



HAL
open science

Modélisation et optimisation de la prise de décision pour la gestion de maladies infectieuses se propageant sur des réseaux de métapopulations animales

Lina Cristancho Fajardo

► To cite this version:

Lina Cristancho Fajardo. Modélisation et optimisation de la prise de décision pour la gestion de maladies infectieuses se propageant sur des réseaux de métapopulations animales. Statistiques [math.ST]. Université Paris-Saclay, 2022. Français. NNT : 2022UPASB008 . tel-03663416

HAL Id: tel-03663416

<https://pastel.hal.science/tel-03663416>

Submitted on 10 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modelling and optimising decision-making
for the control of infectious diseases
spreading on animal metapopulation
networks

*Modélisation et optimisation de la prise de décision pour la
gestion de maladies infectieuses se propageant sur des
réseaux de métapopulations animales*

Thèse de doctorat de l'université Paris-Saclay

Ecole doctorale n°581, Agriculture, Alimentation, Biologie,
Environnement, Santé (ABIES)

Spécialité de doctorat: Mathématiques appliquées

Graduate School : Biosphera, Référent : AgroParisTech

Thèse préparée dans les unités de recherche **MaIAGE (Université Paris-Saclay, INRAE)** et **BIOEPAR (INRAE, Oniris)**, sous la direction de **Elisabeta VERGU**, Directrice de recherche, et la co-direction de **Pauline EZANNO**, Directrice de recherche

Thèse soutenue à Paris-Saclay, le 17 mars 2022, par

Lina CRISTANCHO FAJARDO

Composition du jury

Nicolas VAYATIS

Professeur, ENS Paris-Saclay

Ludovic MAILLERET

Directeur de recherche, INRAE Centre PACA

Rowland KAO

Professor, University of Edinburgh

Anne CORI

Lecturer, Imperial College London

Guillaume FOURNIÉ

Chercheur, Royal Veterinary College London

Elisabeta VERGU

Directrice de recherche, INRAE (Université Paris-Saclay)

Pauline EZANNO

Directrice de recherche, INRAE Centre Pays de la Loire

Président

Rapporteur & Examineur

Rapporteur & Examineur

Examinatrice

Examineur

Directrice de thèse

Co-Directrice de thèse

Titre: Modélisation et optimisation de la prise de décision pour la gestion de maladies infectieuses se propageant sur des réseaux de métapopulations animales

Mots clés: modélisation épidémiologique, réseau de métapopulation, optimisation, comportement humain, allocation de ressources, gestion de maladies infectieuses.

Résumé: Cette thèse porte sur la modélisation et l'optimisation de la maîtrise d'un agent pathogène se propageant dans une métapopulation d'animaux d'élevage via un réseau d'échanges commerciaux, en tenant compte des processus de décision concernant l'adoption de mesures de maîtrise.

D'une part, concernant la prise de décision des éleveurs, un modèle stochastique intégrant la dynamique intra-troupeau de la maladie (composantes démographiques et réseau d'échanges) et la dynamique des décisions des éleveurs a été développé et exploré par simulations intensives et analyses de sensibilité. En particulier, un mécanisme de décision dynamique qui tient compte du comportement aléatoire des éleveurs, de leur apprentissage et de la dynamique d'imitation stratégique a été proposé. Le modèle a été formalisé pour une dynamique d'infection théorique (modèle SIR), et une mesure de contrôle spécifique (vaccination). Ce premier modèle a été étendu et adapté à une maladie réelle, la BVD (diarrhée virale bovine), où tant la transmission de l'agent pathogène que les échanges d'informations

entre éleveurs peuvent passer par le réseau d'échanges, mais aussi par un voisinage géographique.

D'autre part, dans une perspective plus générale, il a été supposé qu'un planificateur social central cherchait à allouer dynamiquement et de manière optimale une ressource limitée entre les différentes sous-populations d'un réseau de métapopulation donné, afin de réduire la propagation d'un agent pathogène. L'approche, basée sur des scores permettant de classer les sous-populations pour l'allocation des ressources, a été formalisée pour le modèle épidémiologique théorique considéré dans la première partie de la thèse, pour deux mesures différentes (vaccination et traitement). De nouveaux scores ont été obtenus par l'adaptation d'une approche d'optimisation gloutonne au cadre des métapopulations. Par le biais de simulations, les performances de ces nouveaux scores ont été comparées à celles de plusieurs heuristiques qui pourraient être appropriées lorsque le réseau de métapopulation correspond à un réseau de commerce d'animaux.

Title: Modelling and optimising decision-making for the control of infectious diseases spreading on animal metapopulation networks

Keywords: epidemiological modelling, metapopulation network, optimisation, human-behaviour, resource allocation, infectious disease control.

Abstract: This thesis focuses on the modeling and optimisation of the control of a pathogen spreading in a livestock metapopulation via a trade network, taking into account the decision processes concerning the adoption of control measures.

On the one hand, regarding farmers' decision-making, a stochastic model integrating the intra-herd disease dynamics (demographic components and trade network) and the dynamics of farmers' decisions was developed and explored through intensive simulations and sensitivity analyses. In particular, a dynamic decision-mechanism that accounts for farmers' random behaviour, their learning and strategic imitation dynamics was proposed. The model was formalised for a theoretical infection dynamics (SIR model), and a specific control measure (vaccination). This first model was extended and adapted to a real-life disease, BVD (bovine viral diarrhoea), where both the pathogen spread and the exchange of information between farmers can oc-

cur through the trade network, but also through a geographical neighbourhood.

On the other hand, from a more general perspective, it was assumed that a central social planner sought to dynamically and optimally allocate a limited resource among the different sub-populations of a given metapopulation network, in order to reduce pathogen spread. The approach, based on scores allowing to rank subpopulations for resource allocation, was formalised for the theoretical epidemiological model considered in the first part of the thesis, for two different measures (vaccination and treatment). New scores were obtained by adapting a greedy optimisation approach to the metapopulation framework. Through simulations, the performances of these new scores were compared to those of several heuristics that might be appropriate when the metapopulation network corresponds to an animal trade network.

Acknowledgements

In these acknowledgements I would like to thank each person that has contributed to me becoming a doctor and that has shaped my life in some way or another. I apologise in advance if I forget anyone (which I probably will).

Je voudrais remercier tout d'abord mes directrices, Eliza et Pauline, qui m'ont guidé tout au long de la thèse tant sur des questions scientifiques comme sur le plan administratif (qui n'a pas été sans complications) et qui m'ont toujours impulsé pour faire mieux. Un merci particulier pour l'encouragement pendant les confinements, la rédaction du manuscrit et la préparation de la soutenance. Merci aussi pour avoir de la confiance sur moi en me proposant de continuer à travailler ensemble un peu plus.

I thank Ludovic Mailleret and Rowland Kao for their detailed, benevolent and constructive review of my thesis manuscript, which greatly helped me for the preparation of the defense. Un grand merci aussi aux examinateurs: Anne Cori, Guillaume Fournié et Nicolas Vayatis pour faire partie du jury pour la soutenance.

Merci aux personnes de BIOEPAR avec qui j'ai pu interagir et/ou travailler, et également aux personnes dans l'unité MaIAGE, l'unité où j'ai passé ces trois dernières années. En particulier, merci aux doctorants avec lesquels j'ai le plus partagé ce parcours: Romain, Henri, Léo, Ajmal, et aussi à ceux qui sont plus récemment arrivés. Merci en spécial à Romain, Henri et Madeleine de m'avoir aidé à gérer le stress de la fin de thèse.

Je voudrais aussi remercier les membres de mon comité de thèse: Nathalie Peyrard, Régis Sabbadin, Argyris Kalogeratos, Alban Thomas et Romulus Breban, pour les discussions riches et variées concernant l'avancement de ma thèse. Ces discussions m'ont permis de faire le point chaque année, ainsi que de mieux orienter mes choix scientifiques. Merci en particulier à Nathalie et Régis, qui m'ont encadrée pour mon stage de M2, grâce auquel j'ai pu confirmer mon intérêt pour faire une thèse. Plus en général, merci à toute l'unité MIAT pour tout ce que vous m'avez appris sur la France (et pas n'importe laquelle).

Merci à tous mes professeurs du M2. En particulier à Pierre Alquier et Christophe Giraud qui à l'époque coordonnaient le Master Stat ML, grâce auquel j'ai compris que je voulais faire de la recherche. J'ai beaucoup apprécié votre qualité pédagogique et je vous remercie pour votre encouragement et votre aide pendant la recherche de stage et de sujet de thèse. Merci aussi à mes autres professeurs de l'ENSAE, avec qui j'ai appris énormément. En particulier merci à Pierre Alquier (à nouveau), Nicolas Chopin, Arnak Dalalyan, Alexandre Tsybakov. C'était un honneur de suivre vos cours. Je remercie également tous les amis que j'ai pu me faire pendant ma scolarité à l'ENSAE.

Merci aussi à toutes les personnes qui m'ont permis de venir en France. A l'équipe de ParisTech, en particulier à Vincent Brenier, pour avoir vu des qualités dans mon examen et mon entretien, et pour toute son aide administrative. Merci d'avoir eu confiance sur le fait que je pouvais passer d'un niveau nul en français à réussir un examen B2 en moins de 6 mois. Ce que j'ai réussi grâce à ma prof de français, Carol, qui a accepté que je rejoigne son groupe en cours de route alors qu'elle n'était pas du tout obligée. Merci de nous avoir préparés non seulement pour l'examen mais aussi plus largement pour notre vie en France. Merci à tous mes camarades du cours de français qui m'ont aidée à rapidement rattraper le groupe. Gracias a las facultades de Ciencias e

Ingeniería por toda su ayuda administrativa. En particular a esta última por haber aceptado que yo siguiera la preparación al examen de ParisTech y los cursos de francés de la facultad.

Gracias a los profesores que tuve durante la carrera de estadística. A Liliana Blanco, en cuyo curso de procesos estocásticos comencé a pensar que me podía interesar la investigación y la modelación matemática de epidemias. A Leonardo Trujillo, en especial por su recomendación para la estancia de investigación en Purdue (a pesar de que finalmente decidiera venir a Europa). A Oscar Melo, por su excelente curso de diseño de experimentos, por haberme aconsejado cuando tenía dudas sobre venir a Francia, y por su recomendación para ser aceptada y becada en el Master Statistics and Machine Learning (beca gracias a la cual pude sostenerme durante el segundo año en Francia). Gracias también a Sergio Calderón por su ayuda en cuanto a la homologación de materias con la UNAL. Gracias a tantos otros profesores que contribuyeron a mi formación: José Alfredo Jimenez, Martha Bohorquez, German Fonseca, entre otros. Gracias también a todas las demás personas que crucé en esta primera etapa universitaria. A mis compañeros y profesores de salsa y danzas afrocolombianas, con quienes aprendí que la danza podía también formarme como persona. Gracias en particular a mis amigos José, Emely y Adriana por los momentos de estudio, de canto, de baile y de compartir chismes.

Gracias a los amigos que he podido hacer en Francia. A Michelle, que fue como mi mamá francesa durante el primer año aquí. A mis amigos colombianos: Juan, Raúl, Santiago, Maria Juliana y David, con quienes afronté el proceso de adaptación a Francia y a las grandes escuelas de ingenieros. A Juan y Raúl, por ser mi mayor soporte durante el primer año. A Juan, por su apoyo durante la thèse. A Juan y Santiago, por toda su ayuda para la soutenance.

Gracias a mis demás amigos que me quieren y apoyan desde la distancia. A mis amigas del colegio, por las travesuras y por alegrarse tanto o incluso más que yo con mis logros académicos. Gracias en general a todas la niñas, ya mujeres, que crucé en esta etapa y para las que espero puedan cumplir sus sueños.

Gracias a mi familia (ese batallón de tíos y tías, primos y primas), que sé que siempre me tiene en sus pensamientos. En particular, gracias a mis abuelos: Agripina, Elizabeth, Diógenes y Luis Alberto. Aunque ya no están, es gracias a ellos, quienes impulsaron sus respectivas familias a pesar de las dificultades de la época y del contexto, que mis padres se convirtieron en los maravillosos seres humanos que nos criaron a mí y a mis hermanos.

Merci à Sébastien, por quererme durante estos años a través de los cuales hemos crecido juntos (no en altura, ça c'est raté). Gracias por toda su ayuda y apoyo durante la thèse. Todavía me sorprende que una niña de los Andes y un bobito d'un petit village français (encore plus petit qu'ArZacq ArraZiguet) se conocieran y encontraran en el otro a su compañero de camino. Merci à Doña Béatrice, Don Jérôme, Raphaël et Hélène de m'avoir accueillie si gentiment au sein de votre famille. Merci pour les repas, Noël, le jardinage, les balades avec les chiens, et tant d'autres choses.

Gracias a mis hermanos y compañía. A la compañía (Sandra, José y Ale), por su apoyo constante. En particular a José, por sus relecturas (a pesar del miedo), y a mi Alejita, por ser la más ñoña de todas, en la que me veo más reflejada, y a la que espero inspirar para que haga lo que sea que le apasione. A mi hermano, César, por las relecturas también, pero sobre todo por ser mi hermanito, por siempre cuidarme y hacerme la ñoña que soy. A mi hermana, Mimi, por ser mi mejor amiga, y por su personalidad que me complementa.

El mayor agradecimiento de todos es para mis papás. A mi papá, William, por las madrugadas al colegio y a la universidad, por acompañarme para aquí y para allá, por ayudarme con todo cuanto podía, y por acompañarme desde la distancia, pensando y rezando por mí desde que vine a Francia. A mi mamá, Miyerlandi, que es la persona más increíble e incansable que alguien pueda conocer. Ella es el motor de toda nuestra familia y de la vida de muchas personas, en particular de la mía. Gracias por venir a acompañarme, cuidarme, ayudarme y soportarme durante la escritura del manuscrito, y desde siempre. Junto con mi papá hacen el mejor equipo, que nos impulsa a actuar siempre desde la honestidad, a trazarnos grandes metas (aunque parezcan inalcanzables) y a esforzarnos por lograrlas. En especial, gracias por 'voltrear' para que estudiáramos idiomas y expandiéramos nuestro mundo.

Contents

1	Introduction	9
1.1	Dynamic control of a disease spreading on a complex network	9
1.2	Endemic livestock diseases spreading through animal trade	10
1.3	Main objective and structure of the thesis	12
2	State of the art and problem formulation	15
2.1	Epidemiological modelling	15
2.1.1	Compartmental models	16
2.1.1.1	Deterministic formalism	16
2.1.1.1.1	SIR-like models	16
2.1.1.1.2	Models with demography	17
2.1.1.1.3	Structured models by host-heterogeneities	18
2.1.1.1.4	Modelling control measures	19
2.1.1.2	Stochastic formalism	21
2.1.1.2.1	Event-driven approaches	21
2.1.1.2.2	Simulation of stochastic models	22
2.1.1.2.3	Measures of persistence	23
2.1.1.3	Generalisation of classic hypotheses in epidemiological modelling	23
2.1.2	Spatial structure in epidemiological models	24
2.1.3	Network epidemiology	25
2.1.3.1	Static networks	25
2.1.3.1.1	Network connectivity measures	26
2.1.3.1.2	Network models for complex networks	27
2.1.3.2	Dynamic networks	28
2.1.3.3	Infection dynamics on metapopulation networks	29
2.1.3.4	Infection dynamics on animal trade networks	30
2.2	Decision-making with application in epidemiology	30
2.2.1	Types of decision problems	31
2.2.2	Decentralised decision-making: modelling human behaviour	32
2.2.2.1	Phenomenological models	32
2.2.2.2	Psychological models	33

2.2.2.3	Classical game theory	34
2.2.2.4	Evolutionary game theory	35
2.2.2.5	Reinforcement Learning	37
2.2.3	Centralised decision-making: optimising the system	40
2.2.3.1	Control levers and practical considerations	40
2.2.3.2	Allocation problems	41
2.2.3.3	Control approaches	41
2.2.3.3.1	Spectral control	42
2.2.3.3.2	Optimal control theory	42
2.2.3.3.3	Markov decision processes	43
2.2.3.3.4	Score-based approaches	43
2.3	Decision-making regarding the control of disease spread on a trade network: problem formulation and selected approaches	44
3	Accounting for farmers' control decisions in a model of pathogen spread through animal trade	49
3.1	Introduction	50
3.2	Integrative model	51
3.2.1	Epidemic model with demography in a metapopulation based on a trade network	51
3.2.2	Farmers' decision-making model	54
3.2.3	An epidemic control measure	56
3.2.4	An economic-epidemiological cost function	57
3.3	Setting for simulations and sensitivity analyses	58
3.3.1	Fixed simulation setting	58
3.3.2	Parameters of the integrative model	64
3.3.3	Sensitivity analysis experiments	64
3.4	Results of the simulation study and sensitivity analyses	66
3.4.1	Model predictions for different decision scenarios	66
3.4.2	Key determinant parameters of decision-making and epidemiological dynamics	74
3.5	Discussion	76
4	Dynamic resource allocation for controlling pathogen spread on a large metapopulation network	79
4.1	Introduction	80
4.2	Dynamic resource allocation in the metapopulation framework	81
4.3	Score-based strategies	82
4.3.1	Greedy scores	82
4.3.2	Heuristic scores	84
4.4	Simulation setting and numerical explorations	84

4.4.1	Setting for the exploration of infection-related dynamics with score-based resource allocation	84
4.4.2	Setting for percolation analysis	86
4.5	Results	87
4.5.1	Greedy scoring functions	87
4.5.2	Results of numerical explorations	88
4.5.2.1	Infection-related dynamics following score-based resource allocation	88
4.5.2.2	Percolation analysis results	95
4.6	Discussion	102
4.7	Appendix: Analytical derivation of the greedy scoring functions	105
4.7.1	Vaccination	106
4.7.1.1	Minimise a function of the number of infected herds	106
4.7.1.2	Minimise a function of the total number of infected animals	109
4.7.2	Treatment	110
4.7.2.1	Minimise a function of the number of infected herds	110
4.7.2.2	Minimise a function of the total number of infected animals	114
5	Application study: Learning and strategic imitation in modelling farmers' dynamic decisions on BVD vaccination	117
5.1	Introduction	117
5.2	The BVD model	118
5.2.1	Data description	118
5.2.2	Within-herd epidemiological-demographic dynamics	122
5.2.2.1	Life-cycle and health-state dynamics	122
5.2.2.2	Handling trade movements	128
5.2.2.3	Epidemiological effect of vaccination	129
5.3	Farmer's decision-making on vaccination	129
5.3.1	Decision-mechanism	129
5.3.2	Economic-epidemiological cost	131
5.4	Simulation setting	133
5.5	Results	134
5.6	Discussion	140
6	Conclusions and perspectives	143
6.1	Main contributions	143
6.1.1	Integrative model for pathogen spread over an animal trade network accounting for farmers' dynamic decisions regarding the adoption of a health measure	144
6.1.2	Strategies for the dynamic resource allocation of a limited resource for controlling disease spread on a metapopulation network	144

6.1.3	Study on farmers' vaccination decisions integrated in a model of BVD's spread at a large scale	145
6.2	Perspectives	145
6.2.1	On decentralised decision-making	145
6.2.1.1	Noisy observation of costs	145
6.2.1.2	Variations of the decision-making algorithm	146
6.2.1.3	Robustness to parameter values	146
6.2.1.4	More relevant field measures	147
6.2.2	On centralised decision-making	147
6.2.2.1	Noisy or partial health information	147
6.2.2.2	Sub-network	148
6.2.2.3	Dynamic network	148
6.2.3	On decentralised and centralised decision-making	148
6.2.3.1	Coupled centralised-decentralised decision-making	148
6.2.3.2	Decision-step optimisation: when to allocate?	149
6.3	Conclusion	149
7	Résumé en français: contexte et contributions de la thèse	151
7.1	Contexte	151
7.1.1	Gestion dynamique de la propagation d'une maladie sur un réseau complexe	151
7.1.2	Maladies endémiques du bétail se propageant par le biais du commerce des animaux	153
7.2	Contributions	155
7.2.1	Modèle intégratif pour la propagation d'agents pathogènes sur un réseau de commerce d'animaux rendant compte des décisions dynamiques des agriculteurs concernant l'adoption d'une mesure sanitaire	155
7.2.2	Stratégies pour l'allocation dynamique d'une ressource limitée pour la gestion de la propagation des maladies sur un réseau de métapopulation	156
7.2.3	Etude sur les décisions de vaccination des agriculteurs intégrées dans un modèle de propagation de la diarrhée virale bovine à grande échelle	157
7.3	Conclusion	157

Chapter 1

Introduction

1.1 Dynamic control of a disease spreading on a complex network

The structure of contacts between individuals is a key element to account for to better understand the spread of infectious diseases, and ultimately to control it (Keeling et al. 2005). This is particularly true for disease transmission between several sub-populations on a large geographic area. Although pathogens can be transmitted between sub-populations in several ways, depending on the particular pathogen and context (e.g. shared environment, wild-animal vectors), one of the most usual paths for transmission occurs through the movement of infected individuals (Danon et al. 2011).

