Simulating contact networks for livestock disease epidemiology: a systematic review

William T M Leung, James W Rudge, Guillaume Fournié

THE ROYAL SOCIETY
PUBLISHING

Simulating contact networks for livestock disease epidemiology: a systematic review

William T. M. Leung[1,2], James W. Rudge[1,3] and Guillaume Fournié[2,4,5]

[1]Communicable Diseases Policy Research Group, Department of Global Health and Development, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[2]Veterinary Epidemiology, Economics and Public Health Group, Pathobiology and Population Sciences Department, Royal Veterinary College, London AL9 7TA, UK
[3]Faculty of Public Health, Mahidol University, Bangkok 10400, Thailand
[4]INRAE, VetAgro Sup, UMR EPIA, Université de Lyon, Marcy l'Etoile 69280, France
[5]INRAE, VetAgro Sup, UMR EPIA, Université Clermont Auvergne, Saint Genes Champanelle 63122, France

WTML, 0000-0003-1616-175X; JWR, 0000-0002-3031-3820; GF, 0000-0002-6998-1201

Contact structure among livestock populations influences the transmission of infectious agents among them. Models simulating realistic contact networks therefore have important applications for generating insights relevant to livestock diseases. This systematic review identifies and compares such models, their applications, data sources and how their validity was assessed. From 52 publications, 37 models were identified comprising seven model frameworks. These included mathematical models ($n = 8$; including generalized random graphs, scale-free, Watts–Strogatz and spatial models), agent-based models ($n = 8$), radiation models ($n = 1$) (collectively, considered 'mechanistic'), gravity models ($n = 4$), exponential random graph models ($n = 9$), other forms of statistical model ($n = 6$) (statistical) and random forests ($n = 1$) (machine learning). Overall, nearly half of the models were used as inputs for network-based epidemiological models. In all models, edges represented livestock movements, sometimes alongside other forms of contact. Statistical models were often applied to infer factors associated with network formation ($n = 12$). Mechanistic models were commonly applied to assess the interaction between network structure and disease dissemination ($n = 6$). Mechanistic, statistical and machine learning models were all applied to generate networks given limited data ($n = 13$). There was considerable variation in the approaches used for model validation. Finally, we discuss the relative strengths and weaknesses of model frameworks in different use cases.

1. Introduction

Livestock holdings may be epidemiologically connected through both direct and indirect contacts. Direct contact typically pertains to the movement of livestock between holdings, while mechanisms for indirect contact include the transfer of biological material, equipment or personnel [1]. These contact patterns can be conceptualized as networks in which nodes may represent livestock populations (given that livestock are often managed in groups or are otherwise spatially clustered) and edges represent the contact(s) of interest between those populations. It is well recognized that the structure of livestock contact networks has important implications for infectious disease transmission dynamics [2–6]. Characterizing the structure of these networks therefore plays a crucial role in understanding transmission patterns of infectious diseases in livestock and, consequently, for informing disease risk assessments and control strategies. This may involve the use of disease transmission models which explicitly account for contact network structure [3,7–11].

Insights about the epidemiological importance of livestock contact networks, especially livestock movement (e.g. trade) networks [1,12–15], have been generated by the analysis of routinely recorded livestock movement data collected via

livestock identification and traceability systems (LITS) [14,16–19]. Where such routine data are unavailable (or insufficient), targeted network surveys can also be conducted [7,20–24].

Such empirical approaches are, however, associated with major challenges. In certain settings, LITS may not be implemented as data collection, and sharing may be restricted by commercial interests and related data privacy concerns [3,25–27]. The costs and infrastructure required to implement and sustain routine systems also constrains their feasibility, especially in low- and middle-income countries [28]. The analysis and utility of such data may be constrained by its vastness [28]. Moreover, a lack of updated or complete data may also limit its use for supporting decision making during disease outbreaks [28–30]. While network surveys have been used when such data are unavailable, these are usually targeted towards specific geographical locations and time periods. Indeed, both routine and non-routine network data capture activities are highly resource intensive and are therefore likely to be targeted towards livestock species or production types of particular interest from a national livestock disease-management perspective [28,31–33].

Model-based approaches are increasingly being used to help address some of these challenges. We therefore conducted a systematic review to provide an overview of the state-of-the-art in modelling livestock contact networks. Our objectives were to identify the main types of models and methods used, compare their applications and data requirements, and examine the extent to which such models have been validated. Based on the findings, we also discuss key challenges and opportunities for future research in this area. In this review, we focus on studies which have employed empirically informed, model-based approaches of network (re)-construction or inference, with a primary interest in epidemiologically relevant (i.e. potentially infectious) contacts between livestock populations.

2. Methods

2.1. Systematic search strategy

This systematic review followed the PRISMA 2020 guidelines for the reporting of systematic reviews [34]. Search terms were developed around four key topics: (i) livestock and poultry, (ii) networks, (iii) models, and (iv) disease. Four databases—Medline, Embase, Web of Science and Scopus—were queried using title, abstract and keyword searches on 22 January 2021 and no date limits. Database searches were repeated on 27 January 2023 to cover all records published up to this date. Relevant subject headings were applied to databases using subject heading indexing (i.e. Medline and Embase; electronic supplementary material, table S1). Search terms within the 'networks' topic were informed by previous reviews of the use of network simulation models in different contexts [35–39]. However, broad terms were also included to ensure identified records were not restricted to known model types. Within each search topic, Boolean 'OR' operators were used to combine search terms and subject headings, while different topics were combined using 'AND' operators (electronic supplementary material, table S1). Wildcards, truncations and adjacency searches were applied using the relevant syntax for each database. Peer-reviewed papers and conference proceedings were all eligible for inclusion. The screening process was expanded to include the reference lists of the included publications, as well as any papers that cited them. For full search terms see electronic supplementary material, table S1.

2.2. Inclusion and exclusion criteria

Inclusion and exclusion criteria were agreed by all authors. A single reviewer screened records but discussed any records for which inclusion was uncertain with the other authors. Screening was split into two stages:

Stage 1: Titles, abstracts and keywords were screened; records were rejected if any of the following statements were true: (i) there was no reference to livestock; (ii) there was no reference to contacts between livestock, contact networks or infectious disease dynamics on networks; (iii) the record was not peer-reviewed, and (iv) the record was not written in English.

Stage 2: Full texts were screened; records were retained if all following statements were true: (i) a model was used to simulate a network of epidemiologically relevant contacts between livestock subpopulations; (ii) the model attempted to reproduce structural properties of an empirical network and/or its underlying generating mechanisms, and (iii) these properties or mechanisms were informed empirically.

Hence, we did not review records which simulated theoretical networks (e.g. to be used as reference or null models) and/or which randomized some aspects of a network to make comparisons with empirical networks (e.g. [40,41]). We also excluded studies that solely reconstructed transmission networks, since these are subsets of the contact networks which are the focus of this review. Where multiple models were used in papers, each model was screened individually for inclusion.

2.3. Data extraction

Information from each study was systematically recorded in a data extraction table. This was designed to record information about: (i) the type of model used; (ii) the applications of models; (iii) characteristics of the empirical network under study (livestock type, geographical location and disease focus); (iv) definition of network nodes and edges; (v) data types and variables used for model fitting, and (vi) how the performance of models was assessed (table 2). Descriptive analyses and visualizations of the frequency of key study characteristics were conducted using R v. 4.2.0 [42].

2.4. Model classifications

Following exploratory scoping of the literature, particularly previous reviews on network simulation models in other disciplines [36–39], we classified models into three groups: mechanistic, statistical and machine learning. Though these categories are not mutually exclusive (e.g. mechanistic model parameters may be estimated using statistical methods), they are useful for describing the general characteristics of the reviewed models, as described below.

Mechanistic models are here defined as mathematical equations *or* an algorithmic set of rules, a 'mechanism', used to generate a set of edges between nodes, i.e. a network. We include in this grouping mechanistic models that span from (i) abstracted and intentionally simplified 'mathematical models' [38], such as scale-free and small-world models (and which include the 'probabilistic' and 'idealized' models/networks described by others) [37,39], to (ii) complex agent-based models (ABMs) explicitly modelling individual-level contact processes. Notably, across both of these subgroups, the generating mechanisms may simply serve as an arbitrary algorithmic tool used to generate networks exhibiting a certain topology, or else they may be configured to reproduce the emergent processes (assumed or otherwise) that generated the observed network, that is, based on 'first principles' [43,44].

Statistical models describe a network as a function of factors hypothesized to be associated with edge formation. They start
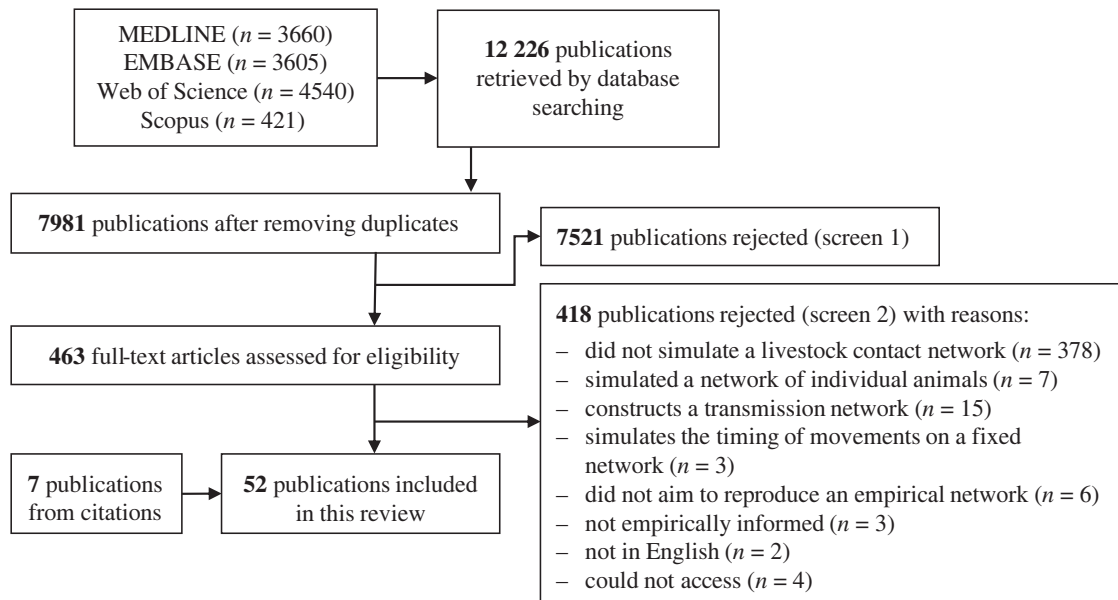
Figure 1. PRISMA flow diagram.

with observations of an empirical network and fit the parameters of a selected model framework to the data through formal statistical inference [36,38,39]. Within this group, we include standard statistical models (e.g. generalized linear models) which may be used to estimate the probability or strength of an edge between nodes given a set of covariates, in addition to network-specific statistical models which explicitly account for the dependencies inherent to network data [36,38,39].

Machine learning models learn patterns in the data without the model being specified by the user and commonly place an emphasis on predictive accuracy rather than causal inference [45,46]. These can be broadly categorized according to whether the model fitting is 'supervised', where the value of the dependent variable is known (i.e. data are 'labelled' in machine learning-terminology), or 'unsupervised', which use 'unlabelled' data and commonly include clustering algorithms [45]. In the context of network simulation, they may be used to solve classification and regression problems.

