarise the model output and data and thus reduce the dimensionality of the space. This choice of summary statistic is crucial for the quality of the approximation (Beaumont et al. 2009). In this scenario, step 3 above would be written

- (3) Accept θ if $\rho(S(\mathbf{X}), S(\mathbf{D})) \leq \delta$, where $S(\cdot)$ denotes a summary statistic.

Of course, a poor choice of summary statistic will add another layer of approximation to that already added by the use of a distance measure and tolerance.

For a detailed explanation of the ABC algorithm and its variants applied to several models see Turner and Van Zandt (2012). This ABC algorithm has been extended recently to approximate Markov Chain Monte Carlo algorithms (Marjoram et al. 2003) and to approximate sequential Monte Carlo algorithms (Sisson et al. 2007).

3.2 Monte Carlo methods

If it is possible to calculate the *Likelihood* function, i.e. probability of observing \mathbf{D} given a set of parameters, $\pi(\mathbf{D} | \theta)$ in a computationally tractable manner, the goal is to find those parameters that maximise this function. If there is some a priori knowledge of the model parameters, these can be incorporated into the search, a method referred to as Maximum a Posteriori (MAP) estimation, and is often more appropriate for the models encountered in ecology and epidemiology. MAP is used to estimate a mode of the posterior distribution (the distribution of θ that maximises the likelihood function). This a priori knowledge (*prior distribution* or simply *priors*) can be as simple as a uniform distribution within some wide limits for priors that are not well known to specific distributions with low measures of spread for well known priors. The posterior distribution of the parameters given the observed data can now be written as

$$\pi(\theta | \mathbf{D}) = \frac{\pi(\mathbf{D} | \theta)g(\theta)}{\int_{\vartheta} \pi(\mathbf{D} | \vartheta)g(\vartheta)d\vartheta} \quad (1)$$

where the integral in the denominator is over the domain of

g , the prior distribution of the parameters θ , and is usually evaluated numerically by sampling the parameters over the prior space. MAP estimates the model parameters, $\hat{\theta}$ for which the posterior distribution has its' maximum (i.e. the mode of the distribution) and is written as

$$\hat{\theta}_{\text{MAX}} = \operatorname{argmax}_{\theta} \pi(\theta | \mathbf{D}) = \operatorname{argmax}_{\theta} \frac{\pi(\mathbf{D} | \theta)g(\theta)}{\int_{\vartheta} \pi(\mathbf{D} | \vartheta)g(\vartheta)d\vartheta} \quad (2)$$

Thus our problem is to find those parameters, θ , that maximise the likelihood $\pi(\mathbf{D} | \theta)$. For complex models we need to explore parameter space to find $\hat{\theta}$, which can be achieved by simulating this distribution using the Markov Chain Monte Carlo. Using this technique gives us a distribution for the estimates of $\hat{\theta}$ rather than the point estimates returned by ML.

Calculating the probability in (1) for most models is intractable and is often approximated using Monte Carlo methods which performs the integration by sampling θ from a distribution and 'saving' those samples that satisfy a condition. This (inefficient) Monte Carlo integration is improved by exploring the parameter space in a manner that hones in on the area of space that we want (i.e. gives those parameters that maximise a likelihood function or minimise a distance function in ABC). In many cases a Markov Chain is used to perform this exploration.

There have been many papers published on MCMC, for example Gilks et al. (1996); Brooks (1998); Berthelsen and Möller (2003); Doucet et al. (2000), which should be consulted for a more rigorous treatment as we will only give an algorithmic outline here.

Suppose we are able to calculate the Likelihood function, $\pi(\mathbf{D} | \theta)$, the steps to perform the MCMC algorithm are:

