Title

Rewiring cattle movements to limit infection spread.

Authors

Thibaut Morel-Journel[a*], Pauline Ezanno[a], Elisabeta Vergu[b]

Affiliations

a: INRAE, Oniris, BIOEPAR, 44300, Nantes, France

b: INRAE, Université Paris-Saclay, MaIAGE, 78350, Jouy-en-Josas, France

Corresponding author

*Thibaut Morel-Journel: thibaut.moreljournel@gmail.com

Abstract

The cattle tracing databases set up over the past decades in Europe have become major resources for representing demographic processes of livestock and assessing potential risk of infections spreading by trade. The herds registered in these databases are parts of a network of commercial movements, which can be altered to lower the risk of disease transmission. In this study, we developed an algorithm aimed at reducing the number of infected animals and herds, by rewiring specific movements responsible for trade flows from high- to low-prevalence herds. The algorithm was coupled with a generic computational model describing infection spread within and between herds, based on data extracted from the French cattle movement tracing database (BDNI). This model was used to simulate a wide array of infections, with either a recent outbreak (epidemic) or an outbreak that occurred five years earlier (endemic), on which the performances of the rewiring algorithm were explored. Results highlighted the effectiveness of rewiring in containing infections to a limited number of herds for all scenarios, but especially if the outbreak was recent and if the estimation of disease prevalence was frequent. Further analysis revealed that the key parameters of the algorithm affecting infection outcome varied with the infection parameters. Allowing any animal movement from high to low-prevalence herds reduced the effectiveness of the algorithm in epidemic settings, while frequent and fine-grained prevalence assessments improved the impact of the algorithm in endemic settings. Overall, our approach, which focuses on a few commercial movements, has led to substantial improvements in the control of a targeted disease, although changes in the network

structure should be monitored for potential vulnerabilities to other diseases. Due to its generality, the developed rewiring algorithm could be applied to any network of controlled individual movements liable to spread disease.

Keywords

Control strategy; Epidemiology; Data-based; Network; Stochastic model

Abbreviation

BDNI: *Base de données nationale d'identification animale*

Introduction

Following bovine spongiform encephalopathy and classical swine fever epidemics in the 1990s, the European Union initiated the mandatory identification and registration of cattle in Europe (EU, 2000). This decision led to the creation of national identification databases, such as the cattle tracing system in the United Kingdom (Kao et al., 2006, Vernon, 2011), the French national bovine identification database (BDNI) (Rautureau et al., 2011, Dutta et al., 2014), the Italian national bovine database (Natale et al., 2009, Bajardi et al., 2011) and the database of the Swedish board of agriculture (Nöremark et al., 2009, 2011). These animal tracing systems have enabled the monitoring of infectious livestock diseases and the development of strategies to prevent their spread (Gilbert et al., 2005, Moslonka-Lefebvre et al., 2016, Beaunée et al., 2017), since animal trade is a major transmission pathway between herds. Indeed, commercial exchanges are not only recorded comprehensively, but also controlled by farmers, unlike animal mobility in the wild. These databases, whose reliability has increased over time since their creation (Green and Kao, 2007), are therefore powerful tools for simulating infectious diseases in cattle (Ezanno et al., 2020) and assessing the impact of livestock movements on epidemics (Ezanno et al., 2021).

The information provided by these commercial animal movements can be used as a basis for representing comprehensively the demographic processes and trades between cattle farms located in a given region, using a metapopulation framework (Liu et al., 2007, Courcoul and Ezanno, 2010). To this end, disease transmission between individuals within a defined set of herds can be modelled, by combining an epidemiological model with existing data on births, deaths and movements. This type of models accounts at least for two ways of spreading the infection: by contact within a herd, or by actually moving animals between herds. This is for instance the case for paratuberculosis, a cattle disease mainly spread by trade (Beaunée et al., 2015, Biemans et al., 2021). Manipulating the structure of cattle movement is expected to have a direct impact on the latter and an indirect impact on the former.

The structure of these trade movements can be understood through the prism of graph theory: herds are the vertices of a commercial exchange network, whose edges are the movements of livestock (Dubé et al., 2009). Thus, each herd can be characterized using graph metrics, e.g. its in- and out-degree, i.e. the number of herds it has respectively bought animals from and sold animals to. Network-based control strategies then aim to modify the structure of the network to reduce infection risks. Removing vertices (Rautureau et al., 2011, Büttner et al., 2013) or edges (Yang et al., 2013, Green et al., 2009) through trade ban or slaughtering is often considered to slow down epidemics. In a context of cattle exchange however, preventing farmers from buying or selling livestock entails high economic costs. Therefore, this strategy cannot be used routinely, and is likely better suited to the management of regulated diseases, the consequences of which are also very costly.

Edge rewiring is a less radical approach able to balance the trade-off between health risks and economic costs. This method corresponds to the modification of one or both vertices that an edge connects (Gross et al., 2006, Piankoranee and Limkumnerd, 2020, Britton et al., 2016, Ball and Britton, 2020). Although most of the theoretical literature on the subject rather considers rewiring in the context of human contact networks, it has also been used to study epidemic spread in cattle movement networks (Gates and Woolhouse, 2015, Mohr et al., 2018, Ezanno et al., 2021). For instance, Gates and Woolhouse (2015) present a rewiring method that creates an entirely new movement network disconnecting large buyers from large sellers, while retaining the total number of animals bought or sold by each herd. This method requires information at the network level, the criteria used being the distributions of in- and out-degrees of all vertices. Global-level information is also generally required for most rewiring methods in contact networks, although Piankoranee and Limkumnerd (2020) proposed a method based on local information. In their study, rewiring is decided at the vertex level, according to its status and those of its direct neighbours. Controlling cattle movements depending on the sanitary status of their origin has been proposed in previous studies, e.g. by Hidano et al. (2016). Their study presents different scenarios regarding farmers' practices, especially their tendency to avoid buying cattle from regions with a higher incidence of bovine tuberculosis. The approach presented here is similar, albeit at a finer grain: preventing farmers from buying cattle from herds with a higher prevalence of the target disease.

This study presents a new rewiring method to reduce the spread of infections in a cattle movement network. To do this, we developed a rewiring algorithm in conjunction with a simulation model. The simulation model coupled demography based on a dataset from the French cattle tracing system (BDNI) with a SIRS epidemiological model. The use of such a generalist epidemiological model made it possible to simulate a wide range of epidemiological settings, only by varying its parameters. The algorithm worked by preventing movements from high- to low-prevalence herds, while maintaining the number of animals bought and sold by each farm. It was based on an edge-level criterion: the estimated difference in prevalence between the herd of origin and the herd of destination of the movement considered. The

following presents the cattle movement network concerned by this study, as well as the model and the algorithm. Then, we consider various outputs of runs of the model with or without rewiring, concerning the the functioning of the algorithm itself, its impact on the propagation of the infection, and its impact on the structure of the network of commercial movements.