3. Results

3.1. Screening process

Database searches retrieved 12 226 publications of which 7981 (65%) were unique. Title, abstract and keyword screening excluded 7521 (94%) unique records (figure 1). A further 418 (5%) were excluded after screening full texts, mostly because they did not simulate a livestock contact network but presented descriptive analyses of empirical networks or simulated infectious disease transmission on empirical networks (figure 1). Six additional publications were identified from the citations of included papers. A single additional publication citing these publications was then identified. Therefore, a total of 52 publications published between 2009 and 2022 were eligible for inclusion (see electronic supplementary material, table S2 for all exclusion reasons).

To identify the number of different models used across all included studies, we considered a model to be 'distinct' from others when a specific framework was applied to a particular dataset. Hence, analyses in 20 publications were based on previously published models (table 2), while two

Table 1. Model frameworks applied to simulate livestock contact networks across 52 included studies. ABM = agent-based model; (T)ERGM = (temporal) exponential random graph model; GM = gravity model; RF = random forests.

category	model framework	number of models	number of publications
mechanistic	mathematical models	8	7
	ABMs	8	15
	radiation models	1	1
statistical	(T)ERGMs	9	7
	GMs	4	4
	other statistical models	6	17
machine learning	RF	1	1
total	-	37	52

publications presented multiple models, applying different model types to a single setting [47], or the same model type to different settings [29]. Consequently, 37 distinct models (tables 1 and 2) were identified and reviewed across the 52 included publications. We refer to unique models using the first published instance.

Following the PRISMA checklist, we highlight nine studies that might appear to meet the inclusion criteria, but were excluded. Three studies rewired empirically observed networks without also attempting to simulate the empirical network [93–95]. Three studies simulated the timing or volume of livestock movements on a predefined (non-modelled) network [96–98]. Two used mechanistic models with entirely hypothetical parameter values [99,100]. One study applied random forests (RFs) to predict the timings of trading events, without using this information to simulate a network [101].

Table 2. Summary of key characteristics and applications of 37 identified models. ABM = agent-based model; GM = gravity model; RF = random forests; (T)ERGM = (temporal) exponential random graph model; LITS = livestock identification and traceability system; limited data = simulating a network from the available data, when empirical networks are incompletely characterized; network-generating processes = inference of factors associated with network (or edge) generation; structure and transmission = analytical exploration of the relationship between network structure and diffusion of phenomena (e.g. disease) on networks; SA disease control = scenario analysis related to assessing the impact of disease control strategies; SA altering network = comparing the impact of alternative network configuration scenarios on simulated disease transmission patterns; SA surveillance = exploring disease surveillance scenarios; behavioural response = modelling adaptive behaviour, e.g. farmers' response to disease on a network.

model ID		model classification			model purpose		network characteristics						data
model [model name]; other papers using the model	no.	year	model category	model framework	model applications	infectious disease model	livestock focus	setting	disease focus	nodes	edges	static / dynamic	data used for calibration
(Ferdousi et al. [48])	1	2019	mechanistic	generalized random graph	structure and transmission, limited data, SA: disease control	yes	pigs	USA	African swine fever	livestock holdings, markets	livestock movement	static	network survey
(Gates & Woolhouse [44])	2	2015	mechanistic	generalized random graph	SA: altering network	yes	cattle	Britain	hypothetical	livestock holdings	livestock movement	static	LITS
(Thakur et al. [47]) (A)	3	2015	mechanistic	Watts–Strogatz	structure and transmission, limited data	yes	pigs	Canada	porcine reproductive and respiratory syndrome	livestock holdings	livestock movement, vehicle	static	network survey
(Thakur et al. [47]) (B)	4	2015	mechanistic	scale-free	structure and transmission, limited data	yes	pigs	Canada	porcine reproductive and respiratory syndrome	livestock holdings	livestock movement, vehicle	static	network survey
(Tago et al. [49])	5	2016	mechanistic	scale-free	SA: disease control, behavioural response	yes	cattle	France	hypothetical	livestock holdings, markets, exchangers	livestock movement	static	LITS
(Lennartsson et al. [50]; [SpecNet])	6	2012	mechanistic	other mathematical	presents model	no	pigs	Sweden	non-specific	livestock holdings, slaughter point	livestock movement	static	LITS
(Rossi et al. [51])	7	2017	mechanistic	spatial	structure and transmission, limited data	yes	cattle	Italy	hypothetical highly contagious	livestock holdings	livestock movement, personnel	static	LITS
(Hu et al. [52])	8	2021	mechanistic	spatial	limited data	yes	pigs	China	African swine fever	livestock holdings, slaughter point	livestock movement	static	LITS
(Wiltshire [53]; [RUSHPNBMI]) (Bucini et al. [54]; Koliba et al. [55]; Wiltshire et al. [27])	9	2018	mechanistic	ABM	SA: altering network, structure and transmission, behavioural response	yes	pigs	USA	porcine epidemic diarrhoea	livestock holdings, slaughter points, feed mills	livestock movement, feed, vehicle	dynamic	emergent
(Yang et al. [56]) (Yang et al. [57,58])	10	2019	mechanistic	ABM	SA: disease control, limited data, behavioural response	yes	cattle	USA	foot and mouth disease	livestock holdings, exchangers, markets	livestock movement, vehicle	dynamic	emergent

#	Reference	Year					Species	Country	Disease	Node	Edge		Data
11	(Ross et al. [59])	2011	mechanistic	ABM	presents model	no	cattle	USA	bovine tuberculosis	livestock holdings, markets	livestock movement	dynamic	emergent
12	(Liu et al. [60]; [Epirur_Cattle])	2012	mechanistic	ABM	limited data	yes	cattle	USA	hypothetical direct contact	livestock holdings	livestock movement	dynamic	emergent
13	(Ansari et al. [61])	2021	mechanistic	ABM	presents model	no	pigs	Germany	non-specific	livestock holdings, exchangers	livestock movement	dynamic	LITS
14	(Brock et al. [62])	2021	mechanistic	ABM	SA: disease control	yes	cattle	Ireland	bovine herpesvirus type 1	livestock holdings	livestock movement	dynamic	emergent
15	(Knight et al. [63,64])	2021	mechanistic	ABM	structure and transmission, behavioural response	yes	cattle	Scotland	hypothetical slowly spreading	livestock holdings	livestock movement	dynamic	LITS
16	(Kim et al. [65]; Pomeroy et al. [66])	2016	mechanistic	ABM	structure and transmission	yes	cattle	Cameroon	foot and mouth disease	geo-locations	livestock movement	dynamic	network survey
17	(Kong et al. [67])	2022	mechanistic	radiation model	limited data	no	poultry	China	non-specific	geo-locations	livestock movement	static	emergent
18	(Valdes-Donoso et al. [68])	2017	machine learning	RF	limited data, network-generating processes	no	pigs	USA	porcine reproductive and respiratory syndrome	livestock holdings, markets	livestock movement	static	network survey
19	(Nicolas et al. [69])	2018	statistical	GM	network-generating processes, limited data	no	cattle, sheep/goats, camels	Mauritania	non-specific	geo-locations	livestock movement	static	network survey
20	(Chaters et al. [28])	2019	statistical	GM	limited data, network-generating processes	yes	cattle	Tanzania	non-specific	geo-locations	livestock movement	static	movement permits
21	(Qiqi Yang et al. [15])	2020	statistical	GM	limited data	no	poultry	China	avian influenza	geo-locations	livestock movement	static	network survey, emergent
22	(Blair and Lowe [70])	2022	statistical	GM	SA: disease control	no	pigs	USA	non-specific	geo-locations, slaughter point	livestock movement	static	network survey
23	(Ortiz-Pelaez et al. [71])	2012	statistical	ERGM	network-generating processes	no	sheep/goats	Ethiopia	non-specific	geo-locations	livestock movement	static	network survey
24	(Relun et al. [29]) (A)	2017	statistical	ERGM	network-generating processes	no	pigs	Bulgaria	non-specific	livestock holdings, exchangers	livestock movement	static	LITS
25	(Relun et al. [29]) (B)	2017	statistical	ERGM	network-generating processes	no	pigs	Spain	non-specific	livestock holdings, exchangers	livestock movement	static	LITS
26	(Relun et al. [29]) (C)	2017	statistical	ERGM	network-generating processes	no	pigs	France	non-specific	livestock holdings, exchangers	livestock movement	static	LITS
27	(Kukielka et al. [32])	2017	statistical	ERGM	network-generating processes	no	pigs	Georgia	African swine fever	geo-locations	livestock movement	static	network survey

(Continued.)

Table 2. (*Continued.*)

model ID		model classification			model purpose		network characteristics						data
no.	model [model name]; other papers using the model	year	model category	model framework	model applications	infectious disease model	livestock focus	setting	disease focus	nodes	edges	static / dynamic	data used for calibration
28	(Poolkhet *et al.* [72])	2018	statistical	ERGM	network-generating processes	no	poultry	Thailand	avian influenza	livestock holdings, exchangers, markets, slaughter point, other	livestock movement, other	static	network survey
29	(Belkhiria *et al.* [73])	2019	statistical	ERGM	network-generating processes	no	cattle, sheep/goats, donkeys	Senegal	Rift valley fever	geo-locations	livestock movement	static	network survey
30	(Hammami *et al.* [74])	2022	statistical	ERGM	network-generating processes	no	pigs	France	non-specific	livestock holdings, slaughter point	livestock movement	static	LITS
31	(Lee *et al.* [75])	2021	statistical	TERGM	structure and transmission, SA: disease control	yes	pigs	Vietnam	African swine fever	livestock holdings	livestock movement, indirect	dynamic	network survey
32	(Lindström *et al.* [25]; [USAMM]) (Brommesson *et al.* [76]; Buhnerkempe *et al.* [77,78]; Gilbertson *et al.* [79]; Gorsich *et al.* [80,81]; Kao *et al.* [82]; Sellman *et al.* [83]; Tsao *et al.* [84])	2013	statistical	statistical other	SA: disease control, limited data, SA: surveillance	yes	cattle	USA	non-specific, foot and mouth disease, bovine tuberculosis	geo-locations	livestock movement	static	movement permits, census
33	(Sellman *et al.* [85])	2022	statistical	statistical other	limited data	no	pigs	USA	porcine epidemic diarrhea	livestock holdings	livestock movement	static	movement permits, census
34	(Lindström *et al.* [86]) (Brommesson *et al.* [87]; (Lindström *et al.* [88–90])	2009	statistical	statistical other	network-generating processes, structure and transmission	yes	cattle, pigs	Sweden	non-specific, hypothetical	livestock holdings	livestock movement	static	LITS
35	(Xiao *et al.* [91]) (Pomeroy *et al.* [66])	2015	statistical	statistical other	network-generating processes	yes	cattle	Cameroon	foot and mouth disease	geo-locations	livestock movement	dynamic	network survey
36	(Moon *et al.* [26])	2019	statistical	statistical other	limited data	no	pigs	USA	non-specific	livestock holdings	livestock movement	static	census
37	(Schumm *et al.* [92])	2015	statistical	statistical other	limited data	no	cattle	USA	non-specific	geo-locations	livestock movement	static	census

3.2. General model characteristics

The identified models were applied to 20 countries in four continents; no eligible models were applied to Australia or South America. The USA was the most well-represented country, with 11 distinct models applied (figure 2a). Most models were applied to a single livestock type, including pigs ($n = 17$), cattle ($n = 13$) and poultry ($n = 3$). Three models were applied to multiple livestock types (figure 2d). All models were applied in a disease context, related to specific ($n = 18$), non-specific ($n = 13$) or hypothetical diseases with specific characteristics ($n = 6$). Infectious disease transmission was simulated on the networks generated by 18 models (13 mechanistic; 5 statistical; table 2).