- (1) Select a starting point in the parameter space, i.e. draw θ from $\pi(\theta)$.
- (2) Calculate the likelihood for this θ . This is usually the most computationally intensive part of the algorithm.
- (3) Take a trial step by selecting a new set of parameters θ_{trial} from $q(\theta_{\text{trial}} | \theta_{\text{current}})$. There is no hard and fast rule about how to select these parameters, taking a large step means the parameter space is explored more quickly but not with any great accuracy, steps that are too small mean that the local area is explored in great detail but it takes longer to explore the whole space. In general selecting a trial step from a normal distribution makes sense where the standard deviation can be used to 'tune' the step size.
- (4) Compare the likelihood for this trial step to the previous step and accept the trial according to a rejection algorithm, if the trial is accepted the parameters are updated $\theta = \theta_{\text{trial}}$ and a new trial step

is sampled. The Metropolis–Hastings algorithm is commonly used to determine whether or not to accept the trial step. The basic algorithm is to accept trial according to a probability proportional to the difference of the likelihoods ($\mathcal{L}_{\text{trial}} - \mathcal{L}$). If $\mathcal{L}_{\text{trial}} \geq \mathcal{L}$ then the trial step is always accepted (thus always moving towards the areas of parameter space that maximise the likelihood), conversely if $\mathcal{L}_{\text{trial}} < \mathcal{L}$ the trial step has a high probability of being accepting if the trial likelihood is *close* to that of the previous step allowing a chance of ‘going downhill’. This means that the walk does not get stuck in a local maximum and thus guarantees that the global maximum will be found (but makes no prediction as to how long it will take to find).

- (5) Several such walks or chains are run, each with a different initial θ until each converge on the same region of the parameter space. This region defines the posterior distribution. The goal of any inference technique is to find this region and draw samples from it, the distribution of these sampled parameters make up the posterior distribution of the parameters. We often refer to *burn-in* when talking about MCMC, this is simply the process of removing those steps in the Markov Chain that are not in the

region of the posterior distribution/maximum likelihood.

MCMC will generate parameters while exploring parameter space in a manner that spends most time in the important regions. In the parlance of inference methods, the samples (parameters) mimic samples drawn from the target distribution (i.e. those parameters we are trying to find).

The efficiency of the MCMC is determined by how well the random walk (Markov Chain) explores the parameter space (how fast it can find the target area). If there are correlations between parameters in the model, these must be taken into account in constructing the trial steps. Failure to take the correlations between parameters into account will result in exploring an area of space that will not contribute to the posterior distribution. A novel method for constructing trial steps was proposed by Haario et al. (2001) and described how it can be applied to epidemiological data in O’Hare (2015).

4 Worked example

To demonstrate an application of the MAP technique to a model describing some spatiotemporal data, we will run a SIR model starting with a single infected animal using the locations and movements in the datasets given in S1 and use the technique outlined above to infer the parameter values in the model.

We model the epidemic as consisting of three distinct stages susceptible, infectious and removed. Infected animals can infect others in the same herd/farm at a rate β per time-step and infectious animals are removed at a rate σ . The method of removal is not important for this example but may be, for example, through culling detected infected animals. We allow for heterogeneity in the size of each farm by sampling the size from a Normal distribution with a mean of 60 and a standard deviation of 20, $N(60, 20)$. When moving animals between farms we allow them to potentially infect others on both farms in that time step. For computational efficiency, we create agents for the infected animals only, updating their location when moving between farms. The movement data describes the source and destination location and the date of the movement, the number of animals moved is sampled from $N(6, 4)$. The number of infected animals that is moved is sampled from a hypergeometric distribution.

We start with a single infected farm on a highly connected farm (farm 2 in the dataset) and solve the model using the Gillespie algorithm, moving animals between farms at each time step according to the movements in the supplementary information. We use parameter values of