Data and methods

Cattle movement network

The dataset used for this study corresponded to an extraction from the French national bovine identification database (BDNI). It concerned all cattle herds in Brittany (a French region) that sold or bought at least one animal during the year 2014. This set of 21,548 herds is referred to as the 'metapopulation' thereafter. The database also included commercial exchanges as well as demographic events, all of which were referred to as 'movements' hereafter. Three types of commercial exchanges were considered: (i) 'internal movements' had an origin and a destination among the herds in the dataset, (ii) 'imports' had only a destination in the dataset and (iii) 'exports' had only an origin in the dataset. They represented respectively 64%, 16% and 20% of the commercial exchanges involving at least one herd of the dataset. Each commercial exchange of animals was assumed to take place directly from one herd to another, neglecting intermediaries. This means that markets and sorting centres were not considered for this study. They differ from herds in that they tend to concentrate a large number of animals, but for a limited period of time (less than a day for markets, a few days for sorting centres). Two types of demographic events were considered: (iv) births had only a destination, corresponding to the herd where the animal was born, and (v) deaths had only an origin, corresponding to the herd where the animal died.

The dataset was represented as a network with herds and internal movements corresponding to the vertices and edges, respectively. This network was (i) dynamic, i.e. movements were characterized by the date at which they occurred, (ii) weighted, i.e. a single edge represented the set of all movements from herd A to herd B, with a weight corresponding to the number of movements, and (iii) directed, i.e. movements from herd A to herd B were accounted for separately from movements from herd B to herd A. The network therefore included 21,548 vertices and 100,088 edges. The total number of internal movements over 2014 was 206,640, thus the average edge weight was 2.06.

Epidemiological model: within and between-herd dynamics and infection settings

The model developed aimed to simulate pathogen transmission within herds, and infection spread between herds through cattle movements. A full description of the model is available in Supplementary material 1. The model is stochastic in discrete time – each time-step corresponding to a day of 2014 – and in discrete space – by integrating the network of herds and movements described above. Commercial exchanges and

demography were data-based: movement m was characterized by its origin O_m, its destination D_m, its date according to the dataset T^*_m and the date at which it was simulated T_m. By default, movements were simulated according to the dataset, i.e. $T_m = T^*_m$. Within-herd dynamics were based on a SIRS model with three parameters: the infection rate β, the recovery rate γ – therefore the average infection duration was $1/\gamma$ – and the rate of return to susceptibility δ. At each time-step t, herd h was characterized by its number of susceptible, infected and recovered individuals, noted respectively $S_h(t)$, $I_h(t)$ and $R_h(t)$. The total herd size $N_h(t)$ was defined as the sum of these three values and infection prevalence as $P_h(t) = I_h(t)/N_h(t)$.

Each simulated infection began with an initial outbreak in a metapopulation without infection, i.e. with only susceptible individuals. At $t = t_I$, the date of the outbreak, 10% of all herds in the metapopulation were infected, by replacing 1 susceptible individual with 1 infected individual in each of the herds. The probability of a herd being part of this 10% was proportional to the number of imports in the herd according to the 2014 dataset. The rationale was that herds receiving the most individuals from herds outside of the metapopulation were the most likely to introduce a new infection.

Two types of infections were considered for the study: epidemic and endemic. An infection was defined as 'epidemic' if it started at the outbreak, i.e. if $t_0 = t_I$. The initial state of the infection was then as described above. An infection was defined as 'endemic' if its start date was five years after the outbreak, i.e. $t_0 = t_I + 1825$ days. The initial state of infection was then the result of a five-year infection, simulated using the same epidemiological model and an extraction from the BDNI over Brittany between 01/01/2009 and 31/12/2013. Endemic simulations for which the infection went extinct before t_0 were discarded, so that only initial states that were not disease-free were considered.

Developed rewiring algorithm

The algorithm developed aimed at preventing movements of cattle from high-prevalence herds to low-prevalence herds. Its functioning was based on prevalence classes, numbered from 1 to c. Class i corresponded to prevalences between b_i and b_{i+1}, with the lowest boundary $b_1 = 0$ and the highest boundary $b_{c+1} = 1$. Each herd was assigned a 'real' and an 'observed' prevalence status, corresponding to one of these classes. The real prevalence status $V^r_h(t)$ of herd h at time t was equal to the prevalence class i if $P_h(t) \in [b_i; b_{i+1}[$, with $V^r_h(t) = c$ if $P_h(t) = 1$. The observed prevalence status was updated periodically, every q time-step. If the real status of herd h was observed at time t_{obs}, then its observed status $V^o_h(t)$ remained the same for q time-steps, i.e. $V^o_h(t) = V^r_h(t_{obs})$ for any $t_{obs} \in [t; t + q[$. No additional errors on the status were assumed, so that $V^r_h(t)$ always corresponded to the correct prevalence class.

Sequential rewiring

The algorithm was executed at each time-step, before the simulation of the epidemiological model. Supplementary material 2 describes its functioning for a single time-step in pseudo-code. First, the algorithm defined all possible quadruplets of prevalence classes $\{c_{OR}, c_{DR}, c_{ON}, c_{DN}\}$, so that $1 < c_{ON} \leq c_{DR} < c_{OR} \leq c_{DN} < c$. The quadruplets were arranged primarily in ascending order of c_{DR}, secondarily in descending order of c_{OR}, thirdly in ascending order of c_{ON} and fourthly in descending order of c_{DN}. The algorithm followed this order, which ensured that no potential rewiring was missed. For each quadruplet, the algorithm selected two lists of internal movements set to occur at time t, based on the status of their origin and destination. A movement mR was in the first set if $V^o_{O_{mR}}(t) = c_{OR}$ and $V^o_{D_{mR}}(t) = c_{DR}$. This movement was considered 'at risk' because, as $c_{OR} > c_{DR}$, the observed prevalence status of their origin was greater than that of their destination. A movement mN was in the second set if $V^o_{O_{mN}}(t) = c_{ON}$ and $V^o_{D_{mN}}(t) = c_{DN}$. This movement was considered 'normal', i.e. not at risk, since $c_{ON} < c_{DN}$. For k the size of the shortest of the two movement sets, the algorithm permuted the origins of the first k movements of each set, so that $V^o_{O_{mR}}(t) = c_{ON}$ and $V^o_{O_{mN}}(t) = c_{OR}$. Since $c_{OR} \leq c_{DN}$ and $c_{ON} \leq c_{DR}$, neither movement was at risk after the permutation.

Once permutations were performed for each quadruplet, the management of remaining movements at risk, i.e. whose origin had a greater observed prevalence status than their destination, depended on two parameters: the maximal delay Δ_{MAX} and the prohibition of movements at risk. Remaining movement m was postponed, i.e. T_m was increased by 1, if $T_m - T^*_m < \Delta_{MAX}$. Otherwise, it was tagged as 'problematic'. If movements at risk were not completely prohibited, problematic movement m would not be modified further. If they were, m would be replaced by an import with D_m as its destination and by an export with O_m as its origin. Overall, four algorithm parameters had to be defined *a priori*: the number of classes c, the update period q, the maximal delay Δ_{MAX} and whether the movements at risk were prohibited.

Simulations

Simulations were performed on the dataset between 01/01/2014 (defined as $t = 0$) and 01/01/2015 ($t = 365$). Different epidemiological settings were explored by manipulating the SIRS model parameters (β, γ and δ) and infection type (epidemic or endemic). Two clustering analyses were performed on the preliminary simulations to define six epidemiological settings (Supplementary material 3): weak, moderate and strong epidemic settings and weak, moderate and strong endemic settings (Fig. S2).