Indeed, the movements of individuals between different sub-populations form a network structure in which the sub-populations are interconnected by the movement links, called a metapopulation network (Keeling et al. 2005). Such a network is directed, as individuals move from one sub-population to another; weighted, as there is a weight (flow of individuals) associated with each link; and dynamic, as flows change over time (different movement connections and different amounts of individuals). Furthermore, when spatial coordinates are explicitly considered, the links can connect sub-populations even if there is no geographical proximity between them (spatial network).

In particular, this context allows to describe two specific phenomena. First, the introduction of a pathogen in a sub-population where it was not present, known as a colonisation event (Donahue et al. 2008). The second type of phenomenon is the persistence of pathogen propagation at large scale despite fade-out at small scale, which is known as a rescue effect (Brown et al. 1977). This is a well known concept in ecology: there are short-lived epidemics at the sub-population level, but the disease is maintained at a large scale due to interactions among sub-populations.

The question of controlling a disease spreading on such a complex network rapidly emerges (Enright et al. 2018). The system is intrinsically dynamic (due to the movements of individuals, demographic changes in the sub-populations over time, and disease spread), as can be the decision-making. Indeed, real-life control of infectious diseases is usually conducted through repeated decisions over time (vaccine campaigns, temporal movement restrictions, etc.).

Moreover, given the fact that the network is large and encompasses many sub-populations, the decisions regarding the adoption of control measures to reduce the disease spread can be taken at two different levels. On the one hand, if there is a human decision-making agent associated to each sub-population of the system, we are concerned with *decentralised human decision-making*. On the other hand, the management can be carried out by a central planner, usually a social planner. This is referred to as *centralised social-planner decision-making*. The distinction between these two types of decision-making is important since depending on the decision-maker, decisions would respond to different motivations. Furthermore, decentralised human decision-making

can naturally introduce strategic behaviour and heterogeneous implementation of control measures among the many sub-populations of the system (Kreps 1997).

In this context, the use of suitable mathematical approaches is really helpful and strongly recommended for describing and controlling such a system. In particular, mechanistic epidemiological models can be useful instruments to represent and understand the complex system involved in pathogen spread (Keeling et al. 2011). Since human behaviour can have a major role in the spread of diseases, especially for those whose control is voluntary, taking into account decision-making related to controlling the spread can increase the accuracy of these models, in terms of understanding and prediction. Furthermore, on the basis of these mechanistic models, strategies can be designed for controlling the propagation of the disease (Manfredi et al. 2013). Yet, in the process of modelling and controlling such a system, crucial methodological challenges arise when accounting for dynamic decision-making, either made by a social-planner or by the human agents in the system.

Though mathematical mechanistic modelling is increasingly being used in the study of infectious diseases, most existing models in literature do not refer to voluntary decisions of interacting agents (Wang et al. 2016), or they do not consider that the decision-making process is dynamic (Rat-Aspert et al. 2010). Furthermore, classical approaches either consider humans as particles (Manfredi et al. 2013), or when economic aspects are taken into account agents are considered to be perfectly rational (e.g. Bauch et al. 2004; Zhang et al. 2012), which is a strong and controversial hypothesis for human decision-making (Wang et al. 2016). Finally, influences between agents decisions are often neglected (Shi et al. 2019), despite the fact that it is an important feature of human behaviour, particularly regarding the management of livestock diseases (Hidano et al. 2018).

Hence, from a modelling perspective, a first challenge consists in building a framework that appropriately formalises the relation between the dynamics of a disease spreading on a large metapopulation network, and the dynamics of the voluntary adoption of control measures in each sub-population, while accounting for relevant psychological, cognitive or economic considerations.

A second challenge consists in developing or adapting methods that can effectively be applied for optimally controlling the spread of the disease on the network. Among the considerations to be taken into account, one of the most important should be of computational nature. Indeed, this is the major challenge for solving an optimisation problem in the present framework, due to the large dimension of the network (Pellis et al. 2015).

For this reason, research on this matter has mostly been focused on two types of situations: either the network is small (e.g. Chernov et al. 2020; Viet et al. 2018), or it is a large network of individuals, i.e. not a meta-population (e.g. Lorch et al. 2018; Zhang et al. 2015). In the first case, studies are mostly based on Markov Decision processes (MDP) (Puterman 2014) or game-theory (Myerson 1997). In the second one, authors usually recur to mean-field approximations inspired from physics (Lasry et al. 2007), which can be inappropriate if one wishes to account for limited rationality or heterogeneities among agents. For example, this can concern the influence that an agent may have on other agents' decisions, which can be determined by their place in the network.

The expected advances from these two perspectives, modelling and optimisation, can contribute to better understand and take into account different characteristics of the complex system involved in pathogen spread on such a network, and ultimately to effectively control it.

1.2 Endemic livestock diseases spreading through animal trade

In essence, the transmission of a disease does not much differ whether the individuals are humans or animals (wildlife or livestock), yet they present some differences. First, humans tend to move freely between households, work places, study places, cities, countries, etc. while livestock are usually in one farm for a long time before going to another one through animal trade

(Brooks-Pollock et al. 2015). Second, economic considerations can be more central for the control of livestock diseases than for human diseases, where the public health aspect generally takes precedence over the rest. Hence, methodological advances for controlling the first are in principle more likely to find the compliance of field agents.

Following the crisis on bovine spongiform encephalopathy (commonly known as mad cow disease) that took place particularly in the United Kingdom between 1986 and 2000, European countries maintain national databases regarding cattle movements between farms. Indeed, animal movements caused by trading are a major path on which livestock pathogens can be transmitted between holdings (Fèvre et al. 2006), e.g. paratuberculosis (Beaunée et al. 2015), foot-and-mouth disease (FMD) (Ferguson et al. 2001) and bovine tuberculosis (bTB) (Donnelly et al. 2003).

These exchanges occur directly between two farms, or can pass through intermediate structures. The first ones generally concern farms that are geographically close, so the disease spread is mostly concentrated in a small geographic area. Also, animal movements can occur via markets or assembling centres. These structures facilitate animal trade at a large geographical scale, increasing exchanges of animals that come from a holding that is geographically far from the destination holding. Hence, they further increase the transmission risk of pathogens over large areas (Robinson et al. 2007). Figure 1.1 illustrates the example of a cattle trade network formed by the dairy animal movements between cattle herds, that occurred in 2009 in France based on the French cattle identification database (FCID), which records the life history of each cattle animal from birth to death. This figure evidences geographically close and far animal exchanges that can take place in an animal trade network, and through which a disease spread is susceptible to attain the whole French territory.



Figure 1.1: Cattle trade flows (dairy animals only) for year 2009 in France, based on the French cattle identification database. Each node is a commune (the smallest French administrative unit). Source: Gaël Beaunée.

A disease spreading through animal exchange has high chances of becoming *endemic* in a given area, i.e. present at a given (generally moderate) prevalence in the whole population for long periods of time (Carlslake et al. 2011). As mentioned for the general metapopulation setting, in animal metapopulations disease persistence can be observed at large scale by the means of two processes. First, a long infection duration in a sub-population gives rise to long metapopulation infection. Second, persistence can be observed if there is a rescue effect due to interactions among sub-populations (see for example Jesse et al. 2008).

Alternatively, a disease can exhibit *epidemic* dynamics, which by definition means that a large and rapid spread in a short period of time is observed. Examples of this type are FMD (Ferguson et al. 2001), African swine fever (ASF) (Nigsch et al. 2013), and avian influenza (Benincà et al.

2020). For a given disease, the distinction between endemic and epidemic dynamics is made using surveillance data. Yet, detailed surveillance data is mainly available for diseases that have already been labelled as epidemics (Carslake et al. 2011).

Once a disease is associated to an epidemic dynamics, significant political efforts are made in order to eradicate it. In particular, public policies generally target livestock epidemic diseases by regulating their control in a mandatory manner. In that case, it is said that the disease is *regulated*. Meanwhile, endemic diseases are generally of less interest to public opinion and policy-makers. Hence, their control is often left to individual or local initiatives, and is therefore not mandatory, causing endemic diseases to be most often *unregulated*.

Nevertheless, as they persist over long time periods, endemic diseases can have a large cumulative incidence, leading to a reduction in farms economic profitability and in animal welfare (Tomley et al. 2009). Furthermore, zoonotic diseases (i.e. infectious diseases that have jumped from animals to humans) play a major role in the increasing number of human emerging diseases (Lefrançois et al. 2014). Indeed, at least 60% of human emerging infectious diseases are zoonotic, and in particular over 30% of emerging infectious zoonotic diseases are associated with food animals (Otte et al. 2021). Hence, the control of livestock endemic diseases represents a major challenge for animal health and for sustainable agrifood systems, particularly in a context of sub-populations that trade animals, and of endemic diseases for which control is not compulsory.

Decision-making for the adoption of one or several health measures is therefore of interest for assessing the control of animal endemic diseases. As farmers production objectives are not only driven by animal health but also by criteria of working time, productivity, profitability, etc, the heterogeneity of the measures implementation is increased (Ezanno et al. 2020). For example, a central decision-maker may decide to isolate a herd if this implies an overall reduction in disease, provided that this is the objective. However, if the farmers themselves decide whether or not to isolate their herd, they will rather make this decision solely on the basis of their own criteria, which is generally not optimal at the collective level (Krebs et al. 2018).

For unregulated diseases, it is natural to think that decisions concerning herd management are made mostly by the farmers themselves. Yet, decisions on control measures can also be taken at a centralised level, even if only by local organisations composed of several farmers. In some French administrative areas, for example, animal health services (GDS: *groupement de défense sanitaire*) may engage in various types of control actions to reduce the local disease prevalence. These actions may involve, for example, the allocation of health resources for intervening in the system without requiring all farmers to implement the measure, or the use of campaigns to encourage the voluntary adoption of the measure by farmers. Intervention can also be achieved by facilitating access to information on the infection-related status of herds, so that farmers can make more informed, and possibly better, decisions.

Moreover, many control or prevention measures are possible (application of vaccines or treatments, testing and culling of positive animals, isolation of infected animals, etc.). It can be difficult to choose, either from the farmer or the social planner point of view, among the different options whose effectiveness is not always known or comparable. Furthermore, anticipating the impact of these individual or collective choices on the large-scale dynamics of infectious diseases remains a challenge.

1.3 Main objective and structure of the thesis

The main objective of this thesis was to identify, adapt and build the suitable approaches for effectively model and optimise the adoption of control measures for limiting a disease spreading on a large metapopulation network. The network was specifically considered to be a livestock trade network through which the disease can spread over a certain region.

The manuscript is structured in the following manner. This introductory chapter is followed by Chapter 2, that presents the state of the art on epidemiological modelling, particularly on networks, and on decision-making approaches in systems with several individuals. These two

categories of modelling approaches are put together for selecting the suitable elements that are used in the rest of the thesis.

Chapter 3 addresses the modelling challenge, through the development of a new integrative mechanistic model, with a feedback loop between the dynamics of a (theoretical) pathogen spreading on an animal trade network, and the decision-making of farmers regarding the voluntary application of control measures. The decision process is governed by a stochastic algorithm. The work in this chapter was published in *Scientific Reports* in 2021 (Cristancho Fajardo et al. 2021).

For the optimisation challenge, in Chapter 4, the point of view is the one of a single social planner distributing a limited resource aimed at reducing the spread of a (theoretical) disease on the network. At first, the focus is on the construction of optimised indicators to target the nodes of a metapopulation network. Then, the performances of these analytical scores are explored as well as those of relevant heuristic indicators, from an epidemiological and feasibility point of view. This chapter has been published in the *Journal of the Royal Society Interface* (Cristancho-Fajardo et al. 2022).

Chapter 5 presents a new variant of the decision making algorithm as an extension of the model developed in Chapter 3, with application to a specific disease, the bovine viral diarrhoea (BVD). Here, farmers exchange information relative to their decisions either through the trade network, or through a geographic proximity network.

The final chapter (Chapter 6) summarises the developed methods and contributions, discusses their relevance and limitations, and presents some perspectives on the basis of this thesis work, for the control of diseases spreading on metapopulation networks.

Chapter 2

State of the art and problem formulation

Contents

2.1 Epidemiological modelling	15
2.1.1 Compartmental models	16
2.1.2 Spatial structure in epidemiological models	24
2.1.3 Network epidemiology	25
2.2 Decision-making with application in epidemiology	30
2.2.1 Types of decision problems	31
2.2.2 Decentralised decision-making: modelling human behaviour	32
2.2.3 Centralised decision-making: optimising the system	40
2.3 Decision-making regarding the control of disease spread on a trade network: problem formulation and selected approaches	44

This thesis relies on two main topics: epidemiological modelling and decision-making applied to epidemiology. In this chapter, I present a non-exhaustive state of the art on these two topics (sections 2.1 and 2.2, respectively), focusing on the mathematical ingredients that could address (or be involved in) the challenges presented in section 1.3. Finally, some choices regarding such ingredients are made in section 2.3 to be integrated to approaches developed in the following chapters.

2.1 Epidemiological modelling

In epidemiology, mathematical models are used to describe the spread of infectious diseases in a formalised manner. The first known such mathematical model is attributed to Daniel Bernoulli in 1760 (Hethcote 2000). The next registers of mathematical epidemiological models appear in the XX century, through the works of Hamer 1906 for measles, and Ross 1910 on malaria. In particular, Hamer’s work introduces the law of mass action into epidemiology, consisting in the idea that the number of new infections within a population is proportional to the number of infected and susceptible individuals. This law implicitly assumes that the population verifies the *homogeneous mixing* hypothesis, so the contact rate between two individuals is the same for any pair of them. The work of Hamer and Ross, as well as that of Kermack and McKendrick in the 1920s, constitute the basic models for describing the spread of infectious diseases: compartmental models. Then, in the mid-1980s, the work of Klovdahl 1985 and May et al. 1987 paved the way for the development of connections between epidemiology and network theory (Danon et al. 2011).

2.1.1 Compartmental models

Compartmental models consist in dividing the population into homogeneous compartments, according to their health status regarding the disease. Each compartment contains the number (proportion) of the individuals in a given state. Their evolution over time is described by a dynamical system.

From the point of view of the mathematical formalism, models can be deterministic or stochastic, even if their compartmental structure is the same. Although the spread of infectious diseases is stochastic, as it is influenced by many random factors, deterministic models are widely used. In particular, when the stochasticity is only related to the population size, and this one is large, it can be neglected by using a deterministic model. Indeed, for such type of intrinsic stochasticity, (and under certain conditions) the deterministic formulation corresponds to the large population limit of the stochastic formulation (Britton et al. 2019).

In the following, for simplicity reasons, I use the deterministic formalism for describing the main types of structures that compartmental models can present.

2.1.1.1 Deterministic formalism

Most deterministic models are represented by ordinary differential equations (ODE) and can therefore be simulated using numerical schemes such as the Euler method (Euler 1794), the simplest algorithm yet not the best one (Keeling et al. 2011).

2.1.1.1.1 SIR-like models The first simple compartmental model is the SIR model of Kermack et al. 1927. In this model, the population is divided into three compartments: the susceptible (S), the infected (I), and the recovered (R) compartments. In the most basic form of the model, the only two possible transitions are: getting infected (going from the S to the I compartment), or recovering (going from I to R). In its deterministic form, the SIR model can be mathematically written through the following ODE:

$$\begin{aligned}\frac{dS(t)}{dt} &= -\lambda(t)S(t), \\ \frac{dI(t)}{dt} &= \lambda(t)S(t) - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t).\end{aligned}\tag{2.1}$$

$S(t)$, $I(t)$ and $R(t)$, are the numbers of susceptible, infected and recovered individuals at time t . The usual initial conditions of the system described by the previous equations are $S(0) > 0$, $I(0) > 0$ and $R(0) = 0$. The parameter γ is the self-recovery rate, and can be defined as the inverse of the mean duration of the infectious period. This constant recovery rate, independent from the time since infection, implies that the infectious period follows an exponential distribution in the stochastic setting (see section 2.1.1.3 for a discussion on this hypothesis).

The force of infection $\lambda(t)$ is defined as the per capita rate at which a susceptible individual contracts the infection, so $\lambda(t)S(t)$ is the rate at which new infections are produced in the population. The model is represented in figure 2.1.

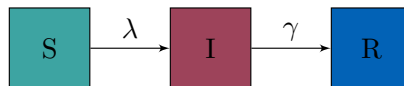


Figure 2.1: SIR flow diagram. $\lambda(t)$ is the force of infection and γ is the recovery rate.