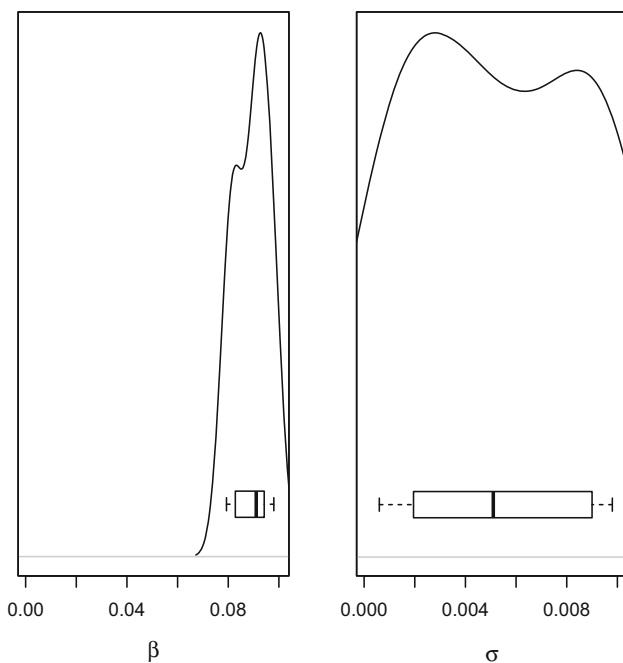


Fig. 7 Posterior distribution of the parameters β, σ in a SIR model using the fictional cattle trading network of Florin. A single simulation was run using the Gillespie algorithm, moving animals around the network according to the movements in Fig. 2. We seeded the outbreak with a single infected animal on farm 2 using $\beta = 0.09, \sigma = 0.007$

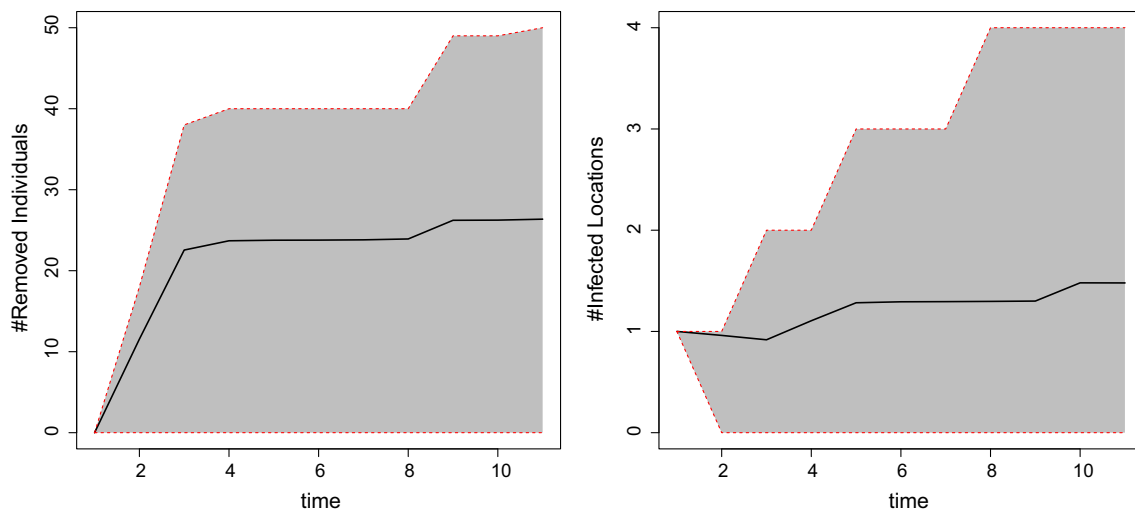


Fig. 8 Time series plot of the number of infected animals (*left*) and infected farms (*right*). The *black line* denotes the mean values and the *grey shaded areas* show the maximum and minimum values. We can

$\beta = 0.09$, $\sigma = 0.007$ to obtain 65 infectious animals and 35 removed (assuming, of course, that there is some physical mechanism to detect and record infectious and removed animals).

We write the likelihood for this model as

$$\mathcal{L} = \frac{n!}{\prod x_i!} \prod p_i^{x_i} \quad (3)$$

where n is the size of the infected population (the number of infectious and removed), x_i, p_i are the numbers in the infectious and removed classes and the probability of observing these numbers respectively. Using uniform priors $\beta = [0.001, 0.1]$, $\sigma = [0.0001, 0.001]$ and running 25 separate Markov Chains each starting at a random point in space each 1000 steps long and calculating the mean number of infectious and removed in 50 simulations at each step we estimate $\beta = 0.0910$ with a 95% credible interval of 0.0768, 0.101 and $\sigma = 0.0051$ with a credible interval of 0.0, 0.012 thus recovering the parameters $\theta = (\beta = 0.09, \sigma = 0.007)$, the kernel density estimates are shown in Fig. 7.

In recovering the posterior values for the parameters in our model we record the time series data of the numbers of infectious individuals and the number of infected farms over time as a measure of the likely number of cases we can expect from a similar outbreak (Fig. 8).