The effectiveness of the algorithm was tested by running simulations with $3 \times 3 \times 3 \times 2$ combinations of the algorithm parameters, respectively (i) the number of prevalence classes c (2, 3 or 4 classes), (ii) the update period q (1, 28 or 91 days), the maximum delay Δ_{MAX} (1, 3 or 7 days) and (iv) the prohibition of movements at risk (yes or no). Each combination, as well as a control without rewiring, were simulated

Outcomes related to	Notation	Description
Algorithm	$n_{rew}(t)$	Number of movements rewired at time t
	$n_{del}(t)$	Number of delayed movements at time t
	$n_{prob}(t)$	Number of problematic movements at time t
	$n_{risk}(t)$	Number of movements at risk at time t
	$n_{err}(t)$	Number of movements undetected as at risk at time t
Infection	n_{inf}	Number of herd infections
	n_{ext}	Number of herds in which the infection goes extinct
	a_{dur}	Average duration of infection
	$c_{inc}(t)$	Cumulated incidence at time t
	$n_{herd}(t)$	Number of infected herds at time t
	$n_{ind}(t)$	Number of infected individuals in the metapopulation at time t
	$a_{prev}(t)$	Average prevalence in the infected herds at time t
Network	n_{SCC}	Number of strongly connected components
	max_{SCC}	Size of the largest strongly connected component
	ind_h	In-degree of herd h
	$outd_h$	Out-degree of herd h

Table 1: List of the outcomes computed from the simulations. The infection-related outcomes were computed for each simulation separately. The algorithm and network-related ones were computed for each simulation with the algorithm.

100 times for each of the six epidemiological settings.

Preliminary simulations were also carried out for each epidemiological setting between 01/01/2009 ($t = -1825$) and 31/12/2013 ($t = -1$), with an initial outbreak at $t_I = -1825$. On the one hand, the number of susceptible, infected and recovered individuals of each herd at $t = -1$ were used as the starting numbers for the endemic simulations (starting at $t = 0$). On the other hand, the boundaries of the prevalence classes b_i used by the algorithm were set as quantiles of the distribution of prevalence values. If fewer than $1/c$ herds had a null prevalence, b_i was the $((i-1)/c)^{th}$ quantile of the distribution. If it was greater than $1/c$, $b_1 = b_2 = 0$ and b_i was the $((i-2)/(c-1))^{th}$ quantile of the distribution.

Outcomes and analyses of numerical explorations

The simulations outcomes are listed in Table 1. They were related either (i) to the functioning of the algorithm, (ii) to the infection or (iii) to the network of internal movements modified by the algorithm.

The algorithm-related outcomes $n_{rew}(t)$, $n_{del}(t)$ and $n_{prob}(t)$ were computed each time-step after rewiring, while $n_{risk}(t)$ and $n_{err}(t)$ were computed before. These latter outcomes were computed by using the real prevalence status of the herds, rather than the observed ones. A movement m was included in $n_{risk}(t)$ if $V^r_{O_m}(t) > V^r_{D_m}(t)$, and also included in $n_{err}(t)$ if $V^o_{O_m}(t) \leq V^o_{D_m}(t)$ at the same time. The proportion of undetected movements at risk was computed on a weekly basis, to account for intra-week variability in the number of livestock movements. Over week w, this proportion $p_{err}(w)$ was :

$$p_{err}(w) = \frac{\sum_{t=7(w-1)+1}^{7w} n_{err}(t)}{\sum_{t=7(w-1)+1}^{7w} n_{risk}(t)}.$$

The Spearman's correlation coefficient ρ between $p_{err}(w)$ and the number of weeks since last update

(from 1 to 4 weeks if $q = 28$ days, from 1 to 13 weeks if $q = 91$ days) was also computed to assess the relationship between errors in herd prevalence status and time. The Spearman's coefficient was preferred because it did not assume any particular distribution of the involved variables.

The impact of the algorithm on the infection dynamic was estimated through $c_{inc}(t)$, i.e. the cumulated number of herds newly infected over the simulation. The variations in $n_{herd}(t)$ and $n_{ind}(t)$ over time are also presented in Supplementary material 4. Besides, the overall impact of the algorithm on the infection was assessed using a global multivariate sensitivity analysis, following Lamboni et al. (2011) and using the *multisensi* package of the *R* software (Bidot et al., 2018). This method allows to perform a sensitivity analysis on a multivariate output. For this analysis, twelve variables were derived from the infection-related outcomes. The three outcomes computed once per simulation n_{inf}, n_{ext} and a_{dur} were used as such. In addition, the maximum, minimum and final values over the whole period simulated (respectively noted $max(u(t))$, $min(u(t))$ and $u(365)$ for outcome $u(t)$) of $n_{herd}(t)$, $n_{ind}(t)$ and $a_{prev}(t)$ were also computed. The analysis included a principal component analysis (PCA) on the scaled variables, which were used as the multivariate output for the analysis. The analysis allowed to compute two generalised sensitivity indices (GSI) for each algorithm parameter, which were weighted means of the sensitivity indices over all the dimensions of the PCA: the total index (tGSI) including interactions with other parameters, and the first-order index (mGSI), not including them. The first principal component of the PCA was also used to assess the distribution of the simulations depending on the algorithm parameters.

The network-related outcomes were based on an static view of the network aggregating all the internal movements performed during the simulation, from $t = 0$ to $t = 365$. Therefore, they took into account the rewiring performed by the algorithm, and the potential removal of problematic movements if movements at risks were completely prohibited. The outcomes recorded for the modified networks were compared to the same metrics for the original network defined by the 2014 dataset. The strongly connected components – from which n_{SCC} and max_{SCC} were computed – corresponded to groups of vertices linked to each other by a directed path. The percentiles of the distributions of ind_h and $outd_h$ of all herds in the static network were used to assess the in-degree and out-degree distributions, respectively.

Results

Outcomes related to the algorithm

The number of movements rewired varied greatly depending on the date of the outbreak. It was negligible in the epidemic settings, with 80% of simulations with a total of rewired movements between 192 (fewer than 0.1% of all movements) and 2250 (1.1%). However, it was larger in the endemic settings, with 80% of simulations with between 17,344 (8.4% of all movements) and 33,640 (16.3%) movements rewired.

Besides, increasing the value of Δ_{MAX} logically increased the number of delayed movements (which was 0 by definition for $\Delta_{MAX} = 0$) and decreased the number of problematic movements. In the endemic settings, the problematic movements represented a small proportion of the movements detected as high risk (median: 5.4%, 9^{th} decile: 17.4%). In the epidemic settings however, they represented a larger part (median: 14.3%, 9^{th} decile: 59.7%), although their absolute numbers remained low (median: 129, 9^{th} decile: 651). Because of the overwhelming number of initially non-infected herds in these simulations, the movements at risk were likely more difficult to rewire, and thus more likely to be tagged as problematic by the algorithm.