In 25 models, nodes represented farms or herds, with 14 of these also accounting for other units such as markets, slaughterhouses and/or livestock traders. Nodes were livestock populations in given administrative areas (e.g. villages, provinces and counties) in the other 12 models. Edges represented livestock movements in all models: either movements of animals among populations ($n = 24$), or transhumant movements of whole livestock populations between geographical areas ($n = 3$). Seven models simulated multi-layer networks with additional sets of edges representing epidemiologically relevant contacts via vehicles, personnel or feed providers. A single model broadly defined an edge as any type of potentially infectious contact in the context of avian influenza without defining transmission routes specifically [72]. Most models ($n = 27$) generated static networks. However, the timing of trades on the simulated static network was sometimes time varying, e.g. based on a probability of trading [49]. Alternatively, nodes or edges were sometimes added or removed by copying empirical records exactly (i.e. without modelling these) [44,52]. Contrastingly, eight ABMs and two statistical models generated dynamically evolving networks.

Most models were statistical ($n = 19$), with the most common frameworks being exponential random graph models (ERGMs; $n = 9$), gravity models (GMs) ($n = 4$) and other statistical models ($n = 6$). Only one machine learning model, based on RFs, was identified. The mechanistic models ($n = 17$) included mathematical models ($n = 8$), ABMs ($n = 8$) and a radiation model ($n = 1$) (table 1). The first model was published in 2009, but most ($n = 31$; 84%) were published between 2015 and 2022 (figure 2c).

In the following sections, we first review the objectives addressed by the different model frameworks and the data sources used. We then introduce the key methodological characteristics of each modelling framework, including how they have been calibrated to data, and review the degree to which their performance was assessed.

3.3. Model applications

Network simulation models were used for a range of applications which varied according to the model type used (figure 3a; table 2). For 13 models, multiple applications were identified.

Approximately half of models (16/37) were used to generate networks based on limited data, for example where total network data was not available but descriptive statistics of that network were, or where models based on complete networks were used for prediction in other settings. These included all model frameworks described above, except

ERGMs. A single study used artificially constrained data on indirect contacts among farms to explore how inferring these contacts using different levels of information and assumptions influenced the outputs of disease transmission models [51].

A third of models ($n = 13$), mostly statistical ($n = 12$), were applied to explore network-generating processes, specifically, the inference of factors associated with network (or edge) generation. For the RF model, the relative importance of predictors was assessed by comparing prediction accuracies of models with and without a given predictor. Nine models, mostly mechanistic ($n = 7$), were applied to analytically explore the relationship between network structure and diffusion of phenomena (e.g. disease) on networks.

Models were also applied to test scenarios related to (i) assessing the impact of disease control strategies ($n = 7$) such as targeted livestock movement restrictions, culling or vaccination; (ii) using simulated livestock movement patterns to inform optimal sites for directing disease surveillance activities ($n = 1$; [80]), and (iii) comparing the impact of alternative network configuration scenarios on simulated disease transmission patterns ($n = 2$; all mechanistic). These scenarios involved, for example, rewiring nodes [44] and changing the composition of the farm population [53]. Mechanistic models were applied to explore the interaction between agents' adaptive behaviour, and network formation or disease spread. Examples of such applications included modelling of farmers' decisions to implement biosecurity measures in response to disease risk [54], or trigger sales in anticipation of movement restrictions [49]. In Knight et al. [64], farmers' adaptive behaviour (i.e. anticipatory response to disease control interventions) influenced the formation of the network itself.

Three models were presented as a proof of principle to demonstrate their ability to reproduce structural features of an empirical livestock contact network, without further application [50,59,61]; these models were therefore omitted from figure 3a.

3.4. Data sources used

Different data sources were exploited for calibrating models, with some variation seen between model types (figure 3b). Most models ($n = 30$) were informed by empirical network data, including data from network surveys ($n = 13$), LITS ($n = 12$), censuses with some data on livestock trade, i.e. capturing total number of animals 'sold or moved' by actors in a given year ($n = 2$), and livestock movement permits which are used in some countries for recording and regulating movements, e.g. across administrative borders ($n = 1$; figure 3b). Contrastingly, mechanistic models sometimes did not use network data, but instead used data to parametrize model processes influencing edge formation (e.g. herd demographic processes ($n = 6$; table 2)).

While models sometimes exploited similar data types, the way that these data were used to calibrate models varied substantially according to model type as detailed in the next section (§3.5).

3.5. Model frameworks
3.5.1. Mechanistic

3.5.1.1. Mathematical models ($n = 8$)

3.5.1.1.1. Generalized random graphs. Random graphs generate edges between sets of nodes at random, either by assigning a
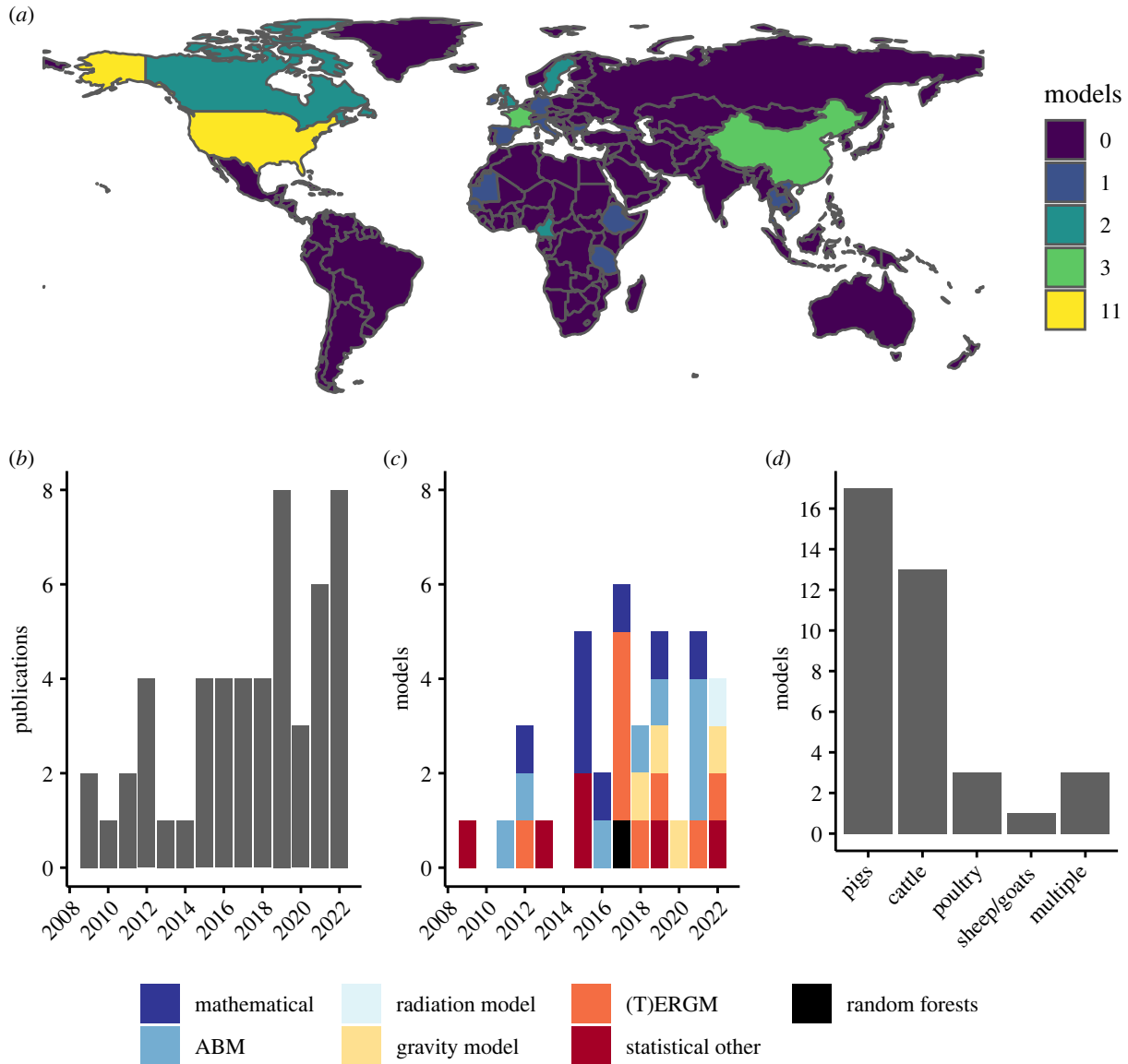
Figure 2. Scope of included papers and models: (*a*) map of countries models were applied to; (*b*) papers published by year; (*c*) models published by year according to model group (blue = mechanistic, red/orange = statistical, black = machine learning); (*d*) livestock types models were applied to.

fixed number of edges [102,103] or by assigning edges with a fixed uniform probability [104]. These models therefore control for network density alone, and the resulting networks fail to capture some important structural features of empirical networks, especially high clustering and a right-skewed degree distribution [38,105,106].

Generalizations may, however, be applied to control for other network structural features beyond density thus permitting the generation of more 'realistic' networks [38,106]. The configuration model, or matching algorithm [107,108], allows for degree distribution to be fixed by algorithmically assigning a number of incoming and outgoing connections (stubs) to nodes, while randomly matching in- and out-stubs between different nodes. Other structural features can be controlled for: for example, in the pig movement network generated in Ferdousi *et al.* [48], connections were only permitted between certain stub combinations, thus additionally controlling for selective mixing among nodes (assortativity). Gates & Woolhouse [44] also adopted a modified configuration algorithm to generate cattle trade networks, preserving farms' empirical daily amounts of purchases and sales, while selectively

matching those reported to have exchanged cattle of the same type (dairy/beef) in the same market, on the same day.

3.5.1.1.2. Scale-free models. Other types of mathematical model seek to reproduce stylized topologies that are common in empirical networks. A key example is the scale-free property which results from the network degree distribution following a power law: $p_k \sim k^{-\gamma}$; where k denotes degree and γ the scaling parameter. The Barabasi & Albert [109] preferential-attachment model generates scale-free networks by progressively adding nodes to a network, with new nodes preferentially forming edges with high-degree nodes. This generates hub-like structures observed in many empirical networks, including those of livestock, where most nodes are poorly connected and a small number of nodes (e.g. markets and breeding farms) have a very high number of connections [105,110].