Concerning the form of $\lambda(t)$ two options are mostly used (McCallum et al. 2001). First, one can assume *density-dependence*, where the contact rate between individuals depends on the size of

the population, $N(t)$, and therefore the rate at which new infections take place depends on the density of infected individuals, $I(t)$. So $\lambda(t) = \beta I(t)$, where β accounts for the product between the contact rate of a susceptible and an infected individual, and the probability of transmission following contact. On the contrary, it is possible to suppose that the transmission is *frequency-dependent*, where the contact rate does not depend on the size of the population, so the new infections depend on the proportion of infected individuals rather than on their density. In this case $\lambda(t) = \beta I(t)/N(t)$. The choice of one of the two hypotheses will be important if $N(t)$ shows great variation over time, otherwise $1/N(t)$ can be considered as a multiplicative factor of the β coefficient. Indeed, it is often assumed that the size of the population is constant in time, in which case both functions are proportional, i.e. $N(t) = N = S(t) + I(t) + R(t), \forall t$. In other cases, when population size varies over time, there is no absolute way to know if, for an application, the density-dependence or the frequency-dependence is the more appropriate hypothesis. Yet, frequency-dependence is most often assumed for human diseases, since usually there are similar social patterns between individuals irrespective of N . Meanwhile, density-dependence is more often assumed for animal diseases, as animals might be more crowded in the same geographical area, which would increase their contact rate. However, it has been evidenced that the frequency-dependence assumption can sometimes show more agreement with experimental and observation data than the density-dependence assumption, even for small livestock populations (De Jong et al. 1994). Hence, the frequency-dependence assumption can also be used for modelling infection transmission in animal populations (e.g. Ruget et al. 2021).

An important note is that the contact rate can be a function of time $\beta(t)$. This illustrates seasonal variation in the transmission due to weather influence, for example for vector-transmitted diseases where vector density depends on the weather season, or periodic behaviour, related for example to the school calendar for childhood diseases (Altizer et al. 2006).

From the analysis of the system of equations that describe the SIR model, a fundamental quantity for the study of compartmental epidemiological models like the SIR model can be obtained: the *basic reproduction rate* R_0 , defined as the number of infections one infected individual can generate in a otherwise susceptible population. If $R_0 < 1$ the outbreak will fade out at an early stage, while if $R_0 > 1$ it will be able to invade the population.

On the basis of the SIR model, other models can be constructed by adding or removing compartments and transitions. The SI model for example, can be thought as a particular case of the SIR model, where there is no recovery from the disease, i.e. $\gamma = 0$. Other examples do not consider the life-lasting immunity assumed in the SIR model. The SIS and the SIRS models are some simple examples of this, where individuals can reacquire the susceptible status immediately after being infected, or where immunity lasts for a limited period, respectively.

Another option is to consider additional compartments to account for specific epidemiological characteristics of individuals. For example, the SEIR model considers an Exposed compartment (E), for individuals that have been infected but are not yet infectious, so their contact with susceptible individuals does not lead to new infections, if the contact occurs during the *latency period*, i.e. the duration between being infected and being infectious.

In the following, I precise some specific extensions of the basic SIR model.

2.1.1.1.2 Models with demography Demographic changes can be easily taken into account through a birth and a death terms in the model. Newborns are usually assumed to be susceptible. However, for some diseases, a *vertical transmission* can be possible. That is, mother-to-child transmission during pregnancy or childbirth, opposed to the *horizontal transmission*, where the pathogen is transmitted among individuals of the same generation. Deaths can occur in any of the health-state. Furthermore, an additional disease-related mortality can be taken into account for infected individuals.

The SIR model with demography (births and deaths), without disease-related mortality nor vertical transmission can be represented by equations 2.2, where μ is the birth rate and τ is the natural death rate. Usually, $\mu = \tau$ to ensure constant population size.

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\lambda(t)S(t) + \mu N(t) - \tau S(t), \\
\frac{dI(t)}{dt} &= \lambda(t)S(t) - \gamma I(t) - \tau I(t), \\
\frac{dR(t)}{dt} &= \gamma I(t) - \tau R(t).
\end{aligned} \tag{2.2}$$

The need of inclusion of demographic terms depends on the lifespan of the population, and on the time scale on which processes are studied. Indeed, for human populations studied on a short time scale (for example a year or less) demography is usually neglected. This is not the case for livestock populations, where births and deaths are important to be taken into account for realistically representing such populations and hence infection-related dynamics.

Furthermore, for some diseases, newborns of mother who have antibodies (i.e. mothers that were infected or have immunity) can acquire maternal antibodies for a certain amount of time, protecting them from getting infected (Hethcote 2000). In that case, it is appropriate to consider a compartment for individuals protected by maternal immunity (M), and newborns will either be S or M, depending on the health-state of the mother (S or other). The following equations (2.3) describe an SIR model with demography and maternal protection, which is lost at a rate α :

$$\begin{aligned}
\frac{dM(t)}{dt} &= \mu[N(t) - S(t)] - \alpha M(t), \\
\frac{dS(t)}{dt} &= \alpha M(t) - \lambda(t)S(t) + \mu S(t) - \tau S(t), \\
\frac{dI(t)}{dt} &= \lambda(t)S(t) - \gamma I(t) - \tau I(t), \\
\frac{dR(t)}{dt} &= \gamma I(t) - \tau R(t).
\end{aligned} \tag{2.3}$$

Considering such a maternal immunity compartment is appropriate for diseases such as the BVD (bovine viral diarrhoea) (Zimmer et al. 2004) or measles (Keeling et al. 2011).

2.1.1.1.3 Structured models by host-heterogeneities In the previous subsections only health related compartments have been considered, as it was implicitly assumed that individuals only differed in their health status, and that contacts were random and homogeneous between all individuals in the population (homogeneous mixing hypothesis). However, populations can be further structured by other host- characteristics, such as risk classes or age classes (Keeling et al. 2011). Another type of structure in the population is spatial structure, a situation to which I refer in section 2.1.2.

To limit the homogeneous mixing assumption, one solution is to introduce discrete subcategories within the compartments involving specific transmission rates for their interactions. In particular, transitions between groups can be considered, for example when the structure is given by age (if ageing is compatible with the times scales considered). Figure 2.2 shows the scheme of a SIR model structured in two categories with possible transitions between them.

The model where individuals can only pass from group 1 to group 2 (age-structured model) can

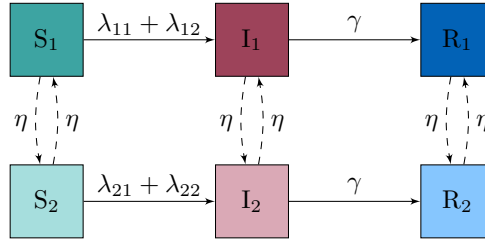


Figure 2.2: SIR model structured in two groups. $\lambda(t)_{kl} = \beta_{kl} \frac{I_l(t)}{N_l(t)}$, where β_{kl} is the infectious contact rate between a susceptible individual of group k and an infected individual of group l , for $k, l = 1, 2$

be formalised through the following ODE:

$$\begin{aligned}
 \frac{dS_1(t)}{dt} &= -(\lambda(t)_{11} + \lambda(t)_{12})S_1(t) - \eta S_1(t), \\
 \frac{dI_1(t)}{dt} &= (\lambda(t)_{11} + \lambda(t)_{12})S_1(t) - \gamma I_1(t) - \eta I_1(t), \\
 \frac{dR_1(t)}{dt} &= \gamma I_1(t) - \eta R_1(t), \\
 \frac{dS_2(t)}{dt} &= -(\lambda(t)_{21} + \lambda(t)_{22})S_2(t) + \eta S_1(t), \\
 \frac{dI_2(t)}{dt} &= (\lambda(t)_{21} + \lambda(t)_{22})S_2(t) - \gamma I_2(t) + \eta I_1(t), \\
 \frac{dR_2(t)}{dt} &= \gamma I_2(t) + \eta R_2(t),
 \end{aligned} \tag{2.4}$$

where η is the rate at which individuals from group 1 pass to group 2. And $\lambda(t)_{kl} = \beta_{kl} \frac{I_l(t)}{N_l(t)}$, where β_{kl} is the infectious contact rate between a susceptible individual of group k and an infected individual of group l , for $k, l = 1, 2$. If groups do not refer to age but to other type of structure (for example, a risk structure for sexually transmitted infections), then transitions can occur in both directions.

Alternatively, host-heterogeneities can be taken into account by assuming a continuum between different categories using partial differential equations (PDE). This seems natural when the heterogeneities are due to age, since it is a continuous variable. Yet, the compartmental approach can be more appropriate if the population is divided in groups in real life. This is the case for childhood infections, since children are often grouped into school classes of a given age range (Keeling et al. 2011).

Regarding the basic reproduction rate R_0 , for deterministic structured epidemiological models it usually corresponds to the spectral radius (i.e. the largest eigenvalue) of the *next-generation matrix*, which depends on the group sizes, and the parameters that quantify the level of interaction between the groups (Diekmann et al. 2010).

2.1.1.1.4 Modelling control measures When considering control strategies, the previous compartmental models can be modified, either by adding new relevant compartments, or by modifying or adding new transitions, hence increasing state and/or parameter space dimension. Among these, one of the most basic models is the SIRV model, represented in figure 2.3.

In this model, susceptible individuals are vaccinated, i.e. pass to the Vaccinated (V) compartment, at a rate ν , which protects them from getting infected. If the vaccine does not provide immunity lasting for the duration of the study, they can lose this protection at a rate ζ , and become susceptible again. Besides the possibility of the effect being limited in time, vaccines can provide an imperfect protection from infection. Let the force of infection of vaccinated individuals be noted as $\lambda(t)_v$; $\lambda(t)_v < \lambda(t)$. In particular, if the vaccine provides a perfect protection, $\lambda(t)_v = 0$.

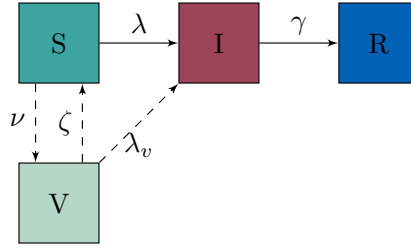


Figure 2.3: SIRV model. ν is the rate at which individuals are vaccinated, and ζ the rate at which vaccinated individuals lose their protection. $\lambda(t)_v$ is the force of infection for vaccinated individuals.

Let $V(t)$ be the number of vaccinated individuals at time t . The SIRV model can therefore be expressed by the following equations:

$$\begin{aligned}
 \frac{dS(t)}{dt} &= -\lambda(t)S(t) - \nu S(t) + \zeta V(t), \\
 \frac{dV(t)}{dt} &= -\lambda(t)_v V(t) + \nu S(t) - \zeta V(t), \\
 \frac{dI(t)}{dt} &= \lambda(t)S(t) + \lambda(t)_v V(t) - \gamma I(t), \\
 \frac{dR(t)}{dt} &= \gamma I(t).
 \end{aligned} \tag{2.5}$$

Other vaccination strategies can be designed, such as vaccinating only a part of the population, e.g. newborns or individuals in risk classes. Furthermore, a similar model can be considered for a treatment administered to infected individuals, by considering an additional compartment T . In particular, the treatment can reduce disease-related mortality (if there is one), or reduce the mean duration of infection for treated individuals. Since infected individuals are supposed to be under treatment, reducing the entire duration of their symptomatic/infectious period, it is usually assumed that treated individuals do not return to the I compartment. This model can be described by the following equations:

$$\begin{aligned}
 \frac{dS(t)}{dt} &= -(\lambda(t) + \lambda(t)')S(t), \\
 \frac{dI(t)}{dt} &= (\lambda(t) + \lambda(t)')S(t) - (\gamma + \nu)I(t), \\
 \frac{dT(t)}{dt} &= \nu I(t) - \gamma' T(t), \\
 \frac{dR(t)}{dt} &= \gamma I(t) + \gamma' T(t),
 \end{aligned} \tag{2.6}$$

where $\lambda(t)' = (\frac{1}{N(t)})\beta' T(t)$, with $\beta' \leq \beta$ the rate at which a susceptible individual contracts the infection from a treated individual. Parameter ν denotes here the rate at which infected individuals are treated and γ' is the rate at which treated individuals recover.

Measures can also be focused on preventing the contact of infected and susceptible individuals. For example, through the isolation of infected individuals, which would require to add a new compartment for these individuals Q (quarantine). The following ODE system describes such model, where ν denotes here the rate of placement in quarantine of infected individuals, and γ' is here the recovery rate of individuals in Q :

$$\begin{aligned}\frac{dS(t)}{dt} &= -\lambda(t)S(t), \\ \frac{dI(t)}{dt} &= \lambda(t)S(t) - (\gamma + \nu)I(t), \\ \frac{dQ(t)}{dt} &= \nu I(t) - \gamma'Q(t), \\ \frac{dR(t)}{dt} &= \gamma I(t) + \gamma'Q(t),\end{aligned}$$

where under a frequency-dependent transmission assumption, $\lambda(t) = \beta \frac{I(t)}{N(t) - Q(t)}$.

2.1.1.2 Stochastic formalism

Following the classification given in Keeling et al. 2011, there are several types of noise that can be considered to affect the infection dynamics.

First, stochasticity can arise from the fact that individuals are different, and transitions between states do not occur in the same way for all of them. This is called *demographic stochasticity*. It is higher for small population sizes, or when the number of infected individuals is small (in which case the probability of an early extinction is positive). This type of stochasticity can be modelled by event-driven approaches, i.e. approaches that explicitly consider (at the unit scale) the different events that can occur in the infection process, or taken into account through the introduction of additional noise terms in the ODE system that describes the deterministic model. A second source of noise can be due to exogenous events that affect the course of the infection dynamics, which is referred to as *environmental stochasticity*, e.g. climatic or individuals' behaviour changes. This is accounted for by considering that the model's parameters are random. Finally, the *observational noise*, i.e. uncertainty in epidemiological data, occurs for instance in case of asymptomatic infectious or under-reporting.

In the following, I focus on event-driven approaches for modelling demographic stochasticity, as they mechanistically represent the randomness that arises from the individual level.

2.1.1.2.1 Event-driven approaches One of the most classical ways to explicitly consider demographic stochasticity is to represent the infection dynamics by a *Markov jump process*. That is, to assume that each compartment of the model is represented by a discrete random variable, and that the stochastic process that determines the evolution of such random variables satisfies the Markov property (Allen 2008). This implicitly assumes that the periods individuals spend in each of the compartments have an exponential distribution (see a discussion on this hypothesis in section 2.1.1.3).

In particular, the Markov jump process analogous to the deterministic SIR model with constant population size (equations 2.1) is the process $X_t = (S_t, I_t)$ with values $x = (s, i)$ where s and i , the number of susceptible and infected individuals, are integer values between 0 and N . Since the number of recovered R_t is deduced from N , the process X_t is bi-dimensional. There are only two possible events at the individual level (infection or recovery), that can increase or decrease the number of individuals in each compartment by one. That is, a hypothesis of this model is that only one transition can occur at each instant. Under the Markov property, the evolution of the system can be described by the following instantaneous transition probabilities:

$$\begin{aligned}p_{infection} &= P[X_{t+dt} = (s-1, i+1) | X_t = (s, i)] = \lambda(t)s \times dt + o(dt), \\ p_{recovery} &= P[X_{t+dt} = (s, i-1) | X_t = (s, i)] = \gamma i \times dt + o(dt).\end{aligned}$$

The probability of no transition taking place within the time interval of length dt , being equal to $1 - p_{infection} - p_{recovery} + o(dt)$.

Bretó et al. 2009 proposed a framework for representing a stochastic compartmental model in terms of flows between compartments, i.e. the number of individuals who have transitioned from one compartment to another. The model is a continuous-time Markov chain (CTMC) that is implicitly-defined via the limit of coupled discrete-time multinomial processes. In particular, this framework allows to also consider environmental stochasticity, through the use of random rates. In the simplest case where the rates are not random, the model corresponds to a Poisson system, i.e. there cannot be simultaneous transitions between compartments Bretó et al. 2009.

An alternative to Markov jump processes that is particularly adapted for modelling the beginning of an epidemic is the *branching process* specifically modelling infectious individuals dynamics. The underlying assumption is that the number of S is sufficiently large to be assumed constant and equal to N (i.e. S depletion is neglected in the first steps of epidemic dynamics). This approximation facilitates the probabilistic study of the epidemic dynamics, in particular regarding the probability of a major outbreak (Britton 2010).

Diffusion processes are another possible approximation valid in the limit of large N . A diffusion process considers a continuous space state, that is the random variables that represent each compartment are real-valued (Andersson et al. 2012). Several studies have been devoted to the study of diffusion processes as approximations of Markov jump processes in the context of epidemiological modelling (e.g. Ethier et al. 2009; Guy et al. 2015).

2.1.1.2.2 Simulation of stochastic models Since the analytical study of stochastic models can prove to be difficult, an alternative is to study their behaviour through the simulations of their trajectories.

One of the most used methods for simulating infectious disease spread, is the *Gillespie algorithm* (Gillespie 1976), which can exactly simulate the Markov jump process that represents infection dynamics. The principle of the method is to simulate the transition of each individual from one state to another, which involves two main steps. The first step consists in simulating the time of the next jump of the process according to an exponential distribution whose parameter is the sum of the transition rates among the different compartments. In the second step, the type of transition is randomly chosen according to probabilities that depend on the state of the process. In a large population, where many transitions can occur at the same time or almost, Gillespie's algorithm can be very costly to use from a computational point of view.

Given this computational cost, several approximations of the direct Gillespie algorithm have been proposed. The most popular one is called *τ -leap algorithm*, of which many variants exist. This approximation, initially proposed in Gillespie 2001, assumes that the jumps take place at a discrete and fixed time step, τ . Therefore, the rates at which transitions occur are constant between two time steps t and $t + \tau$. Yet, in reality such rates can change each time there is an individual transition, so the τ -leap algorithm has a certain approximation error. Hence, the time step τ should be set small enough for the approximation error to be small. Furthermore, in the τ -leaping algorithm, the number of transitions of each type follows independent Poisson laws. Since this law is not bounded, too many reactions can occur during the interval, leading to negative numbers in some compartments. Therefore, alternatives to the original τ -leaping algorithm have been proposed. In particular, Anderson 2008 proposed a technique for postleap checks, which guarantees to never produce negative population values.

The framework developed by Bretó et al. 2009 for representing compartmental models (section 2.1.1.2.1), can be simulated in a straightforward manner through a discrete time-step Euler-scheme, where the number of transitions (between any two compartments) that take place during a small time interval of length $\delta > 0$ are drawn from a multinomial distribution. The use of the multinomial distribution (instead of an unbounded one such as the Poisson distribution) for determining the flows between compartments, allows in particular to always have a positive number of individuals in each compartment.