Increasing the herd status update period q was not associated with a decrease in the number of rewiring events (Fig. 1A, 1B). The value of q was even rather positively correlated with the number of rewiring events in epidemic settings. This suggests that the algorithm performed more erroneous rewiring as q increased. This was confirmed by the distributions of Spearman's correlation coefficient between $p_{err}(w)$ and the number of weeks since last update ρ with $q = 91$ days (Fig. 1D), in endemic settings (80% of values of ρ between -0.01 and 0.50) and in endemic settings (80% of values of ρ between 0.39 and 0.75). This was also somewhat the case with $q = 28$ days (1C), although the correlation were weaker, in endemic (80% of values of values between -0.09 and 0.79) as well as in epidemic settings (80% of values of values between -0.05 and 0.34).

The average proportions of undetected movements at risk $p_{err}(w)$ all tended to increase with the

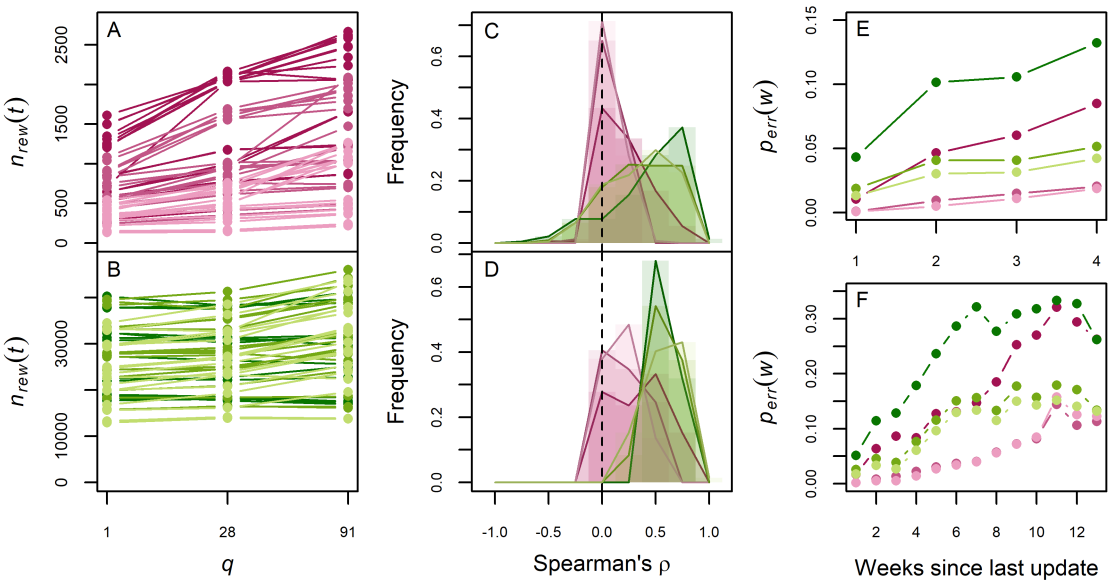

Figure 1: Impact of the update period q on the undetected movements at risk, in epidemic (magenta) or endemic settings (green), weak (light), moderate (medium) or strong (dark). First column: total number of rewiring events as a function of the update frequency q, averaged over all simulations for a same algorithm parameter combination, in epidemic (A) and endemic settings (B). Second column: distribution of Spearman's correlation coefficients (ρ), with $q = 28$ days (C) and $q = 91$ days (D). Third column: average proportion of undetected movements at risk $p_{err}(w)$ as a function of the number of weeks since the last update, with $q = 28$ days (E) and $q = 91$ days (F).

number of weeks since the last update w (Fig. 1E, 1F). This increase was systematically greater for the largest value of q , up to $p_{err}(w) = 0.3$. However, they also appeared to have reached a plateau after 10 weeks. This suggests that a further increase in the update period q would not strongly increase the proportion of undetected movements at risk. As for Spearman's correlation coefficient ρ, the increase was greater in endemic settings than in epidemic settings.

Outcomes related to the infection

Comparison of the results with and without rewiring showed the overall effectiveness of the algorithm in containing the infection (Fig. 2). Regardless of the epidemiological setting and the combination of parameters considered, $c_{inc}(t)$ remained systematically lower after rewiring. The algorithm was particularly effective in weak and moderate epidemic settings, where very few herds were infected during the year. In other epidemiological settings, the impact of the algorithm varied more strongly depending on the scenario considered. Results for $n_{herd}(t)$ and $n_{ind}(t)$ are presented in Supplementary material 4. The algorithm prevented the increase in the number of infected herds in the epidemic setting and even decreased their number in an endemic setting (Fig. S3). However, it had little impact on the number of infected individuals (Fig. S4).

The sensitivity analysis showed differences in the relative importance of the algorithm parameters on the reduction of the infection (Fig. 3). Three different patterns of sensitivity to the algorithm parameters were observed. Firstly, simulations in weak and moderate epidemic settings exhibited an overwhelming

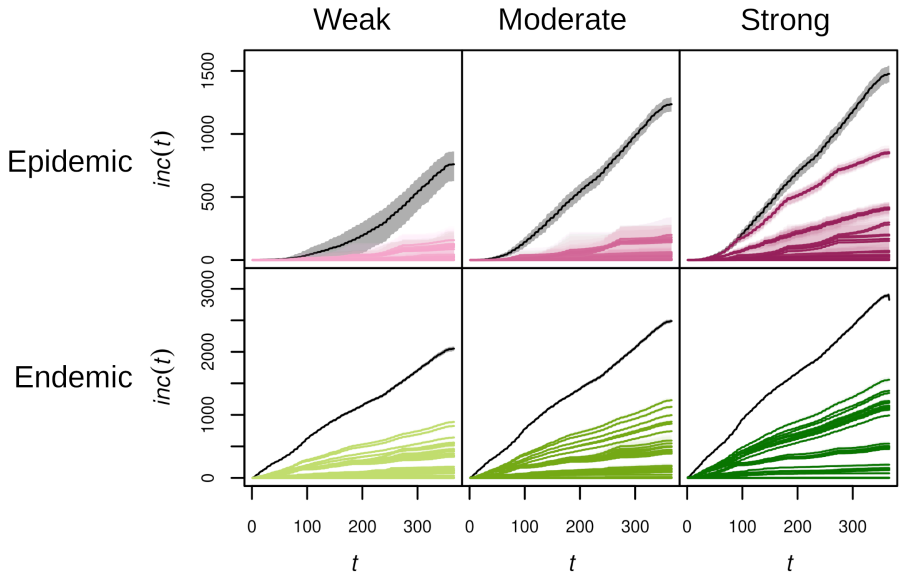

Figure 2: Cumulated incidence $c_{inc}(t)$, in number of herd infections, as a function of time (t, in days), for simulations with (colour) or without rewiring (black), in epidemic (1^{st} row, magenta) or endemic settings (2^{nd} row, green), weak (1^{st} column, light), moderate (2^{nd} column, medium) and strong (3^{nd} column, dark). Each combination of algorithm parameters is represented by its mean over the repetitions (solid line) and an interval of 80% of simulations (envelope).