Thakur *et al.* [47] used the Barabasi–Albert model to simulate scale-free pig trade networks, fitting the model with a scaling parameter derived from empirical studies. The resulting network was altered in a second step by randomly
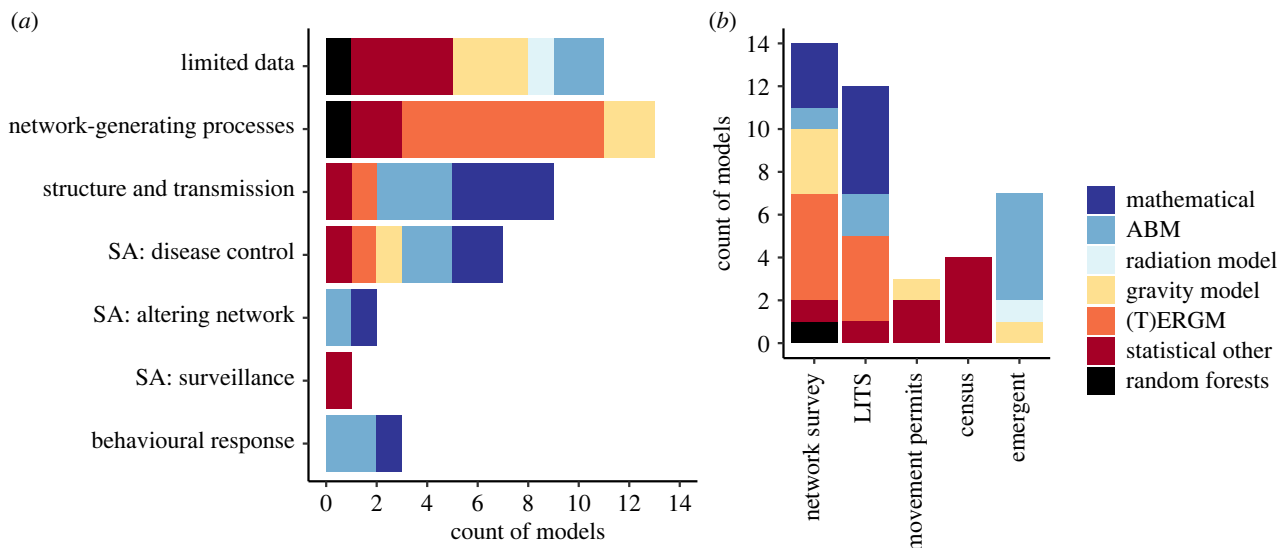
Figure 3. Model applications and data sources by model framework. (a) Model applications (multiple permitted): limited data = simulating a network from the available data, when empirical networks are incompletely characterized; network-generating processes = inference of factors associated with network (or edge) generation; structure and transmission = analytical exploration of the relationship between network structure and diffusion of phenomena (e.g. disease) on networks; SA disease control = scenario analysis related to assessing the impact of disease control strategies; SA altering network = comparing the impact of alternative network configuration scenarios on simulated disease transmission patterns; SA surveillance = exploring disease surveillance scenarios; behavioural response = modelling adaptive behaviour e.g. farmers' response to disease on a network. (b) Data sources used for model calibration. Blue = mechanistic; red/orange = statistical; black = machine learning. LITS = livestock identification and traceability systems; emergent = did not use network data but instead used data to parametrize model processes influencing edge formation.

rewiring edges connecting node types that were not connected in the empirical network, while preserving clustering coefficient and mean degree of the Barabasi–Albert model simulation. Tago et al. [49] generated a scale-free cattle trade network using an empirically derived scaling parameter and mimicked the real network by classifying nodes as markets, dealers or farms, based on the degree of these nodes determined empirically.

3.5.1.1.3. Watts–Strogatz model. The Watts–Strogatz model is another example of a model which reproduces particular features of empirical networks—in this case, 'small-world' properties. The latter refers to networks with short average path lengths, as observed in random graphs, but with higher clustering than is found in random graphs of equivalent size and the same mean degree [111].

This is achieved by taking a ring lattice network, which exhibits high clustering, and randomly rewiring a proportion of its edges such that average path length is reduced. The edge rewiring probability (p) is the single parameter by which the network can be interpolated between the highly clustered lattice and random graph [36,105]. Thakur et al. [47] used this model to generate pig trade networks, choosing a value for p to reproduce clustering coefficients observed in empirical networks.

3.5.1.1.4. Other mathematical models. Other network simulation model frameworks have been devised within different fields of study. Lennartsson et al. [50] describe an algorithm which generates spatially explicit networks of a defined number of nodes and mean degree which can then be tuned to target specified levels of degree-assortativity (selective mixing between nodes of similar degree), clustering coefficient, fragmentation index and spatial aggregation of nodes (random to aggregated). As a proof of principle, the authors generated

networks matching values of these statistics as observed in an empirical swine transportation network.

3.5.1.1.5. Spatial models. With the models described above, the influence of nodes' spatial locations is irrelevant for edge formation. In reality, however, the probability of a connection between livestock populations is likely to be influenced by the geographical distance between them [51,86,112]. While distance may be a variable in other types of model, some of the simplest spatial models express the probability of an edge between nodes as a function of distance alone. For example, in Hu et al. [52], edges between nodes were simply assigned if the Euclidean distance was lower than an empirically informed threshold. In Rossi et al. [51], the probability of contacts between farms via veterinary staff visits was estimated by fitting a logistic regression with distance as the predictor variable.

3.5.1.2. Agent-based models ($n = 8$)
In ABMs, a set of autonomous agents interact with one another and their environment according to defined rules and processes [113,114]. A key feature of ABMs is that they allow complex phenomena to emerge from such processes [114]. Indeed, a livestock contact network can be considered to emerge from the multitude of economic, demographic, husbandry or other behavioural processes occurring at the level of individual agents operating in the system. This may be explicitly modelled within an ABM framework.

In six identified ABMs, network evolution was driven by herd demographic processes (e.g. livestock births, ageing/growth and deaths), and agent trade or partnership generation processes (e.g. selection of trade partners according to geographical distance, and compatibility in terms of industry role and current need to buy or sell) [53,56,57,59–62]. In these models, agents could be defined with distinct industry roles,

holding capacities and geographical locations. In an additional model layer in Liu *et al.* [60], individual animal contacts during grazing were modelled using random walks. In Knight *et al.* [63,64], a dynamic trade network was generated from defined partnership rules—with the rate at which trade partnerships formed and dissolved, dependent on farms' in- and out-flow of animals (i.e. supply and demand). In the most recent paper, farm-level demand, and consequently farmers' edge forming and dissolving behaviours, were adaptive to market shocks such that farms with high-demand sought partnerships at a higher rate. Kim *et al.* [65] simulated a population of mobile pastoralist agents based on seasonal movement rules informed by field surveys. Edges (contact between herds via grazing) were then considered between agents setting up camp within a given distance from one another.

3.5.1.3. Radiation models ($n = 1$)

Radiation models, which were initially developed in the human mobility literature as an alternative to GMs ([115]; see next section), represent a mechanistic approach to predict human movements based on population distributions alone (i.e. distance is not used directly). This method takes analogy from radiation emission and absorption processes in physical sciences and was initially used to describe human commuting patterns, with commuters being 'emitted' from an origin and 'absorbed' by employment opportunities [115]. The model stipulates that the commuting flow (T_{ij}) between an origin (i) and destination (j) is a function of the size of their respective populations (m_i and n_j) and, notably, the 'intervening opportunities' between i and j (alternative employment sinks). The latter are represented by the population (s_{ij}) in the area of the circle with radius r_{ij}, centred at i (excluding m_i and n_j) (equation (3.1)). The variable, T_i represents the overall count of individuals starting their journey at location i $\left(T_i \equiv \sum_{j \neq i} T_{ij} \right)$, which is taken as a proportion of m_i.

$$\langle T_{ij} \rangle = T_i \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})}. \quad (3.1)$$

Kong *et al.* [67] adapted the radiation model to predict country-scale poultry flows in China, with poultry population representing supply (m_i), and human populations representing demand (n_j) and 'intervening' demand (s_{ij}).

3.5.2. Statistical

3.5.2.1. Gravity models ($n = 4$)

GMs were initially developed to model the flow of commodities between pairs of discrete geographical areas (C_{ij}, from origin i to destination j) as a function of their distance (d_{ij}) and Gross National Products representing supply (push) at origin and demand (pull) at destination (p_i and p_j), with normalizing constant k and coefficients α, β and γ (equation (3.2)) [116,117]. The standard formulation of the flow of commodities from node i to j (C_{ij}) is

$$C_{ij} = k \frac{p_i^{\alpha} p_j^{\beta}}{d_{ij}^{\gamma}}. \quad (3.2)$$

This concept has been applied to model livestock trade as a function of livestock population at an origin (supply) and human population at a destination (demand) [15,28,69,70], with different functional relationships (e.g. exponential and power law) between distance and flows having been investigated [15]. Beyond the basic principles of mass and distance, the actual specification of GMs has been loosely defined [118]. GMs may be parametrized in equation (3.2) by fixing the coefficients α, β and γ; an approach which essentially represents a mechanistic parametrization. More often, however, these coefficients are estimated by statistical inference. For example, Qiqi Yang *et al.* [58] used both mechanistic and statistical GM parametrizations to model poultry movements. Equation (3.2) is commonly linearized by logarithmic transformation allowing additional covariates, hypothesized to be relevant for edge formation, to be included in the model (equation (3.3)).

$$\log(C_{ij}) = k + \alpha \log(p_i) + \beta \log(p_j) + \gamma \log(d_{ij}) + \dots \quad (3.3)$$

The coefficients of such models may then be estimated by ordinary least-squares (OLS) regression (e.g. [28,69,70]).

3.5.2.2. Exponential random graph models ($n = 9$)

Under an ERGM formulation, the observed network is considered as just one realization of possible networks (configurations of edges given a set of nodes) with certain characteristics that result from an unknown stochastic process [119]. The ERGM defines a model of this network generation process and a probability distribution over all possible networks. Parameters are selected and estimated, such that the probability of the observed network being generated under the defined model is maximized. It may take a general form as in equation (3.4). Here, the dependent variable is the *whole network* (the probability of drawing the observed network y from the distribution Y), which is modelled as a function of covariates $z_k(y)$ hypothesized to be relevant for network formation. The covariates are weighted by coefficients θk, with c being a normalizing constant [119,120].

$$P_\theta(Y = y | n \text{ nodes}) = c e^{\theta_1 z_1(y) + \dots + \theta_k z_k(y)}. \quad (3.4)$$

A model with a covariate for network density alone is equivalent to a random graph model [119]. However, additional covariates may describe attributes of edges, nodes or notably, local structural features, such as the tendency for reciprocated edges, or the tendency for triangles to form (i.e. where three nodes are completely connected) [121]. Network simulation is achieved by drawing from the probability distribution of possible network configurations given a set of nodes and their attributes. This is the basis for model fitting and assessment of goodness-of-fit: coefficients are fit and the model goodness-of-fit checked based on comparison between characteristics of the simulated and empirical networks [122]. ERGM output is analogous to a logistic regression making their interpretation straightforward [29,71].

ERGMs have been fitted to networks of livestock movements between aggregated spatial units [32,71,73], or actors such as livestock holdings [29,72,74]. These models have sometimes been applied to livestock networks of entire countries (e.g. [29,74]). The use of ERGMs in this context has allowed livestock contact networks to be modelled and simulated as a function of the tendency of farms to form (dis-)assortative trade partnerships with respect to farm

size, type, management practices, company affiliation or location [29,74], in addition to local structural factors [29,32,71,73].

An extension of ERGMs, temporal exponential-family random graph models (TERGMs), enables the statistical modelling of tie dynamics [123]. Here, ERGMs are used to model both tie formation and dissolution, with potentially distinct models for each process. Separable-TERGMs (STERGMs) are used in the latter case. While these models were developed for the statistical modelling of empirical dynamic networks, model parameters may alternatively be defined without being inferred statistically i.e. similar to mechanistic modelling. Lee *et al.* [75] applied TERGMs in this way to simulate dynamic contact networks among pig farms according to a defined mean degree (overall and by node type) and the frequency of contacts.

3.5.2.3. Other statistical models (*n* = 6)

In a series of developments, [86–90] applied a hierarchical Bayesian model to Swedish pig and cattle movement networks incorporating data on between-holding distances, origin and destination production types, and the number of animals in each holding.

Building on these, the USAMM model [25], which has been applied and modified extensively [76–84], uses a Bayesian kernel approach to reconstruct the US cattle trade network. Similarly to GMs, movement probabilities were modelled as a function of the number of cattle premises at the origin and destination, and the distance between them, while also incorporating data on historical state-level cattle inflows. Sellman *et al.* [85] adapted these methods to reconstruct the national US pig movement network.