2.1.1.2.3 Measures of persistence Several indicators exist to quantify the persistence of a disease (Keeling et al. 2011). For certain more or less simple models, some results are available (e.g. time to extinction for the SIR stochastic model). In particular, the classic results regarding the mean time to extinction are due to (Bartlett 1956, 1957). Then, much work has been devoted to approximating the expected time to extinction for slightly more complex epidemiological models (e.g. Andersson et al. 2000; Nåsell 1999). Yet, in general, it can be quite difficult to analytically study this quantity (Andersson et al. 2012).

From a simulation point of view, the time to extinction can be explored either by starting the simulations near the deterministic equilibrium and measuring the mean time when the number of infected individuals extinguishes, or by evaluating the proportion of simulations where this extinction occurs after a certain time. Another relevant measure of persistence that can be calculated is the asymptotic rate of extinction conditional on the disease being present in populations without imports. This can be obtained by simulating populations for many generations, keeping only those where infection is still present, and then further simulating such populations. Finally, it is possible to simulate the population dynamics together with a random infection import, and count the average number of extinctions within a period. Such an approach is biologically realistic yet it is strongly impacted by the pattern of imports, which can be difficult to observe or model (Keeling et al. 2011).

2.1.1.3 Generalisation of classic hypotheses in epidemiological modelling

Although models based on simplified assumptions are mostly used, Kermack et al. 1927 proposed rather generic models considering infection–age dependent infectivity (the infectivity of an individual is dependent on the time since the individual was infected), and infection–age dependent recovery rate (the infectious period follows any continuous distribution). Considering constant rates was only a special case of such generic models that is widely used since it simplifies its analysis (Forien et al. 2021). Indeed, this assumption allows deterministic models to be written as a system of ODEs.

However, assuming constant rates is a strong hypothesis. The underlying assumption is that individuals have the same chance of leaving a compartment no matter the time they have spent in the compartment, which seems biologically not very realistic. This property can be referred to as the *absence of memory* in the deterministic formalism, or the *markovian* property in the stochastic formalism.

The generic models that do not consider absence of memory can be described by integro-differential equations (Kermack et al. 1927) or by PDEs (Kermack et al. 1932). In particular, the latter work additionally considers a recovery-age dependent level of immunity (the susceptibility of a previously infected individual depends on the time since the individual recovered).

Regarding the recovery rate, a simpler alternative to considering generic non-markovian models based on integro-differential equations or PDE is the method of stages (Anderson et al. 1980). Instead of assuming that the infectious period is exponentially distributed with parameter γ , the method considers that the infectious period follows an Erlang distribution (i.e. the distribution of the sum of independent and identically distributed exponential variables). Therefore, it can be modelled by decomposing the infectious compartment into K sub-compartments I_1, \dots, I_K , with the same transition rate for all transitions, such that $\gamma_K = \gamma \times K$ between them, which makes it possible to keep an ODE formulation, while keeping track of the time since infection, and making the chance of recovery dependent on it. This alternative hypothesis has proven to cause a major destabilisation in the system's dynamics, which in the presence of seasonality can lead to complex dynamics patterns with lower levels of seasonality than predicted without this hypothesis (Lloyd 2001).

2.1.2 Spatial structure in epidemiological models

Many models can be used to account for spatial structure in disease spread. Their relevance varies with various factors such as spatial distribution of the population, interaction among hosts, data-availability, and computational constraints (Keeling et al. 2011).

First, *metapopulations models* are useful when the study population is naturally divided in sub-populations, an ecological concept known as metapopulation. Sub-populations have a certain dynamics, and there may be some interaction between them. Regarding infectious diseases, the interaction between sub-populations concerns the mechanisms that allow for the transmission of pathogens. In case of geographical proximity, transmission can occur for example through small particles suspended in the air, i.e. airborne transmission. When the sub-populations are not necessarily geographically close, transmission occurs mostly by movements of individuals.

In such a case, the metapopulation SIR-type model with demography for J sub-populations can be written as:

$$\begin{aligned}\frac{dS_j(t)}{dt} &= -\lambda(t)_j S_j(t) + \mu_j N_j(t) - \tau_j S(t) - \sum_{l=1}^J m_{jl} S_j(t) + \sum_{l=1}^J m_{lj} S_l(t), \\ \frac{dI_j(t)}{dt} &= \lambda(t)_j S_j(t) - \gamma_j I_j(t) - \tau_j I_j(t) - \sum_{l=1}^J m_{jl} I_j(t) + \sum_{l=1}^J m_{lj} I_l(t), \\ \frac{dR_j(t)}{dt} &= \gamma_j I_j(t) - \tau R_j(t) - \sum_{l=1}^J m_{jl} R_j(t) + \sum_{l=1}^J m_{lj} R_l(t),\end{aligned}\tag{2.7}$$

where $S_j(t)$, $I_j(t)$, $R_j(t)$ are respectively the number of susceptible, infected and recovered individuals in sub-population $j = 1, \dots, J$ at time t , and m_{jl} is the rate at which individuals move from sub-population j to sub-population l . Other parameters have the same interpretation as in equations 2.2, but are defined at the sub-population scale.

In a stochastic setting, the extinction of the disease at both levels (the sub-population and the entire metapopulation) emerges from the interaction between the sub-populations and is therefore not easy to predict (Jesse et al. 2011). Indeed, at the sub-population level, an infectious contact with another sub-population can yield a new *invasion* of the pathogen, a *recolonisation*, if the sub-population had already been infected but had recovered, or an increase in the number of infected individuals of the sub-population, if this latter is currently infected. At the metapopulation level, when the pathogen got extinct in some sub-populations, but persisted in others (i.e. *asynchronous* epidemics), the recolonisation of disease-free sub-populations is possible. This asynchronicity further complicates the extinction of the pathogen at the metapopulation level. On the contrary, when epidemics in the sub-populations are *synchronous*, disease extinction is easier to be attained (Hagenaars et al. 2004; Heino et al. 1997).

The spread of a pathogen in a metapopulation occurs at the sub-population level and at the metapopulation level. For an epidemiological model at a given scale, the dynamics at lower scales are usually neglected. In particular, for metapopulation models, if the internal dynamics of each sub-population is ignored, and each one is considered either as infected or not infected, we refer to a *Levins-type metapopulation model*, first introduced by Levins 1969. This model assumes that extinction and recolonisation among the sub-populations are rare events compared to the standard epidemiological dynamics. Yet, if there is a high heterogeneity among the entities at the lower scale it may be necessary to explicitly account for both scales, since predictions will not be accurate otherwise (Keeling 2000). For metapopulation models, this can occur, for example, if the prevalence of infection can be highly different among the infected sub-populations.

If sub-populations can be considered to be uniformly distributed over space, and interactions mostly occur with geographically close sub-populations, one may use a *lattice model*, which is a type of metapopulation model that places the sub-populations in sites of a grid. A particular

lattice-based model is the *cellular-automata*, for which there is a small number of sites in the grid, and each site generally represents only one individual. Lattice-based models have been useful for understanding simple effect of the spatial heterogeneity (e.g. Kao 2003; Keeling et al. 2000), and they generally allow easily to account for stochasticity. Yet, they are rarely relevant for describing the behaviour of complex real-life problems (Keeling et al. 2011).

If the population is densely geographically distributed, an option is to consider continuous and deterministic representations, and use *continuous-space continuous-population models* (Murray 2003). Among these, we find reaction-diffusion models, which use PDEs for describing disease spatial transmission. Other options are integro-differential and PDE models, which consider transmission kernels specifying the form in which the risk of transmission decreases with spatial distance.

Alternatively, it is possible to consider *individual-based models* (IBM), that is, models where the state of each individual is explicitly tracked. This allows to know not only how many individuals are in a certain compartment, for example infected, but also which individuals are in the compartment. Furthermore, IBM allow to consider complex behaviours of the hosts. In particular, they can consider a spatial heterogeneity that is not restricted to a grid. Yet, such level of detail in IBM implies a high number of parameters that can be difficult to estimate or calibrate. Also, they are more computationally intensive than compartmental models (DeAngelis 2018), even if there are ways to accelerate simulations (Keeling et al. 2011).

A final approach for taking into account spatial structure are *networks*, although these are not limited to accounting for spatial heterogeneity. Indeed, they provide an unified representation of the interaction between individuals or populations (Newman 2010), and are particularly useful when the entities of the network mostly have few contacts with others. In particular, *metapopulation networks* can represent a metapopulation in a flexible manner, since they provide an intuitive graphical representation of it when spatial coordinates of sub-populations are considered, while directly accounting for the (time-varying) movements of individuals between sub-populations, or for other transmission ways due to geographical proximity.

2.1.3 Network epidemiology

Although simple compartmental models are useful and relevant when the homogeneous mixing hypothesis is plausible, they neglect specific and complex contact structures in the population, which can be important for correctly reproducing the infection-related dynamics in such a population (Keeling et al. 2011). In this context, network theory is a relevant field to accurately represent and study the contacts between individuals or populations through which pathogens spread. Indeed, networks are made up of a set of nodes, and a set of links between some of those nodes, which allows to account for the heterogeneity of contacts, and can therefore generate a better understanding and control of infectious diseases spread (Keeling et al. 2005).

In the following, I focus on the study of networks and pathogen spreading on networks that are formed by several interacting populations.

2.1.3.1 Static networks

A *network* (or graph) can be noted as $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \dots, J\}$ is defined as the collection of network's *nodes* (or graph's vertices), and $\mathcal{E} = \{(i, j) : i, j \in \mathcal{V}\}$ is the set of *links* between the nodes (or edges between the vertices). Hence, the link $(i, j) \in \mathcal{E}$, if nodes i and j are connected in the network.

A network can either be represented by a list of its links, i.e. of the elements in \mathcal{E} , or by its *adjacency matrix* $A = (a_{ij})$, of size $J \times J$, where $a_{ij} \neq 0$ if there is a link between i to j , i.e. if $(i, j) \in \mathcal{E}$, also noted $i - j$, and $a_{ij} = 0$ otherwise. If $i - j$, we say that i and j are adjacent.

Having $a_{ij} \neq 0 \Leftrightarrow a_{ji} \neq 0$, means the network is *undirected* (figure 2.4a). Otherwise, the network is called *directed* (figure 2.4b), and is noted as $\vec{G} = (\mathcal{V}, \vec{\mathcal{E}})$, where $\vec{\mathcal{E}}$ is the set of directed links.

As the adjacency matrix A is not symmetric, $a_{ij} \neq 0$ if there is a link from i to j , also noted as $i \rightarrow j$, and $a_{ji} \neq 0$ if there is a link from j to i , i.e. $j \rightarrow i$.

Finally, if there is a particular weight associated with each link, the network is said to be *weighted* (figure 2.4c). We can note it as $G_w = (\mathcal{V}, \mathcal{E}_w)$, where $\mathcal{E}_w = \{(i, j, w_{ij}) : i, j \in \mathcal{V}, w_{ij} \neq 0\}$. That is, $(i, j, w_{ij}) \in \mathcal{E}_w$. The adjacency matrix of a weighted network is $W = (w_{ij})$, since we assume $w_{ij} = 0$ if there is no link between two nodes i and j .

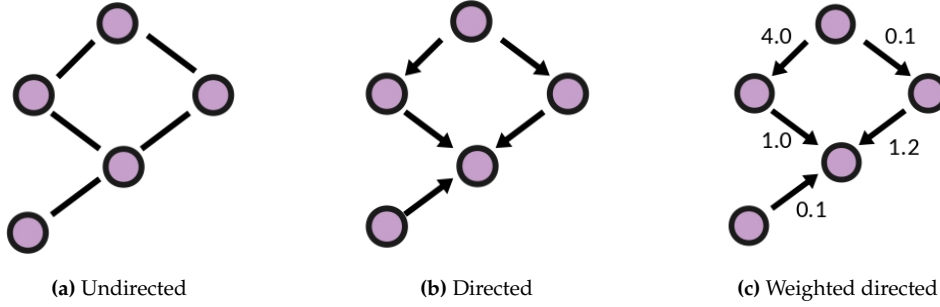


Figure 2.4: Representations of a undirected network (a), a directed network (b), and a weighted directed network. Arrows represent directed links.

We call the *neighbours* of a node j in the network G , $nbr(j, G) := \{i : (i, j) \text{ or } (j, i) \in \mathcal{E}\}$, i.e. the set of nodes connected to j in the network. In a directed network \vec{G} , we call the *parents* of the node j , $Par(j, \vec{G}) := \{i : (i, j) \in \vec{\mathcal{E}}\}$. That is, the set of nodes that directly point to j in \vec{G} . In the same way, we can define the *children* of the node j , $Ch(j, \vec{G}) := \{i : (j, i) \in \vec{\mathcal{E}}\}$, i.e. the set of nodes to which j is directed in \vec{G} . In that case, $nbr(j, \vec{G}) = \{i \in Par(j, \vec{G}) \cup Ch(j, \vec{G})\}$.

We call a *path* the finite or infinite sequence of links which joins a sequence of nodes in which all nodes (and therefore also all links) are distinct. For directed networks, links must be in the same direction. A *shortest path* between two nodes in a network is a path with the minimum number of links, or for a weighted network, a path for which the sum of the links' weights is minimised. Finding a shortest path is a classical problem in network theory, and many algorithms exist for solving it, e.g. Dijkstra's algorithm (Dijkstra 1959). The *distance* between two nodes in a network is the length of a shortest path between them, if there is one, and otherwise the distance is infinite.

A directed network is said to be *strongly connected* if every node is *reachable* from every other node, i.e. if there is a path in each direction between each pair of nodes of the network. That is, for any pair of nodes i, j of the network, a path exists from i to j , and another path exists from j to i . The *strongly connected components* of a directed network form a partition into subgraphs that are themselves strongly connected. A *giant component* of a directed network is a strongly connected component that contains a finite fraction of all the nodes of the network.

2.1.3.1.1 Network connectivity measures The connectivity of a network can be described by several measures (Newman 2010). In the following, I focus on measures of local connectivity, and I describe a particular measure that concerns global connectivity.

Topological centrality measures describe local connectivity, i.e. how important or 'central' is a node given the topological characteristics of the network. The most basic centrality measure is the *degree*, which is defined as the number of links connected to the node, i.e. the number of neighbours of the node. Then, the degree of a node j is the size of the set $nbr(j, G)$. In terms of the adjacency matrix A , for an undirected network the degree of a node j is $deg(j) = \sum_{i \in \mathcal{V}} a_{ij}$.

For a directed network, we can distinguish between incoming degree, *in-degree*, and outgoing degree, *out-degree*. The in-degree of a node j corresponds to the number of links pointing towards j , i.e. the number of parents of j , which is $\overleftarrow{deg}(j) = \sum_{i \in \mathcal{V}} a_{ij}$. The out-degree of j is the number of links coming from it, i.e. the number of children of j , which is $\overrightarrow{deg}(j) = \sum_{j \in \mathcal{V}} a_{ji}$. The *total degree* of the node is therefore the sum of the in and out-degree.

For weighted networks, we talk instead about the *strength* of a node or the weighted degree, i.e. the sum of the weights of its links $\sum_{i \in \mathcal{V}} w_{ij}$. For a directed weighted network, the *in-strength* of node j is therefore the sum of the weights of all links going to j , and *out-strength* is the sum of the weights of all links coming out of j . Like for the total degree, the *total strength* of a node, will be the sum of its in and out strength.

Among other more elaborate centrality measures we find the *closeness*. It is defined, for a node j , as: $closeness(j) = \frac{1}{\sum_{i \in \mathcal{V}} d(j,i)}$, where $d(j,i)$ is the topological distance between j and another node i . According to this measure, a node is central in the network, if it is ‘close’ to all other nodes.

Another measure that is also based on shortest paths is the *betweenness* centrality. For a node j , it is defined as the ratio of the number of shortest paths passing through j over the number of all such paths in the network: $betweenness(j) = \sum_{i \neq j \neq k} \frac{\sigma_{ik}(j)}{\sigma_{ik}}$, where σ_{ik} is the total (weighted) number of shortest paths between node i and k , and $\sigma_{ik}(j)$ is the number of those paths that pass through j .

The *eigenvector* is a centrality measure based on the eigenvectors of the adjacency matrix. The eigenvector score of node j is the j -th element of the eigenvector corresponding to the largest (in absolute value) eigenvalue ξ of the adjacency matrix. Therefore, the eigenvector score of a node j has recursive dependence on the eigenvector score of other nodes that point to it. Since for directed networks, the adjacency matrix is asymmetric, it has two different sets of left- and right-eigenvectors. One can either use the left eigenvectors, to find nodes that have a large number of incoming paths, or the right eigenvector, if the interest is in nodes having many outgoing paths. The (left) eigenvector score for a node j is: $eigenvector(j) = \frac{1}{\xi} \sum_i a_{ij} \times eigenvector(i)$. According to this measure, a node j is central in the network if it is connected to many nodes that are in turn very connected. It is worth noticing that eigenvector centrality can have issues with directed networks. Indeed, real-life directed networks are in general not strongly connected. Due to this, almost only nodes in strongly connected components will have non-zero eigenvector centrality. Meanwhile, all the other nodes will have a null centrality. This happens because nodes with no incoming links have, by definition, a null eigenvector centrality. Therefore, nodes that are pointed only by nodes with a null centrality, will also have a null centrality.

The *PageRank* is a centrality measure conceived for directed networks, based on the eigenvector centrality and random jumps, that among other features corrects the previous issue. It is defined as $PageRank(j) = a \sum_i \frac{a_{ij}}{d_i} PageRank(i) + b$, where a and b are constants, d_i equals $\overrightarrow{deg}(i)$ if $(i,j) \in \vec{\mathcal{E}}$, 1 otherwise.

Finally, an indicator of global connectivity for directed networks, particularly useful in epidemiology, is the *giant strongly connected component* (GSCC), the sub-network of maximal size in which there is a path between each pair of nodes of the sub-graph. Therefore, the GSCC informs about the maximal proportion of nodes that can be reached from any node of this sub-network (Dorogovtsev et al. 2001).

2.1.3.1.2 Network models for complex networks There exist many models that search to explain the probabilistic mechanisms underlying the generation of network structures. They are usually referred to as *network models*, or they are sometimes called *random networks*. The simplest example of an undirected random network model is the *Erdős-Rényi model* (Erdős et al. 2011), initially introduced in 1960. This network can be generated for a certain number n of nodes, by independently connecting each pair of nodes with a probability p . The resulting degrees follow a *Binomial*(n,p) distribution (see an example in figure 2.5a). For its simplicity, much attention has been given to its analysis (Newman 2010).