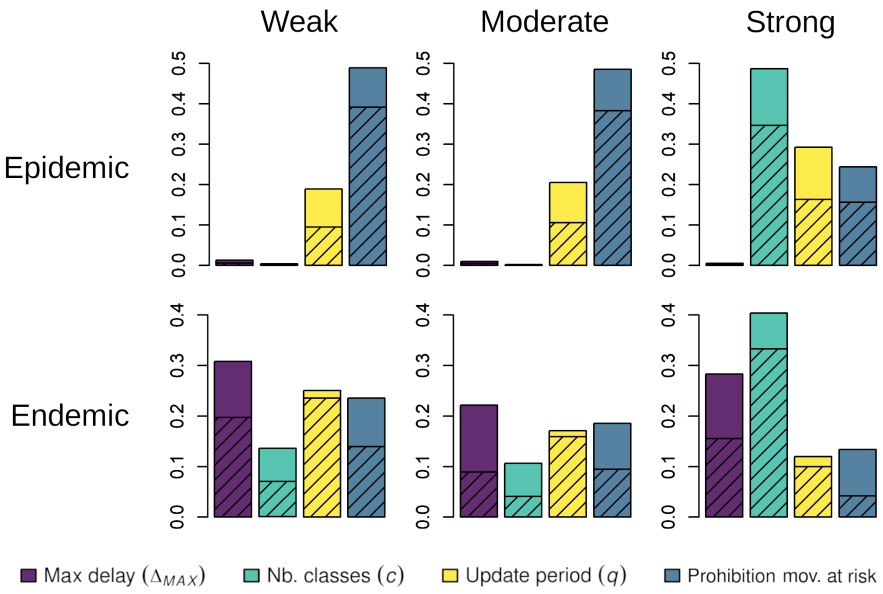

Figure 3: Generalised sensitivity indices (GSI) of the maximum delay Δ_{MAX} (purple) the number of prevalence classes c (cyan), the update period q (yellow) and the prohibition of movements at risk (blue), in epidemic (1^{st} row) or endemic settings (2^{nd} row), weak (1^{st} column), moderate (2^{nd} column) and strong (3^{rd} column). The total indices (tGSI) are in solid colour and the first-order indices (mGSI) are hatched.

sensitivity to the prohibition of movements at risk. Secondly, those in strong epidemic or endemic settings exhibited a strong sensitivity to the number of prevalence classes c. Finally, those in weak and moderate endemic settings exhibited a more balanced sensitivity to all parameters, with a substantial difference between total and first-order indices for the maximum delay Δ_{MAX}, the number of classes and the prohibition of movements at risk. These differences suggest an interaction between the three algorithm parameters. Besides, simulations for every epidemiological setting were somewhat sensitive to the update period q.

The PCA performed as a first step of the sensitivity analysis was used to explore further the way algorithm parameters impacted the infection-related outputs. Supplementary material 5 shows that the first principal component of the PCA was globally positively correlated with outputs describing the extent of the infection. The distributions of simulations along this first principal component therefore provided information about the way algorithm parameter values affected the extent of the infection. Supplementary material 6 presents these distributions for every epidemiological setting and every algorithm parameter, while Fig. 4 displays some of the most relevant distributions. Fig. 4A shows that, in the weak epidemic setting, simulations in which movements at risk were prohibited almost always scored lower on the first principal component than those in which they were not. The distribution was similar in thre moderate epidemic setting (Fig. S6), which had similar sensitivity indices (Fig. 3). Interestingly, distributions of simulations in strong epidemic or endemic settings showed that those with $c = 2$ scored higher on their respective first component, while those with $c = 3$ and $c = 4$ were not different (Fig. 4B, 4F). A

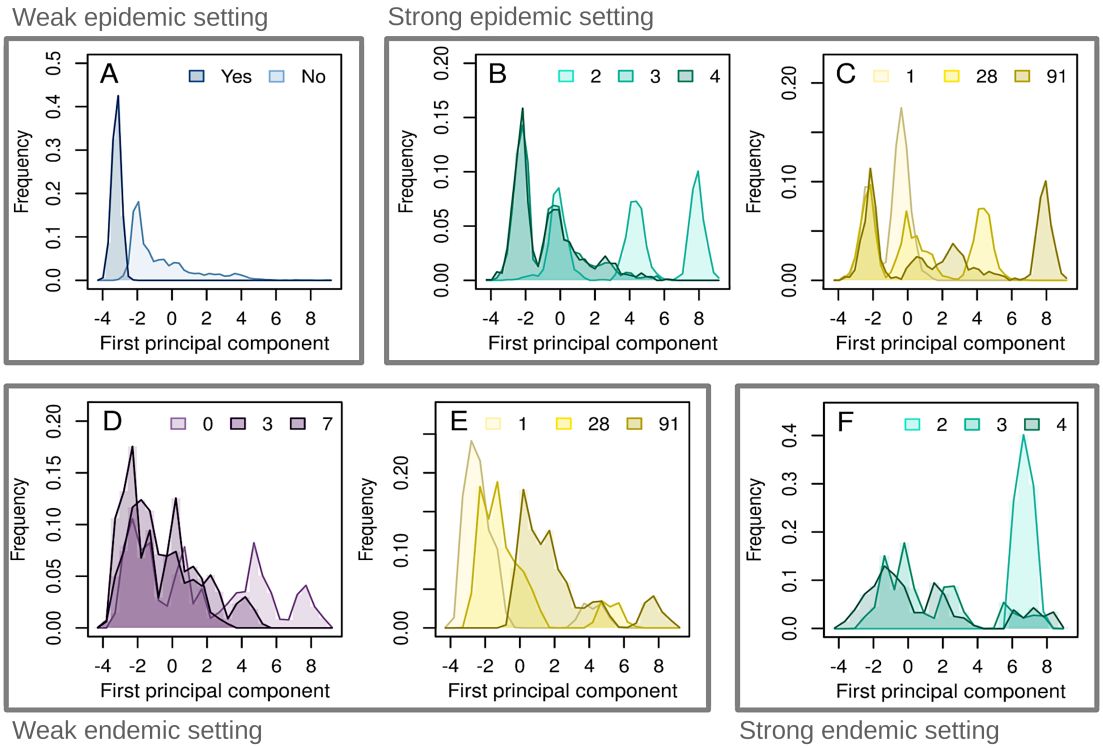

Figure 4: Distribution of the simulations on the first component of the PCA performed as a first step of the sensitivity analysis, in the weak epidemic setting (A), the strong epidemic setting (B, C), the weak endemic setting (D, E) and the strong endemic setting (F). The outputs are divided by maximum delay (purple, D), management of problematic movements (blue, A), number of prevalence classes (cyan, B and F) and herd status update period (yellow, C and E).

similar pattern was observed with the maximum delay in the weak endemic setting: only simulations with $\Delta_{MAX} = 0$ scored higher on the first principal component (Fig. 4D). In the strong epidemic setting, the two high-scoring peaks in the distribution according to c (Fig. 4B) corresponded to the simulations with $q = 28$ and $q = 91$ (Fig. 4C), highlighting an interplay between the number of classes c and the update period q. No interplay between Δ_{MAX} and q was visible in the weak endemic setting, although Fig. 4E showed that the score of simulations on the first principal component was positively correlated with q. Distributions in the moderate endemic setting were similar to those in the weak endemic setting (Fig. S6).