5 Conclusion

The vast amount of animal location, movement and test data that is collected during modern disease outbreaks is a valuable resource for mathematical epidemiologists. Analysing this data is not without difficulty due to the size and

see that in some cases that the disease dies out as both the minimums on the *left* and *right* are zero but overall the epidemic is increasing in size (we have chosen parameter values so that $R_0 > 1$

nature of the collected data but modern inference techniques and advances in pattern extraction in spatiotemporal datasets have aided the control of the spread of diseases. Ignoring temporal affects leads to both an overestimation of the predicted outbreak size and poorly designed control measures. Incorporating the dynamic nature of the network of animal movements can reveal important time windows that can be targeted when designing interventions.

In this paper we have outlined two broad techniques for extracting epidemiological information from mathematical models using these spatiotemporal data sets, giving a step-by-step approach to introduce the concepts and terminology involved. A realistic model using fictitious data demonstrated how the transmission parameters could be recovered (the code is available as one of the examples in the Broadwick framework). Approachable references to more advanced texts are given for the interested reader.

Acknowledgements The authors gratefully acknowledge funding and data from the Scottish Government as part of EPIC: Scotland's Centre of Expertise on Animal Disease Outbreaks

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. In:

- International AAAI conference on weblogs and social media. <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>
- Beaumont MA, Cornuet J-M, Marin J-M, Robert CP (2009) Adaptive approximate Bayesian computation. *Biometrika*, p 20252035. doi:10.1093/biomet/asp052
- Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* 162(4):2025–2035
- Berthelsen KK, Möller J (2003) Likelihood and non-parametric Bayesian MCMC inference for spatial point processes based on perfect simulation and path sampling. *Scand J Stat* 30:549564. doi:10.1111/1467-9469.00348
- Biek R, O'Hare A, Wright D, Mallon T, McCormick C, Orton RJ, McDowell S, Trewby H, Skuce RA, Kao RR (2012) Whole genome sequencing reveals local transmission patterns of *Mycobacterium bovis* in sympatric cattle and badger populations. *PLoS Pathog* 8:e1003008
- Brooks SP (1998) Markov chain Monte Carlo method and its application. *Statistician* 47:69–100
- Cattle Tracing System, Defra. <https://secure.services.defra.gov.uk/wps/portal/ctso>
- Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian computation (ABC) in practice. *Trends Ecol Evol* 25:410418. doi:10.1016/j.tree.2010.04.001
- Doucet A, Godsill S, Andrieu C (2000) On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat Comput* 10(3):197–208. doi:10.1023/A:1008935410038
- Dube C, Ribble C, Kelton D, McNab B (2009) A review of network analysis terminology and its application to foot-and-mouth disease modelling and policy development. *Transbound Emerg Dis* 56(3):73–85. doi:10.1111/j.1865-1682.2008.01064.x
- Estrada E (ed) (2010) *Network science : complexity in nature and technology*. Springer, New York. isbn:978-1-84996-395-4 <http://opac.inria.fr/record=b1132604>
- Gilks WR (ed) (1996) *Markov chain Monte Carlo in practice*. Chapman & Hall, London
- González-Parraa G, Arenasb AJ, Arandac DF, Segoviad L (2011) Modeling the epidemic waves of AH1N1/09 influenza around the world. *Spat Spat Temporal Epidemiol* 2(4):219–226. doi:10.1016/j.sste.2011.05.002
- Gumel AB (2009) Global dynamics of a two-strain avian influenza model. *Int J Comput Math* 86(1):85–108. doi:10.1080/00207160701769625
- Haario H, Saksman E, Tamminen J (2001) An adaptive Metropolis algorithm. *Bernoulli* 7:22342