Yet, real-systems networks often have non-trivial topological features that do not occur in simple networks like those generated by the Erdős-Rényi model, we refer to them as *complex networks* (Albert et al. 2002). One of the most common features of complex networks is the *scale-free property*. We say a network has the scale-free property, or is scale-free, if its degree distribution follows

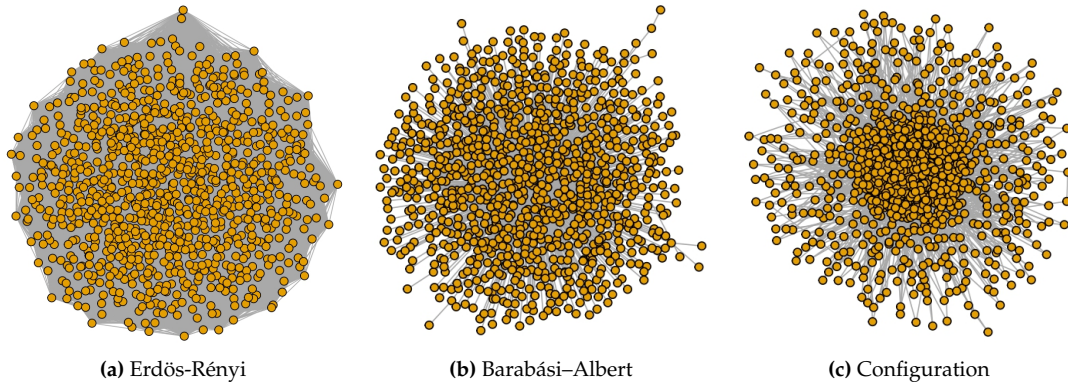


Figure 2.5: Representations of networks of $n = 1000$ nodes generated according to three models. (a) Erdős-Rényi model with $p = 0.5$ (the probability of connecting any pair of nodes). (b) Barabási-Albert model with $v = 2$ (the scale parameter). (c) Configuration model where half of the nodes have a degree 2, and half have degree 100.

a power law, i.e. the proportion of nodes having x neighbours in the network, goes for large values of x as $p(x) \sim x^{-v}$, where v is called the scale parameter. This means that in the network there are nodes, called *hubs*, with much higher degree than the average. A network model that might explain the appearance of power-law distributions in real-world networks is the *preferential attachment* model, a term initially used for the Barabási-Albert model in Barabási et al. 1999. This model lies on the idea that a new node prefers to connect with an existing node in the network that has already established some links with others nodes, which would eventually lead to the appearance of hubs (figure 2.5b).

Contrary to the Erdős-Rényi model, or the Barabási-Albert model, the *configuration model* generates a random network from any given degree sequence (see figure 2.5c). Indeed, instead of having a distribution from which the given degree is chosen, the degree of each node is predefined. It is therefore not restricted to a Binomial degree distribution, or to a power-law distribution. For its flexibility, it is widely used as a model for representing real-life social networks.

The study of scale-free networks are of special interest in epidemiology. One of the reasons is that the most connected individuals or populations, i.e. the hubs, can play a major role in the spread of a disease at a large scale (Keeling et al. 2011).

2.1.3.2 Dynamic networks

Until here, I have implicitly assumed that the network is *static*, since its structure (nodes and links) does not evolve over time. However, several real-life contact patterns have a temporal dimension, as nodes and/or links can appear or disappear over time. The network is therefore referred to as *dynamic* or *temporal* (Masuda et al. 2016).

A dynamic network takes into account the chronological order at which links are established. While the definition of dynamic networks is an extension of static networks, the definition of their properties is not as straightforward. A path, for example, should respect the chronological appearance of links, which complicates its conceptual definition, and in particular makes the shortest path problem more complex. The previously described centrality measures do not take into account the temporal dimension. Several alternative centrality measures have been defined, mostly by generalising static centrality measures. In particular, Kim et al. 2012 consider a temporal betweenness, and Taylor et al. 2017 describe a temporal coupling of eigenvector-based centralities. However, these methods suffer from a high computational cost, due to the additional complexity of temporal networks. In this context, Rocha et al. 2014 proposed a measure based on time-respecting random walks, as an extension to the PageRank centrality. Hoscheit et al. 2021 showed this measure can be efficiently computed using stochastic simulations, so it can be well suited for large temporal networks.

Finally, in dynamic networks, it is possible to consider discrete or continuous time. Although discrete time can theoretically be seen as an approximation, it can be well adapted for certain contexts, particularly from a practical point of view. Indeed, real-life data-sets that can be represented as temporal networks are recorded in discrete-time. For example, records on animal trade movements are usually done on a daily basis, so it is natural to consider daily static networks. Furthermore, data can be aggregated so as to consider snapshots of the dynamic network, if it is assumed or verified that the structural characteristics of the aggregated snapshots are similar (Holme et al. 2012).

2.1.3.3 Infection dynamics on metapopulation networks

The structure of the underlying network (linking individuals of a population, or sub-populations in a metapopulation) on which a pathogen can spread can strongly affect infection-related dynamics (Keeling et al. 2005). In the following, the focus is on infectious diseases spreading on metapopulation networks.

Many studies have focused on pathogens spreading on networks, both from a theoretical and applied point of view (e.g. Pastor-Satorras et al. 2015, 2001), particularly in metapopulation networks (e.g. Colizza et al. 2008b; Hagenaars et al. 2004; Lloyd et al. 2004; Vergu et al. 2010). In particular, mathematical models have been used to explore the spread (over time and space) of a pathogen, by accounting for the individual movements as a source of infection for populations where a pathogen is not present (e.g. Colizza et al. 2006). As mentioned before, infectious contacts can either be undirected, for example in social networks of individuals, or directed, for example in the previously mentioned mobility networks. Additionally, contacts can be assumed homogeneous, or they can depend on node/link specific factors, i.e. weights can be associated to the nodes/links of the network, which ultimately implies a different infection risk.

Furthermore, in a meta-population, it is possible to ignore or not the dynamics internal to each sub-population. The first option is rarely realistic, since it neglects heterogeneities in the severity of infection of an infected sub-population over time, or among infected sub-populations. Also, it supposes that the infection spreads rapidly in a sub-population, compared to the setting up of contacts between sub-populations. On the contrary, considering the within-node infection dynamics of each sub-population can be more accurately represented when processes occur in similar time-scales. Yet, this increases the complexity of the model, as it requires a coupling between the sub-populations internal dynamics, and the dynamics among them, given by interactions. This increased complexity often prohibits the analytical study of model behaviour, and can make model simulation studies very computationally intensive.

Additionally, if the structure of the network varies over time, the order of events occurrence must be taken into account, since it will affect the propagation among nodes. Indeed, it has been shown that not taking into account the temporality of the network can result in major differences in the spread (see figure 2.6 for an example). In particular, static networks fail to capture the predicted infection dynamics behaviour associated with dynamic networks, as they do not account for temporal structures that prohibit or allow infectious contacts (Fefferman et al. 2007; Vernon et al. 2009). The temporal structure of contacts can be taken into account as being exogenous to the infection dynamics, by plugging in network data (e.g. Beaunée et al. 2015). However, the predictions concerning the disease spread will necessarily be associated to past observations of this underlying network. Indeed, there is no guarantee for predictions to hold for another time period, since the links of the network may have probably changed over time. Another option is to consider the network as being endogenous to the pathogen spread, by defining the setting up of links as a transition event jointly occurring with infection dynamics. This approach has the potential of being well-adapted if one is searching for more potentially generic results, although it may be harder to implement.

A final important issue that must be considered when studying infectious diseases on networks is computational complexity (Newman 2010). Indeed, current research is looking at larger and more complex networks. This has been made possible by advances in the collection and analysis

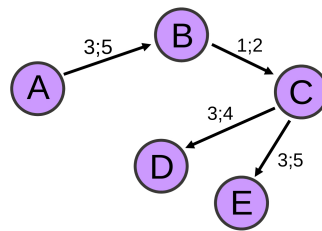


Figure 2.6: Representation of a simple dynamic network. The coefficients on the arrows indicate the time steps at which the link is active. If node A was infected at time 0, the infection could only reach node B, yet according to the static network (i.e. if the temporality of the links is not accounted for) the infection would spread to the whole network.

of these data. Even if many standard algorithms for the study and simulation of networks are already available, computational complexity needs to be considered in order to understand how to use these pre-made tools and not be limited by what a certain package can do. Computation must be taken into account even in the way network data is stored, e.g. as an adjacency list, as an adjacency matrix, etc. Indeed, these aspects can also have an impact on the simulation time and the amount of used memory. This issue must be of particular attention in temporal networks, as links occurrence varies over time, so the memory space and speed of simulations can be further increased.

2.1.3.4 Infection dynamics on animal trade networks

Animal movements databases can all intuitively be represented as directed weighted (dynamic) networks, where holdings are represented by the nodes of each network and animal flows are represented by the weighted directed (dynamic) links between the nodes.

Many studies have been devoted to the statistical description of animal movements databases, such as Dutta et al. 2014 in France, Robinson et al. 2007; Vernon 2011 in the United Kingdom, Martinez-Lopez et al. 2009 in Spain, Matos Baptista et al. 2007 in Portugal, or Natale et al. 2009 in Italy. Furthermore, numerous studies used this representation along with network-analysis techniques (e.g. characterising degree distributions, analysing the giant component or the weakly and strongly components of the network) to study animal trade data, such as Bajardi et al. 2011; Mweu et al. 2013; Natale et al. 2009; Vernon et al. 2009. In particular, it has been observed that animal trade networks have the scale-free property (Rautureau et al. 2011).

Additionally, it has been possible to study the potential spread of infectious diseases at a large scale through these movements. In particular, evaluating the impact of this (time-varying) structure regarding pathogen spread and control (e.g. Bajardi et al. 2012; Beaunée et al. 2017; Bigras-Poulin et al. 2006; Dubé et al. 2008; Ezanno et al. 2006; Gilbert et al. 2005; Kao et al. 2007; Nöremark et al. 2011; Volkova et al. 2010). For example, Bajardi et al. 2011; Vernon et al. 2009 show that the temporal variability of the network is an important factor determining its vulnerability to disease propagation.

Finally, modelling the spread of pathogens between farms represented as a metapopulation network not only allows for contact or spatial heterogeneity, but also naturally considers heterogeneity in the application of control measures across farms in the network (Ezanno et al. 2020) (e.g. Rautureau et al. 2011).

2.2 Decision-making with application in epidemiology

The study of decision problems is a large and interdisciplinary topic that has been treated from different perspectives by mathematicians, economists, psychologists, game-theorists, biologists,

social scientists, computer scientists, etc.

The decision problem can concern a system with only one entity or agent, i.e. decisions solely concern this agent. Yet, in the following I focus on cases where decisions are made for or by many agents. I attempt to give a short and not comprehensive review of such problems, and of the mathematical approaches available to represent and eventually solve them. I emphasise that for both problem types and mathematical approaches, the state of the art is rendered difficult since elements can be found in different fields.

2.2.1 Types of decision problems

Decision problems can be classified in several manners depending on the used criterion. In the following, the first two criteria are relevant both for systems with a single agent or many agents, while the third one is specific to the latter case.

First, a distinction can be made according to when decisions are taken: just once, or through repeated actions. For problems with repeated actions I refer to a *sequential decision problem* (Diederich 2001). Examples of sequential decision problems are optimal stopping problems (Chow et al. 1971), in which the agent makes a sequence of decisions based on a sequence of observations to find a time to perform a particular action, in order to maximise an expected reward. The most studied is the *secretary problem*, in which an administrator has to hire a secretary among n applicants. Yet, they are interviewed one by one and the administrator must decide immediately after the interview whether to hire or reject the person, and cannot call back a previously rejected applicant. The objective is then to find the time at which the best applicant is the hired person, finding a strategy that maximises the probability of selecting the best person for the job (Ferguson 1989). Other well known sequential decision problems are *inventory problems* and *multi-armed bandit (MAB) problems*. The first one is the problem faced by a firm that must regularly decide how much of a product to stock in order to meet future demand for its products (Dvoretzky et al. 1952). The latter takes its name from a situation where a gambler faced with several slot machines (or one-armed bandits) has to decide which machines to play, how many times, in which order, and whether to change the machine. It concerns therefore the decision-maker that must choose at each time among several possibilities in order to maximise the overall expected gain, where there is uncertainty in the result of each choice in advance. The player collects information over time, i.e. learns, in order to better decide in the following steps (Berry et al. 1985). See section 2.2.2.5 for more details on this problem.

Though the terms sequential and dynamic are sometimes indistinguishably used, we talk about a *dynamic decision problem* only if the sequence of decisions are made in order to control and optimise the performance of a dynamic system (Busemeyer 2001). The dynamic aspect of the system is given by a state change over time, both due to the system itself and to the previous actions of the decision maker (Hotelling et al. 2015).

Second, a decision-problem can be classified depending on the degree of certainty that the decision-maker has on the result of his/her actions. Generally, a deterministic setting will result in a decision problem without uncertainty, i.e. the decision-maker knows in advance the consequences of each possible decision. On the contrary, when decisions are made on the basis of outcomes that are unknown in advance, there is *decision-making under uncertainty* (Kochenderfer 2015). When the probability distributions of such outcomes are known by the decision-maker, the decision-making is said to be under risk, and the term uncertainty only refers to situations where these distributions are unknown (Etner et al. 2009). When the uncertainty in the probability distributions is extreme, some authors say that the decision is made under ignorance (e.g. Cohen et al. 1985).

Finally, when the system is composed by more than one agent, one of the main distinctions regards the decision-maker, i.e. the agent who makes the decisions, with two main options. If there is a central decision-maker for all the agents in the system, we talk about *centralised decision-making*, while if each entity or agent of the system makes its own decisions, we talk about *decentralised decision-making* (Xuan et al. 2002).

In general, irrespective of the centralised or decentralised nature of the decision-making problem, its formulation can have as objective either to represent how the decision-maker actually decides i.e. modelling the decision-maker's behaviour, or to find the way in which the decision-maker should decide, i.e. optimising the decision-maker's decisions. However, in epidemiology, the setting of the problem, centralised or decentralised, strongly shapes its conceptual and mathematical formulation, and can therefore define the available approaches to solve it. In particular, since the agents of the system are humans, the overall objective is to model real-life human behaviour. On the contrary, the general objective of a centralised decision problem in epidemiology is to find the optimal decisions that should be taken by a social planner for each agent in order to optimise a certain global criterion of the system, i.e. optimising the system.

Hence, in what follows, the framework is that of a system of many interacting human agents, for which I consider the two main cases given by the previous distinction. That is, on the one hand I study the decentralised case, where I focus on modelling human decision-making. On the other hand, I consider centralised decision-making, where the objective is to propose techniques for the system optimisation.

2.2.2 Decentralised decision-making: modelling human behaviour

In the following I detail the main approaches for representing the decision process of each agent of a system, where each agent is a human. That is, approaches for representing human behaviour, with a special focus on voluntary behaviour when agents are faced to the spread of an infectious disease. However, some of these approaches lie on the assumption that humans are able to optimise their decisions perfectly. So, human behaviour is essentially modelled as the solution to an optimisation problem, or as if the decision-maker was a machine.

The classification is somewhat different to the one found in Wang et al. 2016, yet it is based on it. Indeed, I consider evolutionary game-theory as a category on its own, i.e. not embedded within psychological models, as the former uses elements both from game-theory and psychology. Furthermore, I consider an additional category on reinforcement learning. The five categories are therefore: phenomenological models, psychological models, game-theoretical models, evolutionary game-theoretical models, and reinforcement learning. It should be noted that Markov decision processes are treated both in game-theoretical models and reinforcement learning.

2.2.2.1 Phenomenological models

Phenomenological behaviour models describe the effects of decisions, without posing a mechanism for the decision process.

In this class, we first find models that use mean-field ODEs with a parameter that accounts for the impact of individual behaviour on disease spread. The first model to take this approach was the generalised SIR model of Capasso et al. 1978:

$$\begin{aligned}\frac{dS(t)}{dt} &= -g(I(t))S(t), \\ \frac{dI(t)}{dt} &= g(I(t))S(t) - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t).\end{aligned}$$

The functional form $g(I(t))$ of the force of infection can capture the effects of human behaviour, such as individuals reducing contacts at high levels of prevalence.

The other type of phenomenological models are the ones that lie on the assumption of spread of ideas by a diffusion process, or by an unspecified hypothetical awareness-raising event (e.g. Mendes et al. 2020). This type of model can sometimes neglect important features of real social behaviour (Alshamsi et al. 2015), such as imitation, or more generally the influence of observing

the behaviour of other agents (Bandura et al. 1977). However, they can be useful to easily observe the effect of individuals' behaviour on disease spread, and to describe through simple ideas how complex behaviours can emerge in populations (Wang et al. 2016).

A final example of a model that phenomenologically accounts for behaviour is one that divides population into groups according to a static vaccination profile (such as "vaccinates" vs "does not vaccinate"), and only explores infection dynamics given the interaction of agents in the different assigned groups (e.g. Hill et al. 2021). Indeed, such a model does not consider how the agents can acquire (nor lose) such vaccination profiles.

2.2.2.2 Psychological models

Among the models that do formulate mechanisms to describe individuals' behaviour we find the ones that lie on specific psychological theories.

A first example are models that rely on *subjective expected utility* (SEU) theory (Savage 1972), which is an extension of *expected utility* (EU) theory. EU theory supposes that individual decisions arise from a utility function $U(x_i)$ of the possible outcome $\{x_i\}$ of a given decision i , and the probability of each outcome $p(x_i)$ (Machina 1987). Then, if the individual has another option for the decision with possible outcomes $\{y_i\}$, it chooses the decision with the highest EU. SEU theory adds a psychological aspect to EU theory, supposing that the agent only has a personal belief of the probability of each outcome, and decides based on the subjective expected value of the utility. That is, for computing the SEU of each option, the agent uses their subjective belief on the probabilities, and then chooses the decision with the highest SEU. It is to be stressed that many economic experiments have pointed out that the theory is not consistent with real human decision-making (Karni 2008). In particular, Ellsberg 1961 used simple hypothetical experiments to show how sometimes it can be impossible to infer probabilities from the choices made by an individual, concluding that individuals do not always decide as a function of a probability distribution on the outcomes.

Another important approach is *prospect theory* (Kahneman et al. 1979), which belongs to behavioural economics, a relatively recent economic field that (contrary to classical economics) accounts for psychological, cognitive, emotional, cultural and social factors on individuals decisions (Teitelbaum et al. 2018). Unlike EU theory, prospect theory does not consider perfectly rational agents, i.e. agents that decide by performing a cost-benefit analysis to determine the option that maximises their benefit at the smallest cost. In particular, it describes individuals preferring to avoid losses rather than to acquire equivalent gains. Also, it accounts for a cognitive human bias consisting in over-estimating the probability of very rare events. It is therefore relevant in epidemiology for accounting for the influence of rare events, such as vaccine adverse events or disease complications, on human decision-making (Wang et al. 2016). It has been more often used in mean-field models coupling vaccinating behaviour and disease spread. In particular, Oraby et al. 2015 built a disease-behaviour model in the context of vaccination against a paediatric infectious disease, accounting for bounded rationality using prospect theory. In this work, a comparison with a simpler model where individuals behave rationally is made, showing that taking into account cognitive considerations in the first model could profoundly impact predictions.