Outcomes related to the movement network

In endemic settings, rewiring movements increased the in- and out-degrees of the herds, i.e. the number of different herds they were connected to (see Supplementary material 7). The increase was small but systematic, for every algorithm parameter value (Fig. S7). In addition, the algorithm also affected the strongly connected components of the network in endemic settings. On the one hand, the algorithm reduced their number, all the more that the infection was strong (Fig. 5). On the other hand, the size of the largest strongly connected component was increased in most, but not all simulations (64%, 67% and

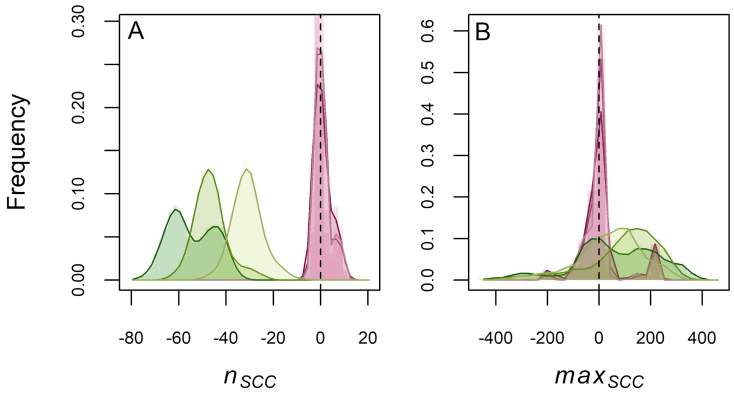

Figure 5: Distributions of the differences in number of strongly connected components (A, n_{SCC}) and in size of the largest strongly connected component B, max_{SCC}) between rewired networks and the original one from the dataset, for epidemic (magenta) and endemic (green) settings, weak(light), moderate (medium) and strong (dark).

80% of simulations in low, moderate and high endemic settings, respectively). It should be noted that the lesser impact of the algorithm on the network in epidemic settings can be explained by a number of rewiring events 25 times smaller on average than in endemic settings.

Discussion

This study aimed at developing an algorithm preventing movements at risk in order to reduce the extent of infections. Our results show that the rewiring algorithm we developed provided an improvement regardless of the disease parameters (infection rate β, recovery rate γ or rate of return to susceptibility δ), although differences were observed in its effectiveness between epidemiological settings. Indeed, infections in weak or moderate epidemic settings were almost completely prevented for each combination of parameters of the algorithm. In the other epidemiological settings, the impact of the algorithm was still significant, although weaker for some parameter combinations. However, the decrease in the number of infected herds was not necessarily coupled with a decrease in the number of infected individuals. This result highlights the tendency of the algorithm to concentrate infected individuals in the already infected herds. The algorithm therefore performed a trade-off that was beneficial to the metapopulation as a whole – with fewer infected herds – but detrimental to the smaller number of already infected herds, in such a situation where movement rewiring is not combined with complementary on-farm measures to reduce within-herd infection prevalence.

Sensitivity analysis on the infection-related outcomes revealed that all four parameters of the algorithm were important, but not for all epidemiological settings. Infections that were neither highly virulent nor already well established remained fully under control as long as movements at risk were completely prohibited, even if they could not be rewired. How these movements were managed seemed less important in other cases, in which infections could not be fully contained. The most virulent infections were strongly

influenced by the number of prevalence classes defined, which had to be greater than two for the algorithm to be effective. Because of the way boundaries between prevalence classes were computed, increasing the number of classes tended to limit the lowest prevalence class to (completely or almost) disease-free herds. Defining only two prevalence classes likely mixed those herds with somewhat infected ones, thus preventing the algorithm from effectively effectively protecting infection-free herds and thereby limiting the spread of the infection. Other endemic infections rather depended on an interaction between the parameters, but they were especially impacted by the frequency at which the prevalence status of herds were updated, as were simulations with a strong epidemic setting.

The update period had a significant impact on the behaviour of the algorithm, making it more error-prone for high values of q. Indeed, the proportion of undetected movements at risk increased with the time since the last update, at least up to ten weeks. This result, together with the impact of q on infection-related outcomes, indicates that any increase in the update frequency of the status of herds should improve the effectiveness of the algorithm. This was not the case for other algorithm parameters. For instance, the sensitivity analysis showed that when Δ_{MAX} played a role, reporting movements over only a few days – a rather minor constraint – was sufficient to improve the effectiveness of the algorithm. Similarly, considering only three prevalence classes reduced the extent of infections in strong epidemic or endemic settings. Conversely, increasing the frequency to daily updates was necessary for maximum algorithm effectiveness, but would likely entail significant costs. Bulk milk-based sampling systems could be used for some diseases in cattle (e.g. Garoussi et al., 2008, Humphry et al., 2012, with bovine viral diarrhoea), but daily prevalence estimation might be outright impossible for others.

Concentrating the prevalence estimation efforts on a few selected herds is a potential way to reduce the sampling costs associated with the algorithm while maintaining its effectiveness. Although the network metrics were not used to drive the algorithm itself, they could be useful for this selection. Indeed, it is expected that the central herds in the network, i.e. those through which a large proportion of animal movements pass, will play a more important role in the spread of infection (Rautureau et al., 2011, Natale et al., 2011). This rationale is notably the one used for rewiring in the study of Gates and Woolhouse (2015). Recently, Hoscheit et al. (2021) reviewed different centrality measures in the BDNI, while taking into account the dynamic nature of the movement network. Among them, the TempoRank index would for example be a good candidate for selecting a subset of herds to be specifically monitored and to which to apply the rewiring algorithm. By targeting specific movements from or to specific herds, the impact on the network structure would be minimal, while potentially keeping a substantial impact on disease spread risks.

It should be noted that the impact of the algorithm on the structure of the commercial network was already limited, as it targeted a few movements only: less than 20% of the movements for endemic infections and less than 2% of them for epidemic infections. Nevertheless, it seems that this rewiring

method tended to increase the overall connectedness of the herds during endemic infections. Indeed, the increases in degrees and size of the largest strong component indicate that the algorithm connected herds that were originally not so. These metrics are generally correlated with higher expected epidemic risks (Kiss et al., 2006, Dubé et al., 2009). This is however not the case in this context, as these changes result from a trade-off made by the algorithm to prevent the single simulated disease. However, such a rewiring in a real-world context would also have to take into account the potentially increased risks of other diseases, whose spread could be facilitated. The algorithm could be extended to assess several diseases at once, but the additional constraints would probably reduce its effectiveness.

For this study, the network-related outcomes concerned only internal movements, i.e. between herds located in Brittany. However, the simulations also took into account interactions with herds located outside Brittany. More precisely, 20% of all movements whose destination was in the metapopulation had an origin outside the metapopulation, and 25% of all movements whose origin was in the metapopulation had a destination outside of it. On the one hand, this confirms that the movements at risk that could not be rewired could probably be replaced by an import and an export if necessary, since those movements are already rather common. On the other hand, all imports were presumed to never be movements at risk, i.e. that the prevalence status of their origin was never higher than that of their destination. This assumption is not trivial, as it assumes that imports did not create greater infection risks than internal movements. In a real-life context, the additional risk because of imports would also have to be taken into account and possibly managed in another way. More generally, the proportion of movements with an origin outside of the metapopulation is expected to decrease when the number of herds in the metapopulation increase. Extending the use of such an algorithm to a national scale, rather than a single region, could therefore mitigate this problem.