Xiao *et al.* [91] modelled pastoralists' movements by fitting statistical models to detailed movement survey data. Distinct seasonal movement trajectories were modelled according to different movement models. For example, origin-destination movements were modelled using a Brownian bridge motion model. This movement model was used to generate dynamic daily contact networks among mobile herds in a separate study [66], with 'contacts' between herds being considered when pastoralists set up camp within a given distance from one another on a given day—corresponding to grazing distances observed in field surveys.

Moon *et al.* [26] and Schumm *et al.* [92] used a statistical inferential method of maximum entropy (which is designed to estimate probability distributions from highly dimensional data) to estimate the movement probabilities of pigs within and between geographical units from census data. Based on the size and number of farms within each county, these movement probabilities were then used to simulate a farm-to-farm pig movement network.

3.5.3. Machine learning

3.5.3.1. Random forest (*n* = 1)

The probability or strength of an edge between two nodes can be treated, respectively, as a classification or regression problem which may be addressed using machine learning models such as classification or regression tree-based approaches. These models perform repeated partitions of the data based on the values of predictor variables, such that the observations in each partition are increasingly

homogeneous with respect to the outcome of interest [124]. The values of observations in the resulting terminal tree-nodes are used as the basis of prediction. RF models combine multiple trees to reduce the variance of predictions and increase predictive performance [124,125]. Predictors may take the form of node or edge attributes. Valdes-Donoso *et al.* [68] used a RF to classify whether livestock movement occurred between pairs of nodes (farms or markets) as a function of geographical distance, node type mixing patterns (i.e. farm, market) and whether or not nodes were under shared ownership. This fitted model was then used to predict edges among nodes in the larger region, for which relevant node attributes were available.

3.6. Model validation

Adopting definitions by Porgo *et al.* [126], model validation (i.e. 'how well a model performs and how applicable the results are to a particular situation') was performed for around two-thirds (23/37) of models. We do not consider model calibration here (see §3.5). There was considerable variation in the methods by which model performance was assessed. This extended from the types of network properties considered, the methods of validation used, and the rigour to which this was carried out.

In terms of the types of validation used, 17 models were internally validated, while nine were externally validated. Approaches for external validation included splitting the data into training and validation sets (e.g. [68]), or through comparison with different datasets [15,56,65,67], such as for different time points [74,85,87]. A single GM was externally validated by assessing whether observed changes in livestock movements resulting from demand changes (i.e. closure of a terminal swine-processing facility) could be reproduced in the model [70]. Lastly, for two models, cross-validation was performed by comparing networks simulated by different models [65,91].

Regarding the types of network statistic considered, a third of models were validated by comparing structural network statistics of simulated and empirical networks (*n* = 14; electronic supplementary material, table S3). For example, model goodness-of-fit for ERGMs (*n* = 9) was assessed by comparing distributions of structural metrics not used for calibration such as in- and out-degree, geodesic distances, edgewise shared partnership and triad census.

Other models were internally validated at the level of the dyad (*n* = 6). Examples of approaches here included computing the predictive accuracy of binary or weighted edges based on, respectively, the area under the receiver operating characteristic curve, or correlation coefficients ([69] GM, [25] other statistical, [68] RF; Kong *et al.* [67] radiation model). Distributions of observed and predicted geographical distances between connected dyads were also sometimes compared ([86] other statistical, [68] RF, [56] ABM).

Alternatively, the outcomes of epidemics modelled on simulated networks were compared (*n* = 4) with either (i) epidemics modelled on empirical networks ([51] spatial), or (ii) empirical disease incidence. For example, the outputs of epidemics simulated on the pastoralist ABM by Kim *et al.* [65] were compared with annual disease incidence data. Meanwhile, Qiqi Yang *et al.* [58] assessed the statistical association between a GM-inferred poultry trade network and the geographical distribution of different avian influenza virus subtypes.

4. Discussion

In this systematic review, we present an overview of empirically informed, model-based approaches of network generation and inference that have been applied to simulate networks of contacts between livestock populations. We found 52 publications presenting 37 distinct models and seven model frameworks being used in this context. The increasing number of publications identified over the past decade illustrates the growing interest in this area. This reflects the considerable interest in applying network science to study the contact networks of livestock more broadly [2,3].

All models were applied to generate insights relevant to livestock diseases, with nearly half being used as inputs of infectious disease transmission models. However, the reviewed models varied greatly in their formulation, complexity and realism, use of data, and in the methods by which their performance was assessed. Consequently, we now turn to a comparison of model frameworks and discuss how their particular features can present opportunities and challenges in different use cases. Finally, we discuss issues and possible solutions around model assessment and validation.

A major application of reviewed mathematical models was to explore the relationship between network structure and disease transmission dynamics. Indeed, the relative simplicity of some of these models and, in particular, their ability to yield analytical solutions, lends them towards such applications. These types of models have consequently been applied extensively to explore the diffusion of phenomena on networks in the network literature [38,105,106]. This simplicity—in particular the ability of these models to be calibrated using few parameters—has also resulted in their application towards generating networks when empirical data are limited [47,48,51], or else totally absent, through the adoption of hypothesized parameter values (e.g. [99,100]). Mechanistic approaches, such as ABMs and radiation models, can also be used in cases where network data are unavailable but the processes underlying the formation of the network are understood and can be parametrized, i.e. based on first principles.

Notably, mechanistic models based on first principles may be more suitable for extrapolating beyond the data to which they were calibrated [127]. Hence, by altering their generative rules, such models can be applied to explore, for example, how counterfactual network configuration scenarios influence disease transmission dynamics [53]. Explicit modelling of the assumed generative mechanisms of the network further allows for an examination of its emergent properties. This makes it possible to explore realistic farm (or node) level disease control interventions that act to modify network structure [44]. Importantly, such approaches also allow complex adaptive properties of the system to be explored [113]. This includes agents' behavioural adaptations as a response to disease [49,54], or as an unintended consequence following regulatory changes or top-down interventions (e.g. [64]), as has been observed empirically [128–131].

Despite these important functions, purely mechanistic approaches commonly rely on calibration to select structural features (e.g. degree distribution and clustering coefficients) with no attempt to assess whether these features are necessary, or adequate, for representing an empirical network [36,39]. A comparative strength of statistical network

models lies in their utility for assessing which features are relevant for network generation, as well as allowing for a measure of the uncertainty of these estimates given the data [36,38–40,132]. This also allows networks to be simulated while accounting for and incorporating this uncertainty which, in the context of infectious disease modelling, can help avoid overfitting epidemic outcomes to observed networks [39,133,134]. Despite this utility, less than a third of models being applied to simulate networks for infectious disease modelling were statistical models, with the remaining being mechanistic. This may broadly reflect the contrasting applications of these different model groupings in our included studies; namely, the emphasis on hypothesis testing for the statistical models, particularly ERGMs which were the most well-represented model framework in this grouping.

As noted, the major application and strength of statistical models reviewed here was the inference of factors associated with network formation. An important limitation that was not addressed in the reviewed literature is that traditional statistical methods, such as GMs using OLS specifications, assume statistical independence between observations. Due to dependencies inherent to network data, such assumptions may not hold, potentially resulting in biased estimates and hence predictions [40,116,135,136]. While standard OLS specifications of GMs cannot explicitly model these dependencies, corrections have been proposed to account for the effects of assumptions about non-independence (summarized by Broekel *et al.* [116]). However, to our knowledge, these have not been used in GMs applied to livestock contact networks.

A major strength of ERGMs lies in their ability to explicitly model and account for such dependencies; networks can be modelled and simulated as a function of parameters describing structural characteristics (e.g. transitivity or mutuality effects) in addition to node and edge factors [120,136,137]. ERGMs are therefore a powerful means of assessing the statistical significance of a range of factors on edge formation, as well as for simulating networks from these parameterizations. In practice, however, it is not always possible to generate a well-fitting model. This can be due to issues with 'model degeneracy' which can occur when high correlations between network effects result in unrealistically dense or sparse networks [29,120,136].

We identified a single model applying RFs to predict and simulate livestock contact networks. More broadly across the network simulation modelling literature, a variety of supervised machine learning approaches have demonstrated high predictive utility when applied to the movements of humans [138,139] and wild animals [140]. Given increasingly widespread application of machine learning approaches across the network prediction literature and the growing volume and complexity of livestock data, including movement data [141], there is likely to be considerable scope in applying machine learning methods to predict and simulate livestock contact networks.

This review has highlighted significant variation in how models were calibrated and assessed. This is of course strongly reflective of the availability of empirical network data and the purpose or intended application of models. In the context of simulating networks relevant for epidemiological study, however, given the fundamental relationship between network structure and disease transmission dynamics, it is clear that meaningful and realistic outputs rely on simulated networks accurately reproducing epidemiologically relevant features of

the empirical networks. A remaining challenge then is understanding which structural features are epidemiologically relevant, and which we should therefore seek to reproduce. Indeed, the importance of these features may be highly disease and context specific [121,122,134,142]. Calibration and validation based on a few select network statistics is unlikely to be sufficient to reproduce networks exhibiting similar structure and diffusion patterns as their empirical counterparts [143–145]. Comparisons based on multiple structural characteristics are likely to be more robust, especially when the selection of these metrics is based on their relevance for diffusion processes, as is routine practice for ERGMs [121,122]. A highly valuable and interpretable form of validation, where data are available, is the comparison of epidemic outcomes on simulated and empirical networks. Comparison of simulated and observed disease incidence or prevalence is also particularly valuable, given that a transmission network is necessarily a subset of the potentially infectious contact network [146].

This review has some limitations. Despite our efforts to keep search terms broadly relevant to network simulation modelling, the lack of standardization in terminology means additional papers may have been missed using our search criteria. We have adopted the term 'network simulation model' from Bellerose et al. [35] and suggest its use in future publications on this topic. This would help to make this area of research more visible and avoid overlap with the related, yet distinct, context in which the term 'network modelling' is commonly applied, i.e. simulating disease spread on (empirical or simulated) networks. To keep the scope adequately focused and the synthesis feasible, we have focused on models which were used to simulate empirical-like and empirically informed contact networks of livestock populations. Hence, we highlight that this review does not present a complete compendium of all possible modelling frameworks, nor was it intended to. Alternative frameworks could be identified from the broader literature, such as from related reviews on network simulation models in other contexts [35–39].

This review serves to synthesize and categorize the heterogeneous group of models that have been applied to simulate the contact networks among livestock populations in the context of livestock disease epidemiology. Despite the important remaining challenges with model validation,

this review highlights a number of unique functions afforded by network simulation models which enable us to advance beyond simple descriptive analyses of livestock networks, or infectious disease modelling on empirical networks. With increasing recognition of the need for evidence-based approaches to livestock production and health, particularly in the context of multitudinous high-profile, and often economically devastating, livestock and zoonotic disease outbreaks in recent decades, it seems reasonable to assume that efforts towards livestock network data collection will continue to gain ground. The types of modelling approaches reviewed here are well positioned to derive key insights from this data. Furthermore, such models can be used to inform the design of future empirical studies and livestock tracking systems, in order to optimize their efficiency and utility in generating data needed for effective disease surveillance and control [26,28].