A third approach belonging to psychological models are *belief-decision models*. These models lie on the Dempster-Shafer theory (Shafer 2020), a general statistical framework for reasoning with uncertainty. This framework allows for an agent to combine evidence from different sources to update the belief on the probability of an event, belief that determines his subsequent decisions. In epidemiology, belief-decision models have been used for coupling vaccination and disease dynamic decisions, particularly to account for the influence of awareness concerning disease prevalence and vaccine safety (see Xia et al. 2014).

2.2.2.3 Classical game theory

Game theory studies mathematical models of strategic interactions among decision-makers (Myerson 1997). In classical game theory, individuals are generally assumed to be perfectly rational, i.e. able to perform a cost-benefit analysis to determine whether an option is good for them and to choose that option.

To be fully defined, a game must specify its players, the information and actions available to each player at each decision point, and the payoffs for each outcome. Furthermore, the game can be either cooperative, if players are supposed to be able to form alliances, or non-cooperative otherwise. Cooperative game theory focuses then on predicting coalitions, joint actions and collective payoffs. Meanwhile, non-cooperative game theory focuses on predicting individual players' actions and payoffs. In particular, it analyses Nash equilibria, a set of strategies for each player such that, when employed, no player can profit by unilaterally changing strategy.

The most studied non-cooperative game is the Prisoner's Dilemma, formalised by Albert W. Tucker in by 1950 (Poundstone 1992). This game consists in two members of an organisation being imprisoned separately without possibility for communicating with each other. Each prisoner *A* and *B* can either betray by saying that the other committed the crime. It is implicitly assumed that prisoners cannot punish the other's betrayal, i.e. the only thing that matters in decision is the years of conviction they can get according to their decision. If both prisoners betray each other, each will have 2 years in prison. If one prisoner betrays, but the other remains silent, the one that betrays will be set free and the other one will have 3 years in prison. Lastly, if both prisoners remain silent, each will have 1 year in prison.

The pay-off matrix of this game is then:

	B silent	B betrays
A silent	(-1, -1)	(-3, 0)
A betrays	(0, -3)	(-2, -2)

Table 2.1: The Prisoner's Dilemma. Terms in parenthesis indicate the pay-off of the players under the given strategy.

From an individual point of view, betraying always results in a better payoff than keeping silent regardless of the other's player action. Then, mutual betrayal is the only strong Nash equilibrium in the game (i.e. the only outcome from which each player could only do worse by unilaterally changing strategy). Hence, the dilemma consists in the fact that, although mutual cooperation would be more beneficial to both players than mutual defection, the most 'rational' strategy is betraying, i.e. is the best strategy from a self-interest point of view, while keeping silent would be irrational.

Classical game theory is commonly used to model human behaviour, i.e. it is commonly assumed that the Nash equilibria of the studied games can predict human decisions on a similar real-life problem. This point of view regarding classical game theory has been extensively criticised, arguing that the assumptions made do not hold for real human decision-making processes, as humans do not always act in a fully rational way (Chang et al. 2019; Newton 2018). On the one hand, individuals can exhibit bounded rationality, given by cognitive, informational or time constraints that do not allow them to compute the optimal strategy, so they rather choose a satisfactory option (Sent 2018). On the other hand, humans can simply make irrational decisions. That is, not choose the best option from their individual point of view, for example when they imitate others' behaviour (Pingle 1995).

These criticisms encouraged the emergence of the field of evolutionary game theory, which searches to resolve these issues by avoiding to assume that individuals are perfectly rational (Wallace et al. 2015).

In epidemiology, classical game theory has been extensively used to study voluntary vaccination

uptake in human populations (e.g. Bauch et al. 2004, 2003; Jijón et al. 2017; Mohr et al. 2020). In particular, the Prisoner’s Dilemma has been used for defining optimal strategies from the individual and from the group point of view. Assuming an initial high level of vaccine coverage, vaccinators are considered to be the Cooperators, while non-vaccinators are the Defectors, which *free-ride*, i.e. benefit from the herd-immunity levels acquired by the vaccinators, without having to bear the cost of the vaccine. At a low initial level of vaccine coverage, the Prisoner’s Dilemma is not however a relevant analogy, since what would be optimal for each individual would be to get vaccinated (Wang et al. 2016).

As mentioned before, classical game theory assumes that individuals optimise their payoffs in a selfish rational way. However, effects like social learning (i.e. learning by imitation) and bounded rationality (cognitive, informational or time constraints) are very important in vaccination decisions (e.g. Sturm et al. 2005). Furthermore, classical game theory mostly studies problems of only two players. In a context of many interacting agents, mean field game theory (Lasry et al. 2007) studies strategic decision making in such large populations by making a mean field simplification. That is, assuming atomic agents determine their strategy by considering the evolution of the group of players as a whole, rather than the set of individual behaviours. Indeed, each player searches to minimise an individual cost that depends on the agent’s state and action, and on a mean criterion based on the global state of the group of players. However, the predictive value of an approach relying on this simplification is limited for cases where disease and behaviour dynamics depend on a non-homogeneous population structure (Wang et al. 2016).

A particular game-theoretic framework that is worth mentioning here, is the one of *Markov decision processes* (MDP). MDP provide a mathematical approach for modelling decision making in situations where outcomes are partly random and partly under the control of a decision maker. They lie on the assumption that the conditional probability distribution of the future states of the dynamic system, given past states and the present state, only depends on the present state and not on past states (‘absence of memory’ of a Markov process). Hence, a MDP is given by a set of states X , a set of actions A , a (Markovian) transition model $P_a(x, x') = \mathbb{P}(x_{t+1} = x' | x_t = x, a_t = a)$ (probability that action a in state x at time t leads to state x' at time $t + 1$), and a reward function $Y_a(x, x') = P(y_{t+1} = y | x_t = x, a_t = a, x_{t+1} = x')$ (probability that y_{t+1} , the reward at time $t + 1$, equals y after transitioning from state x to state x' due to action a).

MDP are not specific to game-theory, as they are used and studied in other fields, such as reinforcement learning (see section 2.2.2.5). From the game-theory point-of-view, a MDP is a stochastic game with only one player. In epidemiology, it has been used by many authors to formulate game-theoretic models of vaccinating behaviour using mean-field and rationality assumptions (Wang et al. 2016). In particular, Reluga et al. 2011 proposes a general framework for vaccination games lying on three models: one CTMC for epidemiological changes at the individual scale (influenced by individual’s strategy and external factors such as population state), one mean-field deterministic model for the infection dynamics at the population level, and one model for individual decisions. The decision model lies on a combination of classical game theory postulates and MDP theory. Indeed, the first one assumes individuals maximise their inter-temporal expected utility, since they can predict their future evolution in probabilistic terms, i.e. integrate over the probability distribution of all possible paths of their future states. Meanwhile, MDP theory is used for the evaluation of decisions, that is, for associating values with events, and calculating the expected value conditional on each choice.

Yet, the criticisms made to classical game-theory can also be pointed out when using MDP within the game-theory framework. Indeed, within this framework, the MDP lies on classical game-theoretical assumptions (i.e. individuals are fully rational), and hence the questioning regarding the extent to which it can represent actual human decision-making still holds.

2.2.2.4 Evolutionary game theory

Here I discuss models that account for psychological or cognitive features of human behaviour but in which decisions are based on a certain payoff. I refer to them as *evolutionary game-theoretical*

models. Since such models are in a way inspired both by classical game theory and by psychology, they are sometimes referred to as game theoretical-psychological models.

Evolutionary game theory is the application of game theory to evolving populations in biology. It originated with the work of Maynard Smith 1982. Maynard Smith's evolutionary version of game theory does not require that players are rational, only that they have a strategy. The results of a game show how good that strategy was, like evolution shows how good strategies are for surviving and reproducing.

Following the works of Bauch 2005 and Fu et al. 2010, several studies subsequently used evolutionary game theory in human disease epidemiology to represent individual's vaccination behaviour in a social network (Tanimoto 2021) (e.g. Cardillo et al. 2013; Fukuda et al. 2014; Li et al. 2017; Liu et al. 2012; Voinson et al. 2015). The decision of the 'vaccination game' is usually played before or after the epidemic season which allows for a 'strategy updating', which determines the form of the influence of the neighbours in the social network on the decisions of each individual. See figure 2.7 for an example of a two-stage vaccination game considered in evolutionary-game theory applied to epidemiology.

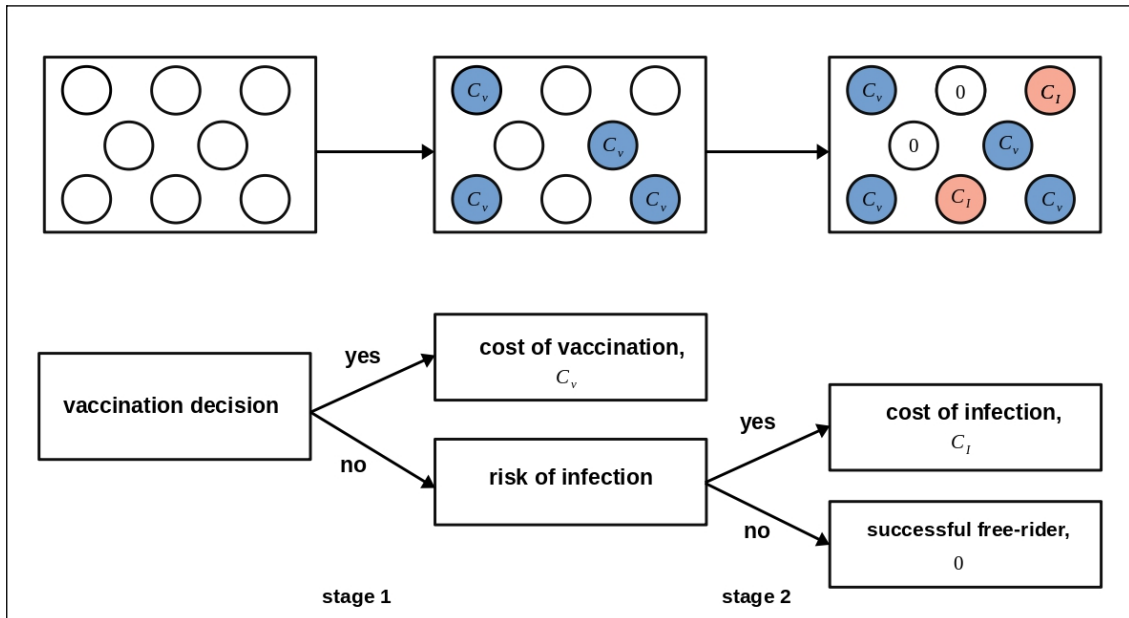


Figure 2.7: Scheme of a two-stage vaccination game. At stage 1, some individuals decide to vaccinate at a cost C_v . At stage 2, vaccinated individuals cannot get infected (vaccine provides perfect immunity) while non-vaccinated individuals can get infected (and hence pay a cost of infection C_I), or remain healthy (free-rider by being indirectly protected by the vaccination of other individuals). Reproduced from Fu et al. 2010.

There are several strategy updating rules in studies of evolutionary game theory applied to epidemiology (Tanimoto 2021). First, the *Imitation Max* (IM) rule, in which each player copies the strategy of the neighbour that has obtained the largest payoff in the previous decision time. At each decision time step, this rule is deterministic since the individual always copies the strategy of such neighbour.

The most studied stochastic updating rule is the *Fermi-Pairwise rule* (individual or socially based). A pairwise rule consists in randomly selecting a neighbour and copying the neighbour's strategy with a probability given by a certain function in which there is a comparison between the individual's payoff and the neighbour's payoff. The Fermi-Pairwise rule uses the Fermi function (Szabó et al. 1998) which reproduces the non-linearity of human decision-making processes (Pingle 1995; Tanaka et al. 2021). According to this rule, the probability of an individual i switching to the strategy of a selected neighbour j is

$$p(d_j \leftarrow d_i) = \frac{1}{1 + \exp[-\kappa(\Pi_j - \Pi_i)]}, \quad (2.8)$$

where d_i is the previous strategy of i , and d_j is the previous strategy of j . The parameter κ captures the individual's sensitivity to the difference on the last payoff of j and i , respectively Π_j and Π_i . In general, the Fermi-Pairwise rule is considered to be the strategy that most closely represents individual's vaccination behaviour, and is therefore the most used one in the literature (Tanaka et al. 2021). It has also been referred to as *IB-RA* (Individual Based Risk Assessment) by Fu et al. 2010. An extension of this rule called *SB-RA* (Strategy Based Risk Assessment) has been proposed by Kuga et al. 2019, where the probability of switching strategy depends on the comparison between the individual's payoff and the average payoff of neighbours with the contrary strategy.

Most of the strategies of evolutionary game theory applied to the vaccination game imply that the possibility of changing strategy only appears if the comparison occurs with a neighbour or group of neighbours with the opposite strategy. Then, if most of the population has initially the same strategy (vaccinate or not vaccinate), there would be a high chance that there are no changes when updating the individual strategies. An updating rule that does not have this issue is the Fermi-like myopic updating rule (Zhang et al. 2013). However, this rule does not consider imitation. Applied to the vaccination game, it assumes that individuals first have a perceived payoff of vaccinating and of not vaccinating. Then, they update their vaccination preference according to their own observation of the consequences of vaccinating or not, using a Fermi function.

The different updating rules have been used for evaluating the effect of subsidy policies on vaccination uptake (e.g. Kuga et al. 2019; Zhang et al. 2013). Indeed, since the cost of vaccination is usually assumed to be simply the cost of a vaccine dose, it is quite straightforward to consider that a subsidy policy that covers such a cost could impact vaccine willingness to vaccinate. Some studies have found that the updating rule that is considered for human decision-making can have an impact on the efficiency of the subsidy policy (e.g. Zhang et al. 2013), or on the pathogen spread (Ning et al. 2020). In particular, Kuga et al. 2019 found that even if results are qualitatively the same, in the case humans behave according to the *IB-RA* rule more individuals should be subsidised when compared to the case when they behave according to the *SB-RA*, since in the latter case there are more people that would vaccinate without subsidies.

The last point introduces the main limitation of this approach: *a priori*, it is unclear to what extent the chosen mechanism of individual decision-making has an impact on the model's dynamics (Tanimoto 2021) and therefore it should be carefully chosen (Wang et al. 2016). Indeed, it is often difficult to test on a representative sample if people actually take decisions in the way postulated by the updating rule, in particular regarding their response to different external incentives. When at least small data samples can be obtained, econometrical techniques can be used to statistically model individual's decisions (e.g. Sok et al. 2020, 2018). Ultimately, these could provide a data-driven baseline to compare models issued from an evolutionary game-theoretical approach.

2.2.2.5 Reinforcement Learning

Reinforcement learning (RL), is an area of machine learning in which optimal control policies are learned by the means of experience (by interacting through actions with a dynamic environment), as the environment's state transition probabilities or the rewards (observations) associated with each possible action are unknown (see figure 2.8). In the following, I review two main instances of RL: MDP (a discrete-time stochastic control process whose basic structure is described in section 2.2.2.3) and *multi-armed bandit* (MAB) problems.

Within RL, both MDP and MABs problems are generally concerned with a decision-maker that is a machine, not a human. Hence, they do not usually search to represent real-life human behaviour (i.e. how humans actually make decisions) but search the optimal sequence of actions that the agent should make, i.e. how the agent should make decisions in order to maximise the expected sum of rewards. Even when social learning is considered, RL approaches are not usually concerned with modelling how a group of human agents would make decisions, but with optimising their collaboration, in particular by enabling or improving communication between

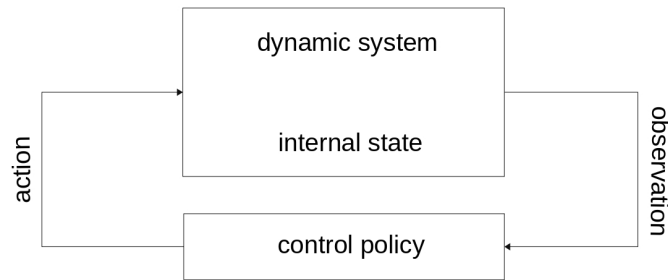


Figure 2.8: Schematic process of reinforcement learning. Reproduced from *RL fundamentals - 2021 RL Virtual School 2022*

them (e.g. Sankararaman et al. 2019). Therefore, MDP and MAB are essentially optimisation approaches. In a context of modelling human decision-making, they can be mostly useful for establishing mechanisms through which humans could behave optimally, regarding the maximisation of rewards in a non-dynamic system (MAB) or a dynamic system (MDP). This is, in a sense, the same criticism made to classical game theory, which supposes humans are able to optimise their expected utility.

However, although RL approaches usually do not refer to human decision-makers (as they do not consider human cognitive, psychological or social aspects), the MDP and MAB settings can be an appropriate framework for considering human-decision making through their relation with evolutionary game-theory. This association could surpass the limitation related to the rationality assumption when considering MDP with classical game theory (see section 2.2.2.3).

Given the previous considerations, the most important difference between the evolutionary game-theoretical models used in epidemiology and RL is that the first one considers that decision is made solely by a human, while the latter one usually considers the decision is taken or at last supported by a machine. Another important difference that is worth mentioning, is that in the latter agents *learn* from their interactions with the system and reinforce the probability of taking certain decisions (even if they can explore other options with a certain probability) given that they have memory. Meanwhile, in most evolutionary game-theoretical models, agents fully recompute at each time the probability of switching decision as a function of the last payoffs, i.e. memory is limited to the way agents' behaviours evolve as a function of the strategies and the proportions of groups of players with each strategy. That is, there is no memory associated to each player, i.e. no explicit impact of past decisions and payoff, nor of previous probabilities. One exception is the recent work of Wang et al. 2020 that considers a certain memory effect.

Regarding MDP, the main obstacle that arises when modelling real-world systems in their framework is the scalability problem. Given that in many applications the space state is combinatorial and extremely large, it is usually impossible to exactly calculate the *action-value function*, i.e. the function that measures long-term cumulative performance of a policy for each state within finite time, where a policy is a function that gives the probability of taking action a when the system is in state s (Sutton et al. 2018).

Hence, if the number of states is very large, simpler RL exploration methods can be more useful (Tokic et al. 2011). Another alternative is to consider an approximation of the value function (Sutton et al. 1999). In particular, Huang et al. 2020 use such an approximation for considering MDP within a setting of many agents (simulations are made on 400 agents) that interact in a dynamic environment. The approach shows new insights for the integration of evolutionary game theory and MDP. However, the association of the two suffers from some limitations related to the use of MDP itself. Apart from the questioning regarding its relevance in situations where decisions are not taken or advised by a machine, the approach lies on considering asynchronous decisions (only one agent decides at a time and other agents still use the actions from the previous decision time). This is necessary for being able to consider that at each time one agent is in interaction with a dynamic environment that evolves as a stationary Markov chain, and hence to formulate the decision-process of each agent as a MDP. This approach implicitly assumes that agents ig-

nore each other, so the problem is reduced to a single-agent problem, and interactions with other agents are possible only through the stochastic environment. Yet, under different conditions (in particular synchronous decisions) the actions of all agents would have to be explicitly considered in a multi-agent MDP (see section 2.2.3.3.3), and the curse of dimensionality for solving MDP would be enhanced (Bloembergen et al. 2015).