The present study builds upon the results from Ezanno et al. (2021), and confirms the effectiveness of a rewiring method based on targeted movements beyond the specific case of bovine paratuberculosis. Indeed, the algorithm presented by Ezanno et al. (2021) was developed specifically to address the control of bovine paratuberculosis, notably characterized by an endemic status in France and a low detection rate. To do so, they used a specific age-structured epidemiological model (Camanes et al., 2018) and an algorithm calibrated to target the disease. This was also the case for instance of Mohr et al. (2018), which specifically targeted foot-and-mouth disease. Conversely, the present study aimed at assessing more comprehensively the effectiveness of the algorithm. Therefore, it was tested for different epidemiological settings – both endemic and epidemic – using a non-specific epidemiological model, and for broad range of parameter values. Ezanno et al. (2021) highlight the low effectiveness of rewiring for low evolving diseases, but our study shows its relevance in other types of infections.

Although the study here show the effectiveness of rewiring on historical data from the BDNI, the algorithm could also be used prospectively as part of decision-making tools. Indeed, the rewiring method

could work without any simulation of infection, if herd statuses were provided otherwise. Given these statuses and the potential movements to occur, the algorithm would also suggest necessary changes to prevent movements at risk. In this context, the implementation of these changes would also depend on the actual decision of the informed farmers. Unless rewiring is enforced, it is expected that constraints other than sanitary ones would affect movements, which would impact the effectiveness of the algorithm. Coupling it with a decision-making model could provide additional insight on this impact. In order to make it easier to use as part of such decision-making tools, the algorithm has been specifically designed to be able to include additional, different constraints. Additional criteria, e.g. concerning the breed or sex of the animals, could be added easily by providing the algorithm with movements for individuals in each category separately. Depending on their number, however, such criteria would reduce the rewiring possibilities of the algorithm, which could reduce its efficiency.

This study demonstrates the effectiveness of a rewiring method targeting specific movements to reduce infection risks. Our approach thus differs radically from that presented by Gates and Woolhouse (2015), as it also aims at generating minimal changes in the structure of the movement network. Although the algorithm was tested on a cattle movement network, it is applicable to a much wider range of networks in animal and plant populations, e.g. among seed exchange networks, which face similar infection risks (Jeger et al., 2007, Pautasso et al., 2010). While the need for controlled movements makes this method more relevant to agricultural systems, the spatial and temporal scales considered can also be adapted depending on the context. Indeed, the daily time-step and the region level were used here as they correspond to the BDNI data structure, but are not necessary for the algorithm to work. The usefulness of our rewiring method could therefore extend beyond cattle concerns, even though the effectiveness of the algorithm in other contexts remains to be tested.

Funding

This work was supported by the French National Agency for Research (*Agence nationale de la recherche*), project CADENCE [grant number ANR-368 16-CE32-0007].

Declarations of interest

None.

Data statement

The code for the algorithm, as well as additional scripts for formatting the data or running preliminary simulations and dummy test data, are freely available at https://sourcesup.renater.fr/projects/pub-rewir-

algo/. The dataset used in the study is an extraction from the French national bovine identification database (BDNI), which is confidential, and therefore cannot be provided publicly.

References

P. Bajardi, A. Barrat, F. Natale, L. Savini, and V. Colizza. Dynamical patterns of cattle trade movements. *PLoS One*, 6(5):e19869, 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0019869.

F. Ball and T. Britton. Epidemics on networks with preventive rewiring. *arXiv preprint arXiv:2008.06375*, 2020.

G. Beaunée, E. Vergu, and P. Ezanno. Modelling of paratuberculosis spread between dairy cattle farms at a regional scale. *Vet Res*, 46(1):1–13, 2015. ISSN 1297-9716. doi: 10.1186/s13567-015-0247-3.

G. Beaunée, E. Vergu, A. Joly, and P. Ezanno. Controlling bovine paratuberculosis at a regional scale: Towards a decision modelling tool. *J Theor Biol*, 435:157–183, 2017. doi: 10.1016/j.jtbi.2017.09.012.

C. Bidot, M. Lamboni, and H. Monod. Multisensi: Multivariate Sensitivity Analysis, 2018.

F. Biemans, R. B. Romdhane, P. Gontier, C. Fourichon, G. Ramsbottom, S. More, and P. Ezanno. Modelling transmission and control of Mycobacterium avium subspecies paratuberculosis within Irish dairy herds with compact spring calving. *Prev Vet Med*, 186:105228, 2021. ISSN 0167-5877. doi: 10.1016/j.prevetmed.2020.105228.

T. Britton, D. Juher, and J. Saldaña. A network epidemic model with preventive rewiring: Comparative analysis of the initial phase. *Bull Math Biol*, 78(12):2427–2454, 2016. doi: 10.1007/s11538-016-0227-4.

K. Büttner, J. Krieter, A. Traulsen, and I. Traulsen. Static network analysis of a pork supply chain in Northern Germany—Characterisation of the potential spread of infectious diseases via animal movements. *Prev Vet Med*, 110(3-4):418–428, 2013. ISSN 0167-5877. doi: 10.1016/j.prevetmed.2013.01.008.

G. Camanes, A. Joly, C. Fourichon, R. Ben Romdhane, and P. Ezanno. Control measures to prevent the increase of paratuberculosis prevalence in dairy cattle herds: An individual-based modelling approach. *Vet Rec*, 49(1):1–13, 2018. doi: 10.1186/s13567-018-0557-3.

A. Courcoul and P. Ezanno. Modelling the spread of Bovine Viral Diarrhoea Virus (BVDV) in a managed metapopulation of cattle herds. *Vet Microbiol*, 142(1-2):119–128, 2010. doi: 10.1016/j.vetmic.2009.09.052.

C. Dubé, C. Ribble, D. Kelton, and B. McNab. 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, 2009. ISSN 1865-1674. doi: 10.1111/j.1865-1682.2008.01064.x.

463 B. L. Dutta, P. Ezanno, and E. Vergu. Characteristics of the spatio-temporal network of cat-
464 tle movements in France over a 5-year period. *Prev Vet Med*, 117(1):79–94, 2014. doi:
465 10.1016/j.prevetmed.2014.09.005.

466 EU. Regulation (EC) No 1760/2000 establishing a system for the identification and registration of bovine
467 animals and regarding the labelling of beef and beef products and repealing Council Regulation (EC)
468 No 820/97, 2000.

469 P. Ezanno, M. Andraud, G. Beaunée, T. Hoch, S. Krebs, A. Rault, S. Touzeau, E. Vergu, and S. Widgren.
470 How mechanistic modelling supports decision making for the control of enzootic infectious diseases.
471 *Epidemics*, 32:100398, 2020. doi: 10.1016/j.epidem.2020.100398.

472 P. Ezanno, S. Arnoux, A. Joly, and R. Vermesse. Rewiring cattle trade movements helps to
473 control bovine paratuberculosis at a regional scale. *Prev Vet Med*, page 105529, 2021. doi:
474 10.1016/j.prevetmed.2021.105529.

475 M. T. Garoussi, A. Haghparast, and H. Estajee. Prevalence of bovine viral diarrhoea virus antibodies
476 in bulk tank milk of industrial dairy cattle herds in suburb of Mashhad-Iran. *Prev Vet Med*, 84(1-2):
477 171–176, 2008. doi: 10.1016/j.prevetmed.2007.12.016.