References

1. Fèvre EM, Bronsvoort BMdC, Hamilton KA, Cleaveland S. 2006 Animal movements and the spread of infectious diseases. *Trends Microbiol.* **14**, 125–131. (doi:10.1016/j.tim.2006.01.004)

2. Craft ME. 2015 Infectious disease transmission and contact networks in wildlife and livestock. *Phil. Trans. R. Soc. B* **370**, 20140107. (doi:10.1098/rstb.2014.0107)

3. Dubé C, Ribble C, Kelton D, McNab B. 2011 Introduction to network analysis and its implications for animal disease modelling. *OIE Revue Scientifique et Tech.* **30**, 425–436. (doi:10.20506/rst.30.2.2043)

4. Meyers LA, Newman MEJ, Pourbohloul B. 2006 Predicting epidemics on directed contact networks. *J. Theor. Biol.* **240**, 400–418. (doi:10.1016/j.jtbi.2005.10.004)

5. Newman MEJ. 2002 Spread of epidemic disease on networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **66**, 016128. (doi:10.1103/PhysRevE.66.016128)

6. Shirley MDF, Rushton SP. 2005 The impacts of network topology on disease spread. *Ecol. Compl.* **2**, 287–299. (doi:10.1016/j.ecocom.2005.04.005)

7. Fournié G, Guitian J, Desvaux S, Cuong VC, Dung DH, Pfeiffer DU, Mangtani P, Ghani AC. 2013 Interventions for avian influenza A (H5N1) risk management in live bird market networks. *Proc. Natl Acad. Sci. USA* **110**, 9177–9182. (doi:10.1073/pnas.1220815110)

8. Hardstaff JL, Häsler B, Rushton JR. 2015 Livestock trade networks for guiding animal health surveillance. *BMC Vet. Res.* **11**, 11–13. (doi:10.1186/s12917-015-0354-4)

9. Marquetoux N, Stevenson MA, Wilson P, Ridler A, Heuer C. 2016 Using social network analysis to inform disease control interventions. *Prev. Vet. Med.* **126**, 94–104. (doi:10.1016/j.prevetmed.2016.01.022)

10. Maurella C et al. 2019 Social network analysis and risk assessment: an example of introducing an exotic animal disease in Italy. *Microbial Risk Analysis* **13**, 100074. (doi:10.1016/j.mran.2019.04.001)

11. Napp S, García-Bocanegra I, Pagès N, Allepuz A, Alba A, Casal J. 2013 Assessment of the risk of a

bluetongue outbreak in Europe caused by *Culicoides* midges introduced through intracontinental transport and trade networks. *Med. Vet. Entomol.* **27**, 19–28. (doi:10.1111/j.1365-2915.2012.01016.x)

12. Gibbens JC, Wilesmith JW, Sharpe CE, Mansley LM, Michalopoulou E, Ryan JBM, Hudson M. 2001 Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months. *Vet. Record* **149**, 729–743. (doi:10.1136/vr.149.24.729)

13. Olugasa BO, Ijagbone IF. 2007 Pattern of spread of African swine fever in south-western Nigeria, 1997–2005. *Vet. Ital.* **43**, 621–628.

14. Ortiz-Pelaez A, Pfeiffer DU, Soares-Magalhães RJ, Guitian FJ. 2006 Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. *Prev. Vet. Med.* **76**, 40–55. (doi:10.1016/j.prevetmed.2006.04.007)

15. Yang Q *et al.* 2020 Assessing the role of live poultry trade in community-structured transmission of avian influenza in China. *Proc. Natl Acad. Sci. USA* **117**, 5949–5954. (doi:10.1073/pnas.1906954117)

16. Aznar MN, Stevenson MA, Zarich L, León EA. 2011 Analysis of cattle movements in Argentina, 2005. *Prev. Vet. Med.* **98**, 119–127. (doi:10.1016/j.prevetmed.2010.11.004)

17. Bigras-Poulin M, Barfod K, Mortensen S, Greiner M. 2007 Relationship of trade patterns of the Danish swine industry animal movements network to potential disease spread. *Prev. Vet. Med.* **80**, 143–165. (doi:10.1016/j.prevetmed.2007.02.004)

18. Guinat C, Durand B, Vergne T, Corre T, Rautureau S, Scoizec A, Lebouquin-Leneveu S, Guérin JL, Paul MC. 2020 Role of live-duck movement networks in transmission of avian influenza, France, 2016–2017. *Emerg. Infect. Dis.* **26**, 472–480. (doi:10.3201/eid2603.190412)

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

20. Apolloni A *et al.* 2018 Towards the description of livestock mobility in Sahelian Africa: some results from a survey in Mauritania. *PLoS ONE* **13**, e0191565. (doi:10.1371/journal.pone.0191565)

21. Baudon E *et al.* 2017 Analysis of swine movements in a province in Northern Vietnam and application in the design of surveillance strategies for infectious diseases. *Transbound Emerg. Dis.* **64**, 411–424. (doi:10.1111/tbed.12380)

22. Moyen N *et al.* 2021 Avian influenza transmission risk along live poultry trading networks in Bangladesh. *Sci. Rep.* **11**, 19962. (doi:10.1038/s41598-021-98989-4)

23. Noopataya S, Thongratsakul S, Poolkhet C. 2015 Social network analysis of cattle movement in Sukhothai Province, Thailand: a study to improve control measurements. *Vet. Med. Int.* **2015**, 587252. (doi:10.1155/2015/587252)

24. O'Hara K, Zhang R, Jung YS, Zhou X, Qian Y, Martínez-López B. 2020 Network analysis of swine shipments in China: the first step to inform disease surveillance and risk mitigation strategies. *Front. Vet. Sci.* **7**, 189. (doi:10.3389/fvets.2020.00189)

25. Lindström T, Grear DA, Buhnerkempe M, Webb CT, Miller RS, Portacci K, Wennergren U. 2013 A Bayesian approach for modeling cattle movements in the United States: scaling up a partially observed network. *PLoS ONE* **8**, e53432. (doi:10.1371/journal.pone.0053432)

26. Moon SA, Ferdousi T, Self A, Scoglio CM. 2019 Estimation of swine movement network at farm level in the US from the Census of Agriculture data. *Sci. Rep.* **9**, 1–14. (doi:10.1038/s41598-019-42616-w)

27. Wiltshire S, Zia A, Koliba C, Buccini G, Clark E, Merrill S, Smith J, Moegenburg S. 2019 Network meta-metrics: using evolutionary computation to identify effective indicators of epidemiological vulnerability in a livestock production system model. *J. Artif. Soc. Soc. Simul.* **22**, 1–8. (doi:10.18564/jasss.3991)

28. Chaters GL *et al.* 2019 Analysing livestock network data for infectious disease control: an argument for routine data collection in emerging economies. *Phil. Trans. R. Soc. B* **374**, 20180264. (doi:10.1098/rstb.2018.0264)

29. Relun A, Grosbois V, Alexandrov T, Sánchez-Vizcaíno JM, Waret-Szkuta A, Molia S, Etter EMC, Martínez-López B. 2017 Prediction of pig trade movements in different European production systems using exponential random graph models. *Front. Vet. Sci.* **4**, 27. (doi:10.3389/fvets.2017.00027)

30. Valdano E, Poletto C, Giovannini A, Palma D, Savini L, Colizza V. 2015 Predicting epidemic risk from past temporal contact data. *PLoS Computational Biology* **11**, e1004152–e1004152. (doi:10.1371/journal.pcbi.1004152)

31. Dawson PM, Werkman M, Brooks-Pollock E, Tildesley MJ. 2015 Epidemic predictions in an imperfect world: modelling disease spread with partial data. *Proc. R. Soc. B* **282**, 1–9.

32. Kukielka EA, Martínez-López B, Beltrán-Alcrudo D. 2017 Modeling the live-pig trade network in Georgia: implications for disease prevention and control. *PLoS ONE* **12**, e0178904. (doi:10.1371/journal.pone.0178904)

33. Nickbakhsh S, Matthews L, Bessell PR, Reid SWJ, Kao RR. 2011 Generating social network data using partially described networks: an example informing avian influenza control in the British poultry industry. *BMC Vet. Res.* **7**, 1–16. (doi:10.1186/1746-6148-7-66)

34. Page MJ *et al.* 2021 The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71. (doi:10.1136/bmj.n71)

35. Bellerose M, Zhu L, Hagan LM, Thompson WW, Randall LM, Malyuta Y, Salomon JA, Linas BP. 2019 A review of network simulation models of hepatitis C virus and HIV among people who inject drugs. *Int. J. Drug Policy* **88**, 102580. (doi:10.1016/j.drugpo.2019.10.006)

36. Goldenberg A, Zheng AX, Fienberg SE, Airoldi EM. 2010 A survey of statistical network models.

Foundations and Trends® in Machine Learning **2**, 129–233. (doi:10.1561/2200000005)

37. Keeling MJ, Eames KTD. 2005 Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307. (doi:10.1098/rsif.2005.0051)

38. Kolaczyk ED. 2009 *Statistical analysis of network data: methods and models*. Springer series in statistics. New York, NY: Springer-Verlag.

39. Welch D, Bansal S, Hunter DR. 2011 Statistical inference to advance network models in epidemiology. *Epidemics* **3**, 38–45. (doi:10.1016/j.epidem.2011.01.002)

40. Croft DP, Madden JR, Franks DW, James R. 2011 Hypothesis testing in animal social networks. *Trends Ecol. Evol.* **26**, 502–507. (doi:10.1016/j.tree.2011.05.012)

41. Hobson EA, Silk MJ, Fefferman NH, Larremore DB, Rombach P, Shai S, Pinter-Wollman N. 2021 A guide to choosing and implementing reference models for social network analysis. *Biol. Rev.* **96**, 2716–2734. (doi:10.1111/brv.12775)

42. R Core Team. 2020 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org.

43. Barabasi A.-L., Pósfai M. 2016 *Network science*, 1st edn. Cambridge, UK: Cambridge University Press.

44. Gates MC, Woolhouse MEJ. 2015 Controlling infectious disease through the targeted manipulation of contact network structure. *Epidemics* **12**, 11–19. (doi:10.1016/j.epidem.2015.02.008)

45. Bi Q, Goodman KE, Kaminsky J, Lessler J. 2019 What is machine learning? A primer for the epidemiologist. *Am. J. Epidemiol.* **188**, 2222–2239. (doi:10.1093/aje/kwz189)

46. Samuel AL. 1959 Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **3**, 210–229. (doi:10.1147/rd.33.0210)

47. Thakur KK, Sanchez J, Hurnik D, Poljak Z, Opps S, Revie CW. 2015 Development of a network based model to simulate the between-farm transmission of the porcine reproductive and respiratory syndrome virus. *Vet. Microbiol.* **180**, 212–222. (doi:10.1016/j.vetmic.2015.09.010)

48. Ferdousi T, Moon SA, Self A, Scoglio C. 2019 Generation of swine movement network and analysis of efficient mitigation strategies for African swine fever virus. *PLoS ONE* **14**, e0225785. (doi:10.1371/journal.pone.0225785)

49. Tago D, Hammitt JK, Thomas A, Raboisson D. 2016 The impact of farmers' strategic behavior on the spread of animal infectious diseases. *PLoS ONE* **11**, e0157450. (doi:10.1371/journal.pone.0157450)

50. Lennartsson J, Hakansson N, Wennergren U, Jonsson A. 2012 SpecNet: a spatial network algorithm that generates a wide range of specific structures. *PLoS ONE* **7**, e42679. (doi:10.1371/journal.pone.0042679)

51. Rossi G, Smith RL, Pongolini S, Bolzoni L. 2017 Modelling farm-to-farm disease transmission through personnel movements: from visits to contacts, and back. *Sci. Rep.* **7**, 2375. (doi:10.1038/s41598-017-02567-6)

52. Hu JH, Pei X, Sun GQ, Jin Z. 2021 Risk analysis of the transmission route for the African swine fever virus in Mainland China. *Front. Phys.* **9**, 785885.