Since there are alternative RL approaches (such as the MAB problem) that could be scalable, they could have the potential of also addressing the concern on representing actual human-behaviour, provided that they imply less cognitive and computational resources for the decision-maker. In particular, if the process is not Markovian, or if the agent does not truly behaves optimally, the decision-making process of the agent can still be formulated within the MAB setting.

Regarding MAB, they are a particular type of RL problems, much simpler than MDP. A MAB is a sequential decision problem under uncertainty where an agent faces a number of options, called arms in the MAB terminology, each with an associated distribution of rewards. Since the agent does not know these distributions, it chooses an arm at each decision time, in order to optimise a previously established objective which depends on the rewards and thus on the agent's choices (Kaufmann 2014). In general, the objective on a MAB problem is to maximise the sum of rewards (in expectation) for a given time horizon T . To do this, the agent must explore the different options to obtain information on the distributions, but also exploit the options that maximise the empirical mean of the realised rewards. It will thus have to manage an exploration-exploitation trade-off.

The most studied MAB problem is the stochastic iid problem: the reward generating process is random, and reward observations are independent and identically distributed (i.i.d.) over time. That is, to each option corresponds an unknown distribution of rewards that is fixed in time, i.e. it is not modified either by the agent's actions or by external elements. The generalisation of the stochastic MAB problem which removes all assumptions on the distribution of the arms is the bandit adversarial problem. It supposes that at each iteration the agent chooses an arm and an 'adversary' simultaneously chooses the payoff structure for each arm. An example for adversarial bandits is the iterated prisoner's dilemma, where two players play prisoner's dilemma more than once in succession and they remember previous actions of their opponent and change their strategy accordingly. A strategy well adapted for the adversarial bandit problem is the *Exp3* adversarial algorithm (algorithm 1) proposed by Auer et al. 2002b. It consists in exploiting good arms (i.e. arms with higher weights) with probability $1 - \psi$, and possibly exploring other arms with probability ψ . After receiving the rewards, the weights are updated by exponentially increasing the weight of good arms.

One can remark the similarity between the exponential weight scheme for updating probabilities used in the *Exp3* adversarial MAB algorithm (algorithm 1), and the one in the Fermi-Pairwise rule used for computing the probability of switching decision (equation 2.8). Similarities between evolutionary game theory (not restricted to its application in epidemiology) and RL (not restricted to MAB) had been previously pointed out, mostly in the economic field (Börgers et al. 1997; Vega-Redondo 2003). Their relation has however received less attention from the RL community (Tuyls et al. 2005). In the field of epidemiology, Shi et al. 2019 considered RL and evolutionary game theory in a context of human decision-making regarding voluntary vaccination. Yet the two approaches are not actually combined together, as they refer to two different steps of the decision-making process. Indeed, a very simple MAB algorithm is used to dynamically determine if a human agent rationally decides according to the local prevalence of the disease, or if he/she imitates his/her neighbours.

More generally, evolutionary games settings can directly be formulated as MAB problems, since most of the heuristics in evolutionary game theory define MAB strategies that do not search to be optimal (i.e. in which human aspects of decision-making are taken into account). Yet, as stated before, agents actually learn in MAB strategies, i.e. they explicitly take into account the history of past decisions to make their following decisions, through the probabilities for the decision options. Furthermore, from a methodological point-of-view, they are generally described by an algorithm that mathematically formalises all the steps of the decision-making process.

Algorithm 1 Exp3 adversarial MAB algorithm (Auer et al. 2002b)**Input:** $\psi \in (0, 1]$, K arms.

-
- | | |
|---|--|
| 1: $w_k(1) = 1$ for $k = 1, \dots, K$ | ▷ Weights initialisation |
| 2: for $t = 1, 2, \dots, T$ do | ▷ At each decision time |
| 3: $p_k(t) = (1 - \psi) \frac{w_k(t)}{\sum_{i=1}^K w_i(t)} + \frac{\psi}{K}; k = 1, \dots, K$ | ▷ Compute the probability of pulling arm k at time t as a function of the weights of all arms and the exploration parameter ψ |
| 4: $a_t \leftarrow \text{Multinom}(1, (p_1(t), \dots, p_K(t)))$ | ▷ Draw an arm a_t randomly according to the probabilities $p_1(t), \dots, p_K(t)$ |
| 5: $x_{a_t}(t) \in [0, 1]$ | ▷ Observe a reward given by the pulled armed and the environment |
| 6: | |
| 7: for $k = 1, \dots, K$ do | ▷ For each arm |
| 8: $\hat{x}_k(t) = \begin{cases} x_k(t)/p_k(t) & \text{if } k = a_t \\ 0, & \text{otherwise} \end{cases}$ | ▷ Standardise the reward of the pulled armed by the probability of pulling it, to compensate for a potentially small probability of getting the observed reward |
| 9: $w_k(t+1) = w_k(t) \exp(\psi \hat{x}_k(t)/K)$ | ▷ Update the weight of the pulled arm k through exponential function of the observed standardised reward. The exponential growth significantly increases the weight of good arms |
| 10: end for | |
| 11: end for | |
-

2.2.3 Centralised decision-making: optimising the system

Controlling infectious diseases is a general and important problem that has been extensively treated in the epidemiology literature. In the following, I begin by precisising the usual control levers that are considered for infectious diseases on metapopulation networks. Then, I describe some strategies and mathematical approaches that can be used to achieve such control from a central social planner's point of view. Even if most of the approaches have been initially and mostly applied on well-mixed populations, the focus here is on the elements that have or could be applied on a metapopulation network.

2.2.3.1 Control levers and practical considerations

Measures for controlling pathogen spread on metapopulation networks are generally focused on either accelerating recovery (reducing illness duration), or decreasing the infection rate of each sub-population (Nowzari et al. 2015). The measures to attain the first of these two objectives focus on providing treatments to infected individuals, or improving existing conditions of treatment (for example through investments in health sector facilities). Regarding the second objective, there are several available measures from a centralised point-of-view, including the two previous ones as they can also impact the infection rate. Another measure is the distribution of vaccines among sub-populations. Other types of measures are movement restrictions between sub-populations (which can particularly imply quarantining some sub-populations), or acting on the awareness of the agents regarding the disease, for example through information campaigns, to increase the adoption of control measures (self-isolation, social distancing, hygiene measures) taken at the individual level. A measure that is specific to animal health is the culling of animals that have been tested positive, or even culling of all animals, not necessarily in infected herds, with a more preventive objective (e.g. for herds in the neighbourhood of an infected one).

The more radical is a measure, the less likely it will be considered and adopted, especially for a disease that does not have a rapid and strong economic and health burden. For example,

though movement restrictions can be very effective, as they seek to completely eliminate the risk of contagion by cutting the transmission path, they are also very costly in practice. In the case of movement restrictions, the agricultural and societal impact should be taken into account. Furthermore, particularly in the context of animal health, the cost of culling measures is not only economical (production losses) but it also affects animal well-being.

2.2.3.2 Allocation problems

The above considerations can be formally taken into account by the means of an *allocation problem*, consisting in distributing a fixed budget of some sort in order to best mitigate the spread of the pathogen on the network. The constraint can be for example on the maximal number of sub-populations to isolate or immunise. The budget constraint prevents the control measure from being applied to the whole metapopulation, which would certainly quickly reduce the pathogen spread but is rarely applicable in real-life, mainly due to high economic costs.

The control levers described in section 2.2.3.1, can be distinguished through a main classification regarding their effect: measures that impact the network structure and measures that impact epidemiological parameters within sub-populations. Following Nowzari et al. 2015, from an optimisation perspective, one can consider two broad problems related to the first type of measures, and a third problem regarding the second type of measures:

Problem 1 Given an original network G and a fixed budget $C > 0$ of nodes, minimise a function F , by removing (isolating or immunising) at most C nodes from G .

Problem 2 Given an original network G and a fixed budget $C > 0$ of links, minimise a function F , by removing at most C links from G .

Problem 3 Given a network G , given epidemiological parameters Θ , and a fixed budget $C > 0$ of a resource (that has an impact on Θ), minimise a function F , by distributing the budget among the N nodes of the network, under the constraint $\sum_i^{N^*} f(i) \leq C$, where $f(i)$ is the cost of allocating the resource to node i , and N^* is the number of nodes to which the resource is actually allocated.

Usually, these problems are formalised for a static network (Cohen et al. 2003; Tong et al. 2012; Wang et al. 2003), yet they could also be considered dynamically. The first problem consists in optimally "removing" (i.e. perfectly immunising or completely disconnecting) some nodes of the network so as to prevent or reduce pathogen spread, which we refer to as *node removal*. The second problem is *link removal*. This is an even more complex problem than the previous one, given that the number of links in the network can be much higher than the number of nodes. Related problems to Problems 1 and 2 can consider, instead of a complete removal, a certain type of control, consisting for example in reducing the intensity of the contacts on the links. The third problem consists in optimally allocating a limited budget of a certain resource, such as a treatment or an (imperfect) vaccine, on the nodes of the network, i.e. *resource allocation*.

An example of an allocation problem related to the generic problem 3 is worth to be mentioned. It can be formulated as:

Problem 4 Minimise the quantity of vaccines, such that herd immunity is reached.

In problem 4, the budget to minimise is usually considered to be a quantity of vaccines that provide perfect and life-long immunity, and the constraint concerns reaching a herd-immunity threshold. That is, the level at which a 'sufficient' percentage of the population has become immune against an infection to hinder infection-related dynamics.

This problem particularly concerns the first of the approaches that are described in section 2.2.3.3.

2.2.3.3 Control approaches

In the following, I present mathematical tools that can be concerned with either of the problems described above, i.e. when the budget is expressed in the number of nodes, number of links, or

in the quantity of a resource. Yet, the focus is made on the third problem, i.e. resource allocation, as the main motivation of this thesis are endemic livestock diseases, for which a strong control of such as movement bans is not very relevant. Furthermore, the nature of the budget to allocate can certainly determine the allocation strategy, e.g. the optimal allocation of imperfect vaccines may not coincide with the optimal node removal strategy. In addition, some approaches that are feasible for determining a node removal or resource allocation strategy, could be too costly, computationally speaking, to be deployed for finding the optimal link removal strategy. See works on network dismantling (Braunstein et al. 2016) or link rewiring (Schwartz et al. 2010) for a focus on approaches that are specifically adapted to node or link removal.

2.2.3.3.1 Spectral control For some of the considered problems, the function F to minimise (or the constraint to satisfy) usually depends on the network structure (Problems 1, 2 and 4). This is usually related to reaching the herd-immunity threshold.

For classical compartmental epidemiological models, the herd-immunity threshold can be expressed with respect to R_0 , the basic reproduction number of the infection (Keeling et al. 2011) (cf. section 2.1.1.1.1). It is possible to show that if a proportion $1 - \frac{1}{R_0}$ of susceptible individuals is perfectly immunised (the herd-immunity threshold), then the disease will be eradicated.

In a context of structured populations, much work has been devoted to determining the minimal fraction of each group that should be immunised (i.e. Problem 4), i.e. the level of vaccination for each sub-population that minimises the spectral radius of the next-generation matrix (cf. section 2.1.1.1.3), which is why such works can be grouped under the term of *spectral control* (Nowzari et al. 2015). These works usually consider a small number of groups, with strong interaction between them (e.g. Feng et al. 2017; Zhao et al. 2019). In particular, Hill et al. 2003 studied the properties of the solution to Problem 4, and computed the solution for an influenza A model stratified in age groups. However, within the framework of networks, the problem of minimising the spectral radius by removing a given number of nodes or links are respectively NP-complete and NP-hard (Mieghem et al. 2011), so heuristics have to be used to approximate its solution. Recently, Delmas et al. 2021, combined problems 1 and 4 into a bi-objective optimisation problem on an infinite-dimensional metapopulation SIS model, studied within the framework of a kernel model, and proved the existence and properties of its solutions.

Spectral control approaches have also been used for solving Problem 3 by assuming it is possible to tune the epidemiological parameters, such as the node or link dependent infection rates. Solutions often recur to geometric programming (Boyd et al. 2007), a type of constrained-optimisation problem with a particular form. The first main drawback is that, since they are based solely on the spectral radius, they do not take into account the state of the system, which can result in uselessly allocating resources to some nodes. Additionally, although some variants of the problem can be solved in polynomial time, this may not be sufficient for very large networks (Nowzari et al. 2015).

2.2.3.3.2 Optimal control theory *Optimal control theory* (OCT) is an analytical approach used for optimising the control of a dynamic system over a period of time (Lenhart et al. 2007). In epidemiology, it has been widely used for finding the optimal deployment of measures such as vaccination, treatment, or quarantine, (e.g. Behncke 2000; Sethi et al. 1978). Various costs have been considered, in particularly integrating economic considerations (e.g. Perrings et al. 2014).

Yet, the mathematical complexity of this method limits its use to the case of dynamic systems that can be described by simple models. Furthermore, the strategies identified as optimal can be highly complex, for example needing for a strategy switch at specific times. This can be beneficial for controlling pathogen spread if the exact switch times are known, notably if there is no uncertainty in model parameters, but can lead to poor results otherwise (Bussell et al. 2019).

Complex population structures are therefore rarely considered in OCT. In particular, Mbah et al. 2011; Rowthorn et al. 2009 apply it for resource allocation in a metapopulation of only two groups. Indeed, Zaric et al. 2001 showed that the resource allocation problem among n cou-

pled sub-populations is intractable. Hence, OCT usually relies on numerical approaches to find approximate solutions to these problems. However, such methods suffer from a lack of intuitive interpretation, which greatly limits their usefulness for the control of pathogen spread (Brandeau et al. 2003).

Bussell et al. 2019 proposed to overcome the limitations related to using optimal control in epidemiology, by the means of techniques consisting in approximating the epidemiological model by a simpler one, which facilitates the mathematical analysis necessary to OCT. A similar approach by Lemaitre et al. 2021, proposes a novel OCT framework for vaccine allocation on a metapopulation network of roughly 100 nodes using a deterministic COVID-19 model. This framework relies on a simplification of the full model that in particular reduces the problem to a tractable size, by considering only the most important links for deriving the optimal strategy, and the full network for evaluating it. The optimal control strategy provides a benchmark for the allocation on such metapopulation, as it outperforms all the other strategies. Yet, as the optimal strategy is complex and difficult to explain, the questioning regarding whether OCT strategies would be adopted by decision-makers in real-life control still holds.

2.2.3.3.3 Markov decision processes RL, in particular MDP, can not only be useful for decentralised decision-making as discussed in section 2.2.2.5, but also refer to the decisions taken by a social planner to control a system of many interacting agents. In this setting, the control can be formalised in two ways.

On the one hand, the social planner could take a decision for each agent of the system, in which case its decision process may be formulated as a centralised *multi-agent MDP*. On the other hand, decisions can concern the system as a whole (i.e. not be specific to each agent), in which case the social planner is the agent whose decision process is formalised as a *single-agent MDP*, as he/she should decide for the action to take at each decision time, and the interacting agents are considered to be part of the dynamic system (i.e. their number only increases the dimension of the MDP state space, not the dimension of the action space). This approach has already been applied in the context of livestock control by Viet et al. 2018, who proposed a MDP framework for a central social planner that searches to dynamically determine the incentive to implement a given measure, in order to optimise the collective financial benefit of a group of farmers. The incentive impacts farmers decisions, which are taken into account in a phenomenological manner, and farmers decisions impact the costs related to the disease and control.

In both cases, the decision-maker does not explicitly consider learning processes corresponding to each agent (i.e. each agent is not a decision maker on its own), which allows to use standard MDP policy evaluation algorithms (Xuan et al. 2002). However, this implies that irrespective to the approach, using MDP for controlling a very large multi-agent system suffers from the classic limitations of MDP regarding the *curse of dimensionality* in the system's space state, which is combinatorial in the number of agents. In particular Viet et al. 2018 consider a system of only $N = 50$ agents (where each agents is a herd). As each herd could either be in one of five health compartments, the number of states was given by $\binom{N+4}{4}$. Additionally, from an implementation point-of-view having such a large state space renders the optimal policy too complex to be used in the field, which is why authors proposes an approximated policy.

2.2.3.3.4 Score-based approaches The three previous approaches formalise and address the allocation problem by searching to directly determine the optimal allocation decisions. I refer here to approaches that, instead of seeking to directly determine the order in which the resource should be prioritised to the nodes of a network, aim to find the optimal score to rank the nodes (and not directly the optimal decisions) for the resource allocation. Such approaches are *score-based*.

The scores or indicators are most often *heuristics*, but can alternatively be derived from a certain optimisation procedure, i.e. *optimisation-based*. Furthermore, they can be *dynamic* or *static*, either from the network topology perspective, i.e. consider a dynamic or a static network, or from the pathogen spread perspective, i.e. taking or not into account the infection dynamics.

For a complex network of individuals, Pastor-Satorras et al. 2002 showed that the strategy consisting in randomly immunising individuals does not allow to eradicate the disease. Particularly for scale-free networks, authors argue that their structural properties should be accounted for. This implies to design the immunisation strategy taking into account nodes' connectivity, which would consist in identifying and targeting the most connected individuals. This is usually achieved through local metrics of the network topology, i.e. centrality measures, that can be directly used as heuristic scores to target the nodes. This includes classic static centrality measures (degree, PageRank, etc.) (cf. 2.1.3.1.1) and more recent temporal centrality measures, such as TempoRank (cf. section 2.1.3.2). Alternatively, the global network topology can be used to derive optimal scores. Examples of this are the *CURE* policy (Drakopoulos et al. 2014) and the *maxcut minimization* policy proposed by Scaman et al. 2016, where treatments are allocated according to a priority-order pre-computed offline, and the chosen priority-order is the one associated to the solution of a minimisation problem based on the network topology.

All these topologically-based scores do not explicitly consider the dynamics of the pathogen spread, but only the potential role each node may play in the pathogen spread based on the node's place in the network. Yet, this is not necessarily a shortcoming of topological scores, since detailed information on pathogen spread is rarely available. Therefore, although scores taking into account such information could be suited for a better control, scores that do not use updated epidemiological information and yet perform well are eligible for disease control.

An alternative is to use strategies that rely on epidemiological information. For example, it is possible to use data-driven or simulation-based approaches (Alamo et al. 2021). In particular, Venkatramanan et al. 2019 propose an optimisation procedure that at each decision step allocates vaccines to the sub-population that would lead to the maximum marginal reduction in the objective function according to simulations. This implies to run, at each decision step, a number of one-step simulations that is equal to the number of sub-populations.