478 M. C. Gates and M. E. Woolhouse. Controlling infectious disease through the targeted ma-
479 nipulation of contact network structure. *Epidemics*, 12:11–19, 2015. ISSN 1755-4365. doi:
480 10.1016/j.epidem.2015.02.008.

481 M. Gilbert, A. Mitchell, D. Bourn, J. Mawdsley, R. Clifton-Hadley, and W. Wint. Cattle movements and
482 bovine tuberculosis in Great Britain. *Nature*, 435(7041):491–496, 2005. doi: 10.1038/nature03548.

483 D. Green and R. Kao. Data quality of the cattle tracing system in Great Britain. *Vet Rec*, 161(13):
484 439–443, 2007. ISSN 0042-4900. doi: 10.1136/vr.161.13.439.

485 D. M. Green, A. Gregory, and L. A. Munro. Small-and large-scale network structure of live fish movements
486 in Scotland. *Prev Vet Med*, 91(2-4):261–269, 2009. doi: 10.1016/j.prevetmed.2009.05.031.

487 T. Gross, C. J. D. D'Lima, and B. Blasius. Epidemic dynamics on an adaptive network. *Phys Rev Lett*,
488 96(20):208701, 2006. doi: 10.1103/PhysRevLett.96.208701.

489 A. Hidano, T. E. Carpenter, M. A. Stevenson, and M. C. Gates. Evaluating the efficacy of regionalisation
490 in limiting high-risk livestock trade movements. *Prev Vet Med*, 133:31–41, 2016. ISSN 0167-5877. doi:
491 10.1016/j.prevetmed.2016.09.015.

492 P. Hoscheit, É. Anthony, and E. Vergu. Dynamic centrality measures for cattle trade networks. *Appl*
493 *Netw Sci*, 6(1):1–17, 2021. ISSN 2364-8228. doi: 10.1007/s41109-021-00368-5.

494 R. Humphry, F. Brülisauer, I. McKendrick, P. Nettleton, and G. Gunn. Prevalence of antibodies to
495 bovine viral diarrhoea virus in bulk tank milk and associated risk factors in Scottish dairy herds. *Vet*
496 *Rec*, 171(18):445–445, 2012. doi: 10.1136/vr.100542.

497 M. J. Jeger, M. Pautasso, O. Holdenrieder, and M. W. Shaw. Modelling disease spread and control in
498 networks: Implications for plant sciences. *New Phytol*, 174(2):279–297, 2007. ISSN 0028-646X. doi:
499 doi:10.1111/j.1469-8137.2007.02028.x.

500 R. R. Kao, L. Danon, D. M. Green, and I. Z. Kiss. Demographic structure and pathogen dynamics on
501 the network of livestock movements in Great Britain. *Proc R Soc B*, 273(1597):1999–2007, 2006. doi:
502 10.1098/rsif.2006.0129.

503 I. Z. Kiss, D. M. Green, and R. R. Kao. The network of sheep movements within Great Britain: Network
504 properties and their implications for infectious disease spread. *J R Soc Interface*, 3(10):669–677, 2006.
505 ISSN 1742-5689. doi: doi:10.1098/rsif.2006.0129.

506 M. Lamboni, H. Monod, and D. Makowski. Multivariate sensitivity analysis to measure global con-
507 tribution of input factors in dynamic models. *Reliab Eng Syst Saf*, 96(4):450–459, 2011. doi:
508 10.1016/j.ress.2010.12.002.

509 W.-c. Liu, L. Matthews, M. Chase-Topping, N. J. Savill, D. J. Shaw, and M. E. Woolhouse. Metapopu-
510 lation dynamics of Escherichia coli O157 in cattle: An exploratory model. *Journal of the Royal Society*
511 *Interface*, 4(16):917–924, 2007. ISSN 1742-5689.

512 S. Mohr, M. Deason, M. Churakov, T. Doherty, and R. R. Kao. Manipulation of contact network
513 structure and the impact on foot-and-mouth disease transmission. *Prev Vet Med*, 157:8–18, 2018. ISSN
514 0167-5877. doi: 10.1016/j.prevetmed.2018.05.006.

515 M. Moslonka-Lefebvre, C. A. Gilligan, H. Monod, C. Belloc, P. Ezanno, J. A. Filipe, and E. Vergu. Market
516 analyses of livestock trade networks to inform the prevention of joint economic and epidemiological
517 risks. *J R Soc Interface*, 13(116):20151099, 2016. ISSN 1742-5689. doi: 10.1098/rsif.2015.1099.

518 F. Natale, A. Giovannini, L. Savini, D. Palma, L. Possenti, G. Fiore, and P. Calistri. Network analysis
519 of Italian cattle trade patterns and evaluation of risks for potential disease spread. *Prev Vet Med*, 92
520 (4):341–350, 2009. doi: 10.1016/j.prevetmed.2009.08.026.

521 F. Natale, L. Savini, A. Giovannini, P. Calistri, L. Candeloro, and G. Fiore. Evaluation of risk and
522 vulnerability using a Disease Flow Centrality measure in dynamic cattle trade networks. *Prev Vet*
523 *Med*, 98(2-3):111–118, 2011. doi: 10.1016/j.prevetmed.2010.11.013.

M. Nöremark, N. Håkansson, T. Lindström, U. Wennergren, and S. S. Lewerin. Spatial and temporal investigations of reported movements, births and deaths of cattle and pigs in Sweden. *Acta Vet*, 51(1): 1–15, 2009. doi: 10.1186/1751-0147-51-37.

M. Nöremark, N. Håkansson, S. S. Lewerin, A. Lindberg, and A. Jonsson. Network analysis of cattle and pig movements in Sweden: Measures relevant for disease control and risk based surveillance. *Prev Vet Med*, 99(2-4):78–90, 2011. doi: 10.1016/j.prevetmed.2010.12.009.

M. Pautasso, M. Moslonka-Lefebvre, and M. J. Jeger. The number of links to and from the starting node as a predictor of epidemic size in small-size directed networks. *Ecol Complex*, 7(4):424–432, 2010. ISSN 1476-945X. doi: doi:10.1016/j.ecocom.2009.10.003.

S. Piankoranee and S. Limkumnerd. Effect of local rewiring in adaptive epidemic networks. *Phys Lett A*, 384(15):126308, 2020. ISSN 0375-9601. doi: 10.1016/j.physleta.2020.126308.

S. Rautureau, B. Dufour, and B. Durand. Vulnerability of animal trade networks to the spread of infectious diseases: A methodological approach applied to evaluation and emergency control strategies in cattle, France, 2005. *Transbound Emerg Dis*, 58(2):110–120, 2011. ISSN 1865-1674. doi: 10.1111/j.1865-1682.2010.01187.x.

M. C. Vernon. Demographics of cattle movements in the United Kingdom. *BMC Vet Res*, 7(1):1–16, 2011. doi: 10.1186/1746-6148-7-31.

H.-X. Yang, Z.-X. Wu, and B.-H. Wang. Suppressing traffic-driven epidemic spreading by edge-removal strategies. *Phys Rev E*, 87(6):064801, 2013. doi: 10.1103/PhysRevE.87.064801.