53. Wiltshire SW. 2018 Using an agent-based model to evaluate the effect of producer specialization on the epidemiological resilience of livestock production networks. *PLoS ONE* **13**, e0194013. (doi:10.1371/journal.pone.0194013)

54. Bucini G, Merrill SC, Clark E, Moegenburg SM, Zia A, Koliba CJ, Wiltshire S, Trinity L, Smith JM. 2019 Risk attitudes affect livestock biosecurity decisions with ramifications for disease control in a simulated production system. *Front. Vet. Sci.* **6**, 196. (doi:10.3389/fvets.2019.00196)

55. Koliba C, Merrill SC, Zia A, Bucini G, Clark E, Shrum TR, Wiltshire S, Smith JM. 2022 Assessing strategic, tactical, and operational decision-making and risk in a livestock production chain through experimental simulation platforms. *Front. Vet. Sci.* **9**, 962788. (doi:10.3389/fvets.2022.962788)

56. Yang Q, Gruenbacher D, Stamm JLH, Brase GL, DeLoach SA, Amrine DE, Scoglio C. 2019 Developing an agent-based model to simulate the beef cattle production and transportation in southwest Kansas. *Phys. A Stat. Mech. Appl.* **526**, 120856. (doi:10.1016/j.physa.2019.04.092)

57. Yang Q, Gruenbacher DM, Brase GL, Stamm JLH, DeLoach SA, Scoglio CM. 2021 Simulating human behavioral changes in livestock production systems during an epidemic: the case of the US beef cattle industry. *PLoS ONE* **16**, e0253498. (doi:10.1371/journal.pone.0253498)

58. Yang Q, Gruenbacher DM, Heier Stamm JL, Amrine DE, Brase GL, DeLoach SA, Scoglio CM. 2020 Impact of truck contamination and information sharing on foot-and-mouth disease spreading in beef cattle production systems. *PLoS ONE* **15**, e0240819. (doi:10.1371/journal.pone.0240819)

59. Ross CW, Glass RJ, Harger J, Conrad S, Zagonel AA, Beyeler W, Finley M. 2011 Development of an agent-based epidemiological model of beef cattle. In *29th Int. Conf. of the Systems Dynamics Society, Washington, DC, 24–28 July.* Sandia National Laboratories.

60. Liu H, Schumm P, Lyubinin A, Scoglio C. 2012 Epirur-cattle: a spatially explicit agent-based simulator of beef cattle movements. *Procedia Comp. Sci.* **9**, 857–865. (doi:10.1016/j.procs.2012.04.092)

61. Ansari S, Heitzig J, Brzoska L, Lentz HHK, Mihatsch J, Fritzemeier J, Moosavi MR. 2021 A temporal network model for livestock trade systems. *Front. Vet. Sci.* **8**, 766547. (doi:10.3389/fvets.2021.766547)

62. Brock J, Lange M, Tratalos JA, More SJ, Guelbenzu-Gonzalo M, Graham DA, Thulke HH. 2021 A large-scale epidemiological model of BoHV-1 spread in the Irish cattle population to support decision-making in conformity with the European Animal Health Law. *Prev. Vet. Med.* **192**, 105375. (doi:10.1016/j.prevetmed.2021.105375)

63. Knight MA, White PCL, Hutchings MR, Davidson RS, Marion G. 2021 Generative models of network dynamics provide insight into the effects of trade on endemic livestock disease. *R. Soc. Open Sci.* **8**, 201715. (doi:10.1098/rsos.201715)

64. Knight MA, Hutchings MR, White PCL, Davidson RS, Marion G. 2022 A mechanistic model captures livestock trading, disease dynamics, and compensatory behaviour in response to control measures. *J. Theor. Biol.* **539**, 111059. (doi:10.1016/j.jtbi.2022.111059)

65. Kim H, Xiao N, Moritz M, Garabed R, Pomeroy LW. 2016 Simulating the transmission of foot-and-mouth disease among mobile herds in the Far North Region, Cameroon. *J. Artif. Soc. Soc. Simul.* **19**, 6. (doi:10.18564/jasss.3064)

66. Pomeroy LW, Kim H, Xiao N, Moritz M, Garabed R. 2019 Network analyses to quantify effects of host movement in multilevel disease transmission models using foot and mouth disease in Cameroon as a case study. *PLoS Comput. Biol.* **15**, e1007184. (doi:10.1371/journal.pcbi.1007184)

67. Kong L, Zhou X, Duan M, Yang X, Zhou C, Liao Y. 2022 A dataset of a proxy variable for the poultry trade flows in China. *Sci. Data* **9**, 690. (doi:10.1038/s41597-022-01793-6)

68. Valdes-Donoso P, VanderWaal K, Jarvis LS, Wayne SR, Perez AM. 2017 Using machine learning to predict swine movements within a regional program to improve control of infectious diseases in the US. *Front. Vet. Sci.* **4**, 2. (doi:10.3389/fvets.2017.00002)

69. Nicolas G, Apolloni A, Coste C, Wint GRW, Lancelot R, Gilbert M. 2018 Predictive gravity models of livestock mobility in Mauritania: the effects of supply, demand and cultural factors. *PLoS ONE* **13**, e0199547. (doi:10.1371/journal.pone.0199547)

70. Blair B, Lowe J. 2022 The application of an augmented gravity model to measure the effects of a regionalization of potential risk distribution of the US cull sow market. *Vet. Sci.* **9**, 215. (doi:10.3390/vetsci9050215)

71. Ortiz-Pelaez A, Ashenafi G, Roger F, Waret-Szkuta A. 2012 Can geographical factors determine the choices of farmers in the Ethiopian Highlands to trade in livestock markets? *PLoS ONE* **7**, e30710. (doi:10.1371/journal.pone.0030710)

72. Poolkhet C, Makita K, Thongratsakul S, Leelehapongsathon K. 2018 Exponential random graph models to evaluate the movement of backyard chickens after the avian influenza crisis in 2004–2005, Thailand. *Prev. Vet. Med.* **158**, 71–77. (doi:10.1016/j.prevetmed.2018.07.015)

73. Belkhiria J, Lo MM, Sow F, Martínez-López B, Chevalier V. 2019 Application of exponential random graph models to determine nomadic herders' movements in Senegal. *Transboundary Emerg. Dis.* **66**, 1642–1652. (doi:10.1111/tbed.13198)

74. Hammami P, Widgren S, Grosbois V, Apolloni A, Rose N, Andraud M. 2022 Complex network analysis to understand trading partnership in French swine production. *PLoS ONE* **17**, e0266457. (doi:10.1371/journal.pone.0266457)

75. Lee HS, Thakur KK, Pham-Thanh L, Dao TD, Bui AN, Bui VN, Quang HN. 2021 A stochastic network-based model to simulate farm-level transmission of African swine fever virus in Vietnam. *PLoS ONE* **16**, e0247770. (doi:10.1371/journal.pone.0247770)

76. Brommesson P, Sellman S, Beck-Johnson L, Hallman C, Murrieta D, Webb CT, Miller RS, Portacci K, Lindström T. 2021 Assessing intrastate shipments from interstate data and expert opinion. *R. Soc. Open Sci.* **8**, 192042. (doi:10.1098/rsos.192042)

77. Buhnerkempe MG et al. 2014 The impact of movements and animal density on continental scale cattle disease outbreaks in the United States. *PLoS ONE* **9**, e91724. (doi:10.1371/journal.pone.0091724)

78. Buhnerkempe MG, Grear DA, Portacci K, Miller RS, Lombard JE, Webb CT. 2013 A national-scale picture of U.S. cattle movements obtained from Interstate Certificate of Veterinary Inspection data. *Prev. Vet. Med.* **112**, 318–329. (doi:10.1016/j.prevetmed.2013.08.002)

79. Gilbertson K et al. 2022 The importance of livestock demography and infrastructure in driving foot and mouth disease dynamics. *Life (Basel)* **12**, 1604. (doi:10.3390/life12101604)

80. Gorsich EE, McKee CD, Grear DA, Miller RS, Portacci K, Lindström T, Webb CT. 2018 Model-guided suggestions for targeted surveillance based on cattle shipments in the U.S. *Prev. Vet. Med.* **150**, 52–59. (doi:10.1016/j.prevetmed.2017.12.004)

81. Gorsich EE, Luis AD, Buhnerkempe MG, Grear DA, Portacci K, Miller RS, Webb CT. 2016 Mapping U.S. cattle shipment networks: spatial and temporal patterns of trade communities from 2009 to 2011. *Prev. Vet. Med.* **134**, 82–91. (doi:10.1016/j.prevetmed.2016.09.023)

82. Kao SYZ, VanderWaal K, Enns EA, Craft ME, Alvarez J, Picasso C, Wells SJ. 2018 Modeling cost-effectiveness of risk-based bovine tuberculosis surveillance in Minnesota. *Prev. Vet. Med.* **159**, 1–11. (doi:10.1016/j.prevetmed.2018.08.011)

83. Sellman S, Beck-Johnson LM, Hallman C, Miller RS, Bonner KAO, Portacci K, Webb CT, Lindström T. 2022 Modeling U.S. cattle movements until the cows come home: who ships to whom and how many? *Comput. Elect. Agricult.* **203**, 107483. (doi:10.1016/j.compag.2022.107483)

84. Tsao K et al. 2020 Effects of regional differences and demography in modelling foot-and-mouth disease in cattle at the national scale. *Interface Focus* **10**, 20190054. (doi:10.1098/rsfs.2019.0054)

85. Sellman S, Beck-Johnson LM, Hallman C, Miller RS, Owers Bonner KA, Portacci K, Webb CT, Lindström T. 2022 Modeling nation-wide U.S. swine movement networks at the resolution of the individual premises. *Epidemics* **41**, 100636. (doi:10.1016/j.epidem.2022.100636)

86. Lindström T, Sisson SA, Nöremark M, Jonsson A, Wennergren U. 2009 Estimation of distance related probability of animal movements between holdings and implications for disease spread modeling. *Prev. Vet. Med.* **91**, 85–94. (doi:10.1016/j.prevetmed.2009.05.022)

87. Brommesson P, Wennergren U, Lindström T. 2016 Spatiotemporal variation in distance dependent animal movement contacts: one size doesn't fit all. *PLoS ONE* **11**, e0164008. (doi:10.1371/journal.pone.0164008)

88. Lindström T, Lewerin SS, Wennergren U. 2012 Influence on disease spread dynamics of herd

characteristics in a structured livestock industry. *J. R. Soc. Interface* **9**, 1287–1294. (doi:10.1098/rsif. 2011.0625)

89. Lindström T, Sisson SA, Lewerin SS, Wennergren U. 2011 Bayesian analysis of animal movements related to factors at herd and between herd levels: implications for disease spread modeling. *Prev. Vet. Med.* **98**, 230–242. (doi:10.1016/j.prevetmed.2010.11.005)

90. Lindström T, Sisson SA, Lewerin SS, Wennergren U. 2010 Estimating animal movement contacts between holdings of different production types. *Prev. Vet. Med.* **95**, 23–31. (doi:10.1016/j. prevetmed.2010.03.002)

91. Xiao N, Cai S, Moritz M, Garabed R, Pomeroy LW. 2015 Spatial and temporal characteristics of pastoral mobility in the Far North Region, Cameroon: data analysis and modeling. *PLoS ONE* **10**, e0131697. (doi:10.1371/journal.pone.0131697)

92. Schumm P, Scoglio C, Scott HM. 2015 An estimation of cattle movement parameters in the Central States of the US. *Comput. Electr. Agricult.* **116**, 191–200. (doi:10.1016/j.compag.2015.06.016)

93. Ezanno P, Arnoux S, Joly A, Vermesse R. 2022 Rewiring cattle trade movements helps to control bovine paratuberculosis at a regional scale. *Prev. Vet. Med.* **198**, 105529. (doi:10.1016/j.prevetmed. 2021.105529)