- Hagberg AA, Schult DA, Swart PJ (2008) Exploring network structure, dynamics, and function using NetworkX In: Gäel Varoquaux, Travis Vaught, and Jarrod Millman (eds) *Proceedings of the 7th python in science conference (SciPy2008)*. <http://math.lanl.gov/~hagberg/Papers/hagberg-2008-exploring>
- Holme P (2013) Epidemiologically optimal static networks from temporal network data. *PLoS Comp Biol* 9(7):e1003142. doi:10.1371/journal.pcbi.1003142
- Hong YT, Paik B (2012) Inference model derivation with a pattern analysis for predicting the risk of microbial pollution in a sewer system. *Stoch Environ Res Risk Assess* 26:695. doi:10.1007/s00477-011-0538-9
- Kao RR, Green DM, Johnson J, Kiss IZ (2007) Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the UK *J Royal Soc. Interface* 4(16):907–916. doi:10.1098/rsif.2007.1129
- Keeling MJ, Woolhouse MEJ, Shaw DJ, Matthews L, Chase-Topping M, Haydon DT, Cornell SJ, Kappey J, Wilesmith J, Grenfell BT (2001) Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science* 294(5543):813–817. doi:10.1126/science.1065973
- Kiss IZ, Green DM, Kao RR (2006) The network of sheep movements within Great Britain: network properties and their implications for infectious disease spread. *J R Soc Interface* 3(10):669–677. doi:10.1098/rsif.2006.0129
- Lopes JS, Beaumont MA (2010) ABC: a useful Bayesian tool for the analysis of population data. *Infect Genet Evol* 10(6):825–832. doi:10.1016/j.meegid.2009.10.010
- Lowe R, Cazelles B, Paul R et al (2015) Quantifying the added value of climate information in a spatio-temporal dengue model. *Stoch Environ Res Risk Assess*, pp 1–12. doi:10.1007/s00477-015-1053-1
- Marjoram P, Molitor J, Plagnol V, Tavaré S (2003) Markov chain Monte Carlo without likelihoods. *Proc Natl Acad Sci USA* 100:1532415328. doi:10.1073/pnas.0306899100
- Martínez-López B, Perez AM, Sánchez-Vizcaíno JM (2009) Social network analysis. Review of general concepts and use in preventive veterinary medicine. *Transbound Emerg Dis* 56(4):109–120
- Moustakas A, Evans MR (2015) Coupling models of cattle and farms with models of badgers for predicting the dynamics of bovine tuberculosis (TB). *Stoch Environ Res Risk Assess* 29:623. doi:10.1007/s00477-014-1016-y
- Moustakas A, Evans MR (2016) Regional and temporal characteristics of bovine tuberculosis of cattle in Great Britain. *Stoch Environ Res Risk Assess* 30:989. doi:10.1007/s00477-015-1140-3
- Nöremark M, Widgren S (2014) EpiContactTrace: an R-package for contact tracing during livestock disease outbreaks and for risk-based surveillance. *BMC Vet Res* 10:71. doi:10.1186/1746-6148-10-71
- O'Hare A (2015) Inference in high dimensional parameter space. *J Comp Biol* 22(11):997–1004. doi:10.1089/cmb.2015.0086
- O'Hare A, Lycett SJ, Doherty T, Salvador LCM, Kao RR (2016) Broadwick: a framework for computational epidemiology. *BMC Bioinform* 17(1):1–5. doi:10.1186/s12859-016-0903-2
- Reiczigel J, Brugger K, Rubel F et al (2010) Bayesian analysis of a dynamical model for the spread of the Usutu virus. *Stoch Environ Res Risk Assess* 24:455. doi:10.1007/s00477-009-0333-z
- ScotEID—scottish EID livestock traceability research. <https://www.scoteid.com>
- Sisson SA, Fan Y, Tanaka MM (2007) Sequential Monte Carlo without likelihoods. *Proc Natl Acad Sci USA* 104:17601765. doi:10.1073/pnas.0607208104
- Stärk KDC, Regula G, Hernandez J, Knopf L, Fuchs K, Morris RS, Davies P (2006) Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches. *BMC Health Serv Res* 6(1):1–8. doi:10.1186/1472-6963-6-20
- Tanaka MM, Francis AR, Luciani F, Sisson SA (2006) Using approximate bayesian computation to estimate tuberculosis transmission parameters from genotype data. *Genetics* 3:1511–1520. doi:10.1534/genetics.106.055574
- Turner BM, Van Zandt T (2012) A tutorial on approximate Bayesian computation. *J Math Psychol* 56(2):69–85. doi:10.1016/j.jmp.2012.02.005