94. Hidano A, Carpenter TE, Stevenson MA, Gates MC. 2016 Evaluating the efficacy of regionalisation in limiting high-risk livestock trade movements. *Prev. Vet. Med.* **133**, 31–41. (doi:10.1016/j.prevetmed. 2016.09.015)

95. Mohr S, Deason M, Churakov M, Doherty T, Kao RR. 2018 Manipulation of contact network structure and the impact on foot-and-mouth disease transmission. *Prev. Vet. Med.* **157**, 8–18. (doi:10.1016/j. prevetmed.2018.05.006)

96. Hoscheit P, Geeraert S, Beaunée G, Monod H, Gilligan CA, Filipe JAN, Vergu E, Moslonka-Lefebvre M. 2017 Dynamical network models for cattle trade: towards economy-based epidemic risk assessment. *J. Complex Netw.* **5**, 604–624. (doi:10.1093/comnet/cnw026)

97. Sottile MJ, Iles RA, McConnel CS, Amram O, Lofgren E. 2021 PastoralScape: an environment-driven model of vaccination decision making within pastoralist groups in East Africa. *J. Artif. Soc. Soc. Simul.* **24**, 1–11. (doi:10.18564/jasss.4686)

98. Tennant WSD *et al.* 2021 Modelling the persistence and control of Rift Valley fever virus in a spatially heterogeneous landscape. *Nat. Commun.* **12**, 5593. (doi:10.1038/s41467-021-25833-8)

99. Scoglio CM, Bosca C, Riad MH, Sahneh FD, Britch SC, Cohnstaedt LW, Linthicum KJ. 2016 Biologically informed individual-based network model for Rift Valley fever in the US and evaluation of mitigation strategies. *PLoS ONE* **11**, e0162759. (doi:10.1371/ journal.pone.0162759)

100. Sekamatte M, Riad MH, Tekleghiorghis T, Linthicum KJ, Britch SC, Richt JA, Gonzalez JP, Scoglio CM. 2019 Individual-based network model for Rift Valley fever in Kabale District, Uganda. *PLoS ONE* **14**, e0202721. (doi:10.1371/journal.pone. 0202721)

101. Marsot M, Canini L, Janicot S, Lambert J, Vergu E, Durand B. 2022 Predicting veal-calf trading events in France. *Prev. Vet. Med.* **209**, 105782. (doi:10. 1016/j.prevetmed.2022.105782)

102. Erdős P, Rényi A. 1960 On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–60.

103. Erdős P, Rényi A. 1959 On random graphs. *Publ. Math. Debrecen* **6**, 290–297.

104. Gilbert EN. 1959 Random graphs. *Ann. Math. Stat.* **30**, 1141–1144.

105. Albert R, Barabasi AL. 2002 Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97. (doi:10.1103/RevModPhys.74.47)

106. Newman MEJ. 2003 The structure and function of complex networks. *SIAM Rev.* **45**, 167–256. (doi:10. 1137/S003614450342480)

107. Bollobás B. 1980 A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combinatorics* **1**, 311–316. (doi:10.1016/S0195-6698(80)80030-8)

108. Britton T, Deijfen M, Martin-Löf A. 2006 Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* **124**, 1377–1397. (doi:10. 1007/s10955-006-9168-x)

109. Barabasi A-L, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512. (doi:10.1126/science.286.5439.509)

110. Fielding HR, McKinley TJ, Silk MJ, Delahay RJ, McDonald RA. 2019 Contact chains of cattle farms in Great Britain. *R. Soc. Open Sci.* **6**, 180719. (doi:10. 1098/rsos.180719)

111. Watts DJ, Strogatz SH. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442. (doi:10.1038/30918)

112. Boender GJ, Meester R, Gies E, De Jong MCM. 2007 The local threshold for geographical spread of infectious diseases between farms. *Prev. Vet. Med.* **82**, 90–101. (doi:10.1016/j.prevetmed.2007.05.016)

113. Grimm V *et al.* 2006 A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* **198**, 115–126. (doi:10.1016/j. ecolmodel.2006.04.023)

114. Lanzas C, Chen S. 2015 Complex system modelling for veterinary epidemiology. *Prev. Vet. Med.* **118**, 207–214. (doi:10.1016/j.prevetmed.2014.09.012)

115. Simini F, González MC, Maritan A, Barabási AL. 2012 A universal model for mobility and migration patterns. *Nature* **484**, 96–100. (doi:10.1038/ nature10856)

116. Broekel T, Balland PA, Burger M, van Oort F. 2014 Modeling knowledge networks in economic geography: a discussion of four methods. *Ann. Reg. Sci.* **53**, 423–452. (doi:10.1007/s00168-014-0616-2)

117. Tinbergen JJ, 1962. *Shaping the world economy: suggestions for an international economic policy*. New York, NY: Twentieth Century Fund.

118. Conlan AJ, Klepac P, Kucharski AJ, Kissler S, Tang ML, Fry H, Gog JR. 2021 Human mobility data from the BBC pandemic project. *medRxiv*. (doi:10.1101/2021. 02.19.21252079)

119. Robins G, Snijders T, Wang P, Handcock M, Pattison P. 2007 Recent developments in exponential random graph (*p**) models for social networks. *Social Netw.* **29**, 192–215. (doi:10.1016/j.socnet. 2006.08.003)

120. Lusher D, Koskinen J, Robins G. 2012 *Exponential random graph models for social networks: theory, methods, and applications*. (Structural analysis in the social sciences no. 35). Cambridge, UK: Cambridge University Press. (doi:10.1017/ CBO9780511894701)

121. Hunter DR, Goodreau SM, Handcock MS. 2008 Goodness of fit of social network models. *J. Am. Stat. Assoc.* **103**, 248–258. (doi:10.1198/ 016214507000000446)

122. Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M. 2008 ergm: a package to fit, simulate and diagnose exponential-family models for networks. *J. Stat. Softw.* **24**, nihpa54860.

123. Krivitsky PN, Handcock MS. 2014 A separable model for dynamic networks. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **76**, 29–46. (doi:10.1111/rssb.12014)

124. Kuhn M, Johnson K. 2013 *Applied predictive modeling*. New York, NY: Springer New York.

125. Lichtenwalter RN, Lussier JT, Chawla NV. 2010 New perspectives and methods in link prediction In *Proc. of the 16th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'10), Washington, DC, 24–28 July.* Association for Computing Machinery. (doi:10.1145/1835804.1835837)

126. Porgo TV, Norris SL, Salanti G, Johnson LF, Simpson JA, Low N, Egger M, Althaus CL. 2019 The use of mathematical modeling studies for evidence synthesis and guideline development: a glossary. *Res. Syn. Meth.* **10**, 125–133. (doi:10.1002/jrsm. 1333)

127. Bolker BM. 2008 *Ecological models and data in R*. Princeton, NJ: Princeton University Press.

128. Christley RM, Robinson SE, Moore B, Setzkorn C, Donald I. 2011 Responses of farmers to introduction in England and Wales of pre-movement testing for bovine tuberculosis. *Prev. Vet. Med.* **100**, 126–133. (doi:10.1016/j.prevetmed.2011.02.005)

129. Gates MC, Volkova VV, Woolhouse MEJ. 2013 Impact of changes in cattle movement regulations on the risks of bovine tuberculosis for Scottish farms. *Prev. Vet. Med.* **108**, 125–136. (doi:10.1016/j.prevetmed.2012.07.016)

130. Robinson SE, Everett MG, Christley RM. 2007 Recent network evolution increases the potential for large epidemics in the British cattle population. *J. R. Soc. Interface* **4**, 669–674. (doi:10.1098/rsif.2007.0214)

131. Vernon MC, Keeling MJ. 2012 Impact of regulatory perturbations to disease spread through cattle movements in Great Britain. *Prev. Vet. Med.* **105**, 110–117. (doi:10.1016/j.prevetmed.2011.12.016)

132. Robins G, Pattison P, Kalish Y, Lusher D. 2007 An introduction to exponential random graph (*p**) models for social networks. *Soc. Netw.* **29**, 173–191. (doi:10.1016/j.socnet.2006.08.002)

133. Reynolds JJH, Hirsch BT, Gehrt SD, Craft ME. 2015 Raccoon contact networks predict seasonal susceptibility to rabies outbreaks and limitations of vaccination. *J. Anim. Ecol.* **84**, 1720–1731. (doi:10. 1111/1365-2656.12422)

134. Rolls DA, Wang P, McBryde E, Pattison P, Robins G. 2015 A simulation study comparing epidemic

dynamics on exponential random graph and edge-triangle configuration type contact network models. *PLoS ONE* **10**, e0142181. (doi:10.1371/journal.pone.0142181)

135. Lubbers MJ, Snijders TAB. 2007 A comparison of various approaches to the exponential random graph model: a reanalysis of 102 student networks in school classes. *Soc. Netw.* **29**, 489–507. (doi:10.1016/j.socnet.2007.03.002)

136. Silk MJ, Croft DP, Delahay RJ, Hodgson DJ, Weber N, Boots M, McDonald RA. 2017 The application of statistical network models in disease research. *Methods Ecol. Evol.* **8**, 1026–1041. (doi:10.1111/2041-210X.12770)

137. Goodreau SM, Kitts JA, Morris M. 2009 Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography* **46**, 103–125. (doi:10.1353/dem.0.0045)

138. Robinson C, Dilkina B. 2017 A machine learning approach to modeling human migration. In *Proc. of the 1st ACM SIGCAS Conf. on Computing and Sustainable Societies*. New York, NY: Association for Computing Machinery. (doi:10.1145/3209811.3209868)

139. Spadon G, Carvalho ACPLFde, Rodrigues Jr JF, Alves LGA. 2019 Reconstructing commuters network using machine learning and urban indicators. *Sci. Rep.* **9**, 1–13. doi:10.1038/s41598-019-48295-x

140. Wijeyakulasuriya DA, Eisenhauer EW, Shaby BA, Hanks EM. 2020 Machine learning for modeling animal movement. *PLoS ONE* **15**, e0235750. (doi:10.1371/journal.pone.0235750)

141. VanderWaal K, Morrison RB, Neuhauser C, Vilalta C, Perez AM. 2017 Translating big data into smart data for veterinary epidemiology. *Front. Vet. Sci.* **4**, 110.

142. Pellis L, Ball F, Bansal S, Eames K, House T, Isham V, Trapman P. 2015 Eight challenges for network epidemic models. *Epidemics Challenges Model. Infect. Dis. Dyn.* **10**, 58–62. (doi:10.1016/j.epidem.2014.07.003)

143. Ball F, Britton T, Sirl D. 2013 A network with tunable clustering, degree correlation and degree distribution, and an epidemic thereon. *J. Math. Biol.* **66**, 979–1019. (doi:10.1007/s00285-012-0609-7)

144. Green DM, Kiss IZ. 2010 Large-scale properties of clustered networks: implications for disease dynamics. *J. Biol. Dyn.* **4**, 431–445. (doi:10.1080/17513758.2010.487158)

145. Ritchie M, Berthouze L, House T, Kiss IZ. 2014 Higher-order structure and epidemic dynamics in clustered networks. *J. Theor. Biol.* **348**, 21–32. (doi:10.1016/j.jtbi.2014.01.025)

146. Friedman SR, Aral S. 2001 Social networks, risk-potential networks, health, and disease. *J. Urban Health* **78**, 411–418. (doi:10.1093/jurban/78.3.411)

147. Leung WTM, Rudge JW, Fournié G. 2023 Simulating contact networks for livestock disease epidemiology: a systematic review. Figshare. (doi:10.6084/m9.figshare.c.6631153)