Otherwise, it is possible to build approximations of the objective function (the function related to pathogen spread that the social planner searches to minimise, e.g. cumulative number of infected sub-populations over time) and seek to optimise these approximations. For example Dangerfield et al. 2019, consider the resource allocation problem in a meta-population of independent sub-populations, and approximate it as a *0-1 knapsack problem*, a problem that consists in finding the subset of items (each with a certain value and a weight) that maximise the total value of the selected items under a constraint on their total weight. Another example is the LRIE (Largest Reduction in Infectious Edges) strategy for treatment allocation found in Scaman et al. 2015, where the LRIE score is derived through a greedy minimisation of a certain function associated to the pathogen spread on an arbitrary and large network.

Finally, it can be possible to propose heuristics that are ultimately proven to correspond to optimal, or at least locally optimal solutions of a certain objective on the pathogen spread. As an example, the LRIE strategy was initially proposed in Scaman et al. 2014 as an heuristic consisting in targeting nodes whose recovery would minimise the number of links between infected and susceptible nodes. Then, it was proven to be the local solution to a certain optimisation problem (Scaman et al. 2015).

2.3 Decision-making regarding the control of disease spread on a trade network: problem formulation and selected approaches

In order to address the methodological challenges posed by dynamic decision-making regarding the adoption of control measures to limit the spread of a disease over a metapopulation network, and in the light of the state of the art presented in this chapter, I made the following modelling and formalism choices.

First, regarding the *modelling framework*, the motivation of the work (unregulated livestock dis-

eases spreading on an animal trade network) strongly suggests the use of a metapopulation network. Also, the internal sub-population infection dynamics and demographic changes are important features to be taken into account. Oversimplifying this internal dynamics could result in a loss of important characteristics of the system that emerge from this complexity. Furthermore, regarding the type of measures to consider, I decided to focus on within sub-population health management such as vaccination, instead of more stringent measures (such as trade movement restrictions). Indeed, for unregulated diseases and from a decentralised point-of-view it seemed unlikely that farmers would impose such strong restrictions on themselves. In addition, decisions involving measures that alter the network are more difficult to consider, especially in optimisation from a centralised perspective.

Second, concerning the *decentralised decision approach*, the choice was driven by the considerations in table 2.2, which summarises the main advantages and limitations of each approach presented in the state of the art. The following remarks were accounted for in the studies presented in chapters 3 and 5. Here, the main considered factor was the extent to which the approach could represent real-life human behaviour. Therefore, I chose not to consider phenomenological approaches as they could not show how behaviours emerge in the population. Also, even if classical game-theory studies strategic interactions among agents, the rationality assumption seemed a strong limitation in the representation of actual human behaviour under uncertainty. This issue is overcome by psychological models, yet these generally do not explicitly account for phenomena that can be important in the emergence of human behaviours, such as imitation dynamics. Finally, although a complex reinforcement learning approach (using Markov decision processes) could have been used, it offered little to account for real-life considerations, particularly for cognitive constraints in the farmers' decision-process.

<i>Approach</i>	<i>Advantages</i>	<i>Limitations</i>
Phenomenological	Simple way of accounting for human behaviour	Directly assumes an effect that could be caused by a decision, not the decision-making process. Hence, it is unclear whether the effect can actually emerge from the behaviour.
Classical game-theory	Rigorous mathematical framework under rationality assumption	Does not account for humans having limited rationality (either for cognitive or psychological reasons, or because they imitate each other).
Psychological	Accounts for individuals' limited rationality for cognitive and psychological reasons.	In general, does not consider limited rationality that can arise in a system of many interacting agents, for example through imitation.
Evolutionary game-theory	Can account for stochastic behaviour and human decision-making, as it makes no assumptions on the individual's rationality.	In general, individuals are not assumed to actually learn, as the probability of switching decision is not updated based on previous decision times, but only as a function of the new payoffs.
Reinforcement Learning	Can account for stochastic behaviour, learning (reinforcing good choices over time), and exploring. Algorithmic formalisation flexible enough for generic dynamic decision problems.	The decision-maker is essentially a machine, i.e. not intended to represent voluntary human behaviour, but to optimally solve a sequential decision problem by learning. Hence, algorithms do not usually consider human cognitive, psychological, and social aspects.

Table 2.2: Summary of the main advantages and limitations of decentralised decision-making approaches for modelling human-behaviour regarding voluntary control of an infectious disease.

There are similarities between the other two approaches, evolutionary game-theory and rein-

forcement learning, in particular between the Exp3 MAB adversarial algorithm of Auer et al. 2002b (algorithm 1) and the most used rule-of-thumb of evolutionary game-theory in epidemiology, the Fermi-Pairwise rule (equation 2.8). Both approaches, MAB and evolutionary game-theory, have limitations regarding their suitability to the context presented in this thesis. Indeed, on the one hand, the simple heuristics of evolutionary game-theory in epidemiology are based on cognitive considerations for representing human decision-making. Yet, such heuristics do not usually consider that the decision-makers learn, and are not formalised as algorithms, which limits their generalisation. On the other hand, MAB algorithms are not really intended to represent human decision-making. Yet, they seem appropriate for formalising a decision-making mechanism, since they explicitly consider a learning process.

For the above reasons, the two approaches (evolutionary game-theory and reinforcement learning) seemed the most suitable way of addressing the challenges regarding decentralised decision-making, particularly through the use of adapted multi-armed bandit algorithms inspired by the heuristics of evolutionary game theory.

Third, regarding the *centralised decision approach*, i.e. the optimisation of the system made by a social planner, the choice of the method was made by integrating the considerations in table 2.3 that summarises the reviewed approaches concerning allocation problems. These considerations underlie the work presented in chapter 4.

<i>Approach</i>	<i>Advantages</i>	<i>Limitations</i>
Spectral control	Directly addresses the topology of the population network	Does not consider the system state. Intractable for very high dimensions.
Optimal control theory	Can take into account the system state. Theoretical guarantees on the true optimal strategy.	Intractable for very high dimensions. When approximation-based approaches are used to address dimensionality, no intuitive interpretation of the optimal strategy.
MDP	Can take into account the system state. Theoretical guarantees on the true optimal strategy. Useful in the absence of a mathematical model of the dynamic system, when used with a RL perspective (learn through experience).	Intractable for very high dimensions. When approximation-based approaches are used to address dimensionality, no intuitive interpretation of the optimal strategy.
Score-based	Some can take into account the system state. Approaches can be scalable and strategies have an intuitive interpretation.	No guarantee on achieving a global optimum, i.e. may not solve the allocation problem in an optimal way.

Table 2.3: Summary of the main advantages and limitations of centralised decision-making approaches for optimising the control of an infectious disease on a network-structured system.

Here, the main factor that motivated the choice was the scalability of the approach. Indeed, the motivation of this work concerns very large networks of animal populations. Approaches ensued from spectral control did not appear as computationally feasible for explicitly taking into account the structure of very large networks. Furthermore, they did not allow to integrate the system's epidemiological state. Optimal control theory and MDP approaches seemed in theory particularly well-suited for controlling the system by determining the sequence of allocation decisions that would optimise a global objective. However, they could not be used for such a large problem without resorting to approximations that no longer have guarantees of optimality and that neglect the network structure or the intra-node dynamics. Furthermore, while using such approximations allows for dealing with relatively large networks, the policies have no intuitive interpretation and seem therefore not suitable for providing advice to real-life decision-makers on how to allocate the available resource.

Hence, the choice concerning centralised decision-making fell on score-based approaches. Al-

though those methods do not necessarily have optimality guarantees, they are scalable and can allow an intuitive interpretation of the allocation policies, which could motivate their adoption by a central social planner.

Chapter 3

Accounting for farmers' control decisions in a model of pathogen spread through animal trade

Cristancho Fajardo, L., Ezanno, P., and Vergu, E. (2021). "Accounting for farmers' control decisions in a model of pathogen spread through animal trade". In: *Scientific Reports* 11. DOI: 10.1038/s41598-021-88471-6

Contents

3.1	Introduction	50
3.2	Integrative model	51
3.2.1	Epidemic model with demography in a metapopulation based on a trade network	51
3.2.2	Farmers' decision-making model	54
3.2.3	An epidemic control measure	56
3.2.4	An economic-epidemiological cost function	57
3.3	Setting for simulations and sensitivity analyses	58
3.3.1	Fixed simulation setting	58
3.3.2	Parameters of the integrative model	64
3.3.3	Sensitivity analysis experiments	64
3.4	Results of the simulation study and sensitivity analyses	66
3.4.1	Model predictions for different decision scenarios	66
3.4.2	Key determinant parameters of decision-making and epidemiological dynamics	74
3.5	Discussion	76

Accounting for individual decisions in mechanistic epidemiological models remains a challenge, especially for unregulated endemic animal diseases for which control is not compulsory (section 3.1). We propose an original integrative model built by combining two sub-models. The first one corresponds to the dynamics of a livestock epidemic on a metapopulation network, grounded on demographic and animal trade data (section 3.2.1). The second one describes farmers' behaviour regarding the adoption of a control measure against the disease spread in their herd, i.e. decentralised decision-making (section 3.2.2). The measure is specified as a protective vaccine (section 3.2.3) with given economic implications (section 3.2.4), and the model is numerically studied through intensive simulations and sensitivity analyses (sections 3.3 and 3.4). This study is relevant for the understanding of the interplay between decision related human behaviour and

livestock epidemic dynamics (detailed discussion in section 3.5). The model can be used for other structures of epidemic models or different interventions, by adapting its components.

The modelling choices in this chapter were motivated by the state of the art in section 2.1.1.1.2, section 2.1.1.1.4 and section 2.1.3.4, regarding epidemiological modelling, and in section 2.2.2 and section 2.3, regarding decentralised decision-making modelling.

3.1 Introduction

Fighting livestock diseases spreading through animal trade is a major issue to guarantee sustainable farming, competitive agrifood chains and public health (Tomley et al. 2009). Epidemic prevention and reduction of prevalence require improved methods of control and compliance of the actors, especially for non-regulated diseases for which control decisions are left to individual or collective initiatives (Carslake et al. 2011). Mechanistic epidemiological models can provide a refined mathematical description and understanding of the complex system involved in pathogen spread, and be used to assess the effectiveness of control measures. They are complementary to observational or experimental approaches (Keeling et al. 2011). However, accounting for human behaviour in such models in order to increase their predictive power remains a challenge (Hidano et al. 2018; Manfredi et al. 2013), in particular for livestock diseases spreading through a trade network.

Indeed, most works on infectious diseases that consider the adoption of control measures usually do not account for human decision-making (Wang et al. 2016) or they do not consider a structured population (Morin et al. 2013; Perrings et al. 2014). In particular, in models based on a metapopulation over an explicit network (Keeling et al. 2005) such as Colizza et al. 2008a, control decisions are usually assumed to be taken at a centralised level (e.g. Braunstein et al. 2016; Scaman et al. 2016). Furthermore, when human decision-making is explicitly taken into account, it generally focuses on the context of human diseases (e.g. Fu et al. 2010; Kuga et al. 2019; Wang et al. 2012), but it has barely been applied to veterinary epidemiology yet (Horan et al. 2010). However, in the context of animal diseases, the decision of implementing control measures can be much more influenced by economic considerations than for human diseases, an aspect that should be taken into account in the decision model. Finally, in the field of veterinary epidemiology, studies have been mostly focused on regulated diseases, so human behaviour mainly consists in delaying the application of a central policy (e.g. Tago et al. 2016). In the few works that investigate control measures for unregulated animal diseases, there is generally no dynamic feedback on decision due to epidemic evolution (Rat-Aspert et al. 2010). Additionally, some real-life aspects, such as farmers having limited rationality or learning are generally missing (Krebs et al. 2018; Shi et al. 2019). There is therefore a special need for models combining the dynamics of an epidemic process that takes place on a livestock trade network, and the behaviour of farmers regarding the voluntary implementation of control measures (Brooks-Pollock et al. 2015; Heesterbeek et al. 2015).

We build an integrative model that can meet this need by coupling the dynamic spread of a livestock disease over a structured metapopulation, and the dynamics of the human decision-making process for applying a sanitary measure against the epidemic spread. To model the epidemic spread through a trade network we use a stochastic compartmental model that takes into account demographic dynamics and animal exchanges. The population structure of the model is calibrated using real data. Our decision model is inspired by previous studies (Fu et al. 2010; Kuga et al. 2019; Wang et al. 2012), in which the result of a decision regarding the voluntary adoption of a control measure for a human disease is evaluated after being applied, and preferences over the possible decisions are updated through time. We specifically formalise the dynamic decision problem that each farmer faces, and propose a mechanism that represents farmers' decision-making process in such a context. Our formalisation considers some real-life phenomena that can be present in the context of human decision-making: stochastic behaviour, learning, and the emergence of imitation and free-riding (Hershey et al. 1994).

First, we present the two components of the integrative model: the epidemic-demographic model, and the decision model. Then, we describe the specific control measure we consider, as well as its economic implications. Later, we describe the setting for simulations and sensitivity analyses we perform on the model, as well as the results of this simulation studies. Finally, we examine the model as well as the results of numerical explorations.

3.2 Integrative model

In the following, we describe the two main components of our integrative model: the epidemic-demographic one, and the decision-making one. We then detail the integrative model by considering vaccination as a specific control measure.

3.2.1 Epidemic model with demography in a metapopulation based on a trade network

For this work we place ourselves in the context of a hypothetical livestock infectious disease that is transmitted only through a contact network structure consisting in herds that exchange animals. This population structure is inspired by real data on animal movements, extracted from the French Cattle Identification Database (FCID). We assume this is a closed metapopulation, that is, we neglect exchanges with herds outside of it. This livestock trade network can be described as a directed weighted time-varying network, where nodes represent herds and links represent animal trade. The direction of each link is determined by the transfer's direction, and its weight corresponds to the amount of animals exchanged. By nature, this network is time-varying since links may change over time. In fact, not only trade connections may appear or disappear, but the amount of animals exchanged can vary on a daily basis.

Given this trade network, we consider an infectious livestock disease that can potentially be spread on it, and that can only be directly transmitted between animals within the same herd. The disease is assumed to be spread between herds only by animal transfers, as can be observed for diseases such as paratuberculosis (Beaunée et al. 2015), bovine tuberculosis when there is no contact with wildlife (Cousins 2001), and porcine reproductive and respiratory syndrome virus (Mortensen et al. 2002). In addition, the infection risk and status are assumed independent of animal breed, age or sex. In the absence of any intervention, the intra-herd disease spread is described by a stochastic SIR model (Keeling et al. 2011) with demography, accounting for animal transfers over the trade network. In a compartmental SIR model, the population is divided into three compartments: Susceptible (S), Infected (I), and Recovered (R), according to their health status. The only two possible transitions in a basic SIR model correspond to infection ($S \rightarrow I$) and recovery ($I \rightarrow R$). The implicit modelling assumptions we make are the following: intra-herd homogeneous mixing, meaning that the contact rate is the same among all the animals in a given herd; absence of a latent period, i.e. animals become infectious as soon as they are infected; acquisition of immunity after recovery; no vertical transmission, i.e. no mother-to-child transmission during pregnancy or childbirth; frequency-dependent intra-herd transmission, i.e. the transmission rate depends on the proportion of infected animals in the herd, rather than on their number; variation in time of herd size due to births, deaths and animal transfers, which we assume are not affected by the disease prevalence. The intra-herd transmission of the disease (without any intervention) can therefore be described by the scheme in figure 3.1.

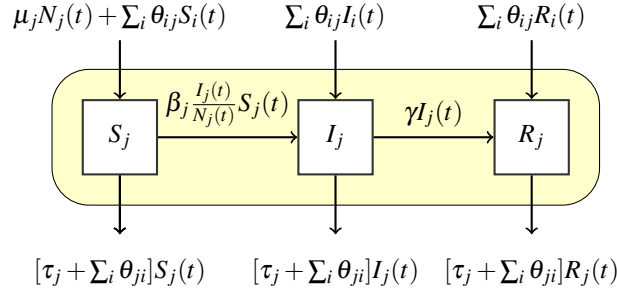


Figure 3.1: Schematic representation of the intra-herd epidemic-demographic dynamics for a herd j , without any control measure. Horizontal arrows represent transitions between health-related compartments, corresponding to the course of infection inside the herd (yellow rectangle), while vertical arrows represent population flows to and from the herd. The coefficients on the arrows are the transition rates.

Formally, we consider J herds in the population. Without any control-related measure, for each herd $j = 1, \dots, J$ the intra-herd transmission of the disease between times $t \geq 0$ and $t + \delta_t$ where $\delta_t > 0$, is described by a stochastic SIR model with demography, characterised by the following updating equations:

$$S_j(t + \delta_t) = S_j^*(t) + BS_j(t) + \overleftarrow{S}_j(t), \quad (3.1)$$

$$I_j(t + \delta_t) = I_j^*(t) + N_{S_j \rightarrow I_j}(t) + \overleftarrow{I}_j(t), \quad (3.2)$$

$$R_j(t + \delta_t) = R_j^*(t) + N_{I_j \rightarrow R_j}(t) + \overleftarrow{R}_j(t), \quad (3.3)$$

where $S_j(t)$, $I_j(t)$, $R_j(t)$ are the number of susceptible, infected and recovered animals in herd j for $t \geq 0$. We suppose $S_j(0) > 0$ for all j , $I_j(0) > 0$ for at least one herd j , $R_j(0) = 0$ for all j . We note as $N_j(t) = S_j(t) + I_j(t) + R_j(t)$ the size of herd j at time t . The terms $N_{X_j \rightarrow Y_j}(t)$ represent the number of animals going from epidemic state X to Y in herd j in the time interval $]t, t + \delta_t]$. $BS_j(t)$ is the number of births. $\overleftarrow{S}_j(t)$, $\overleftarrow{I}_j(t)$, $\overleftarrow{R}_j(t)$, are the number of susceptible, infected and recovered animals (respectively) purchased by herd j . Finally, $S_j^*(t)$, $I_j^*(t)$, $R_j^*(t)$ are the number of susceptible, infected and recovered at time t that remain in the same state after all outgoing flows are accounted for. Each one of the terms results from a stochastic process. In order to precise the probabilities of the different possible change events in the status of an animal, we first define the following epidemic and demographic daily rates:

- β_j : daily rate of transmission of the disease in herd j . That is, the rate at which a susceptible animal gets infected by contact with an infected one in herd j . It is equal to the product between the contact rate in the herd, and the transmission rate of the disease if there is contact. Then, the rate at which a susceptible becomes infected by contact with the $I_j(t)$ infected is $\beta_j \frac{I_j(t)}{N_j(t)}$.
- γ : self-recovery rate of the disease. That is, the rate at which the infected individuals recover from the disease. Equivalent to the inverse of the mean duration (in days) of the infectious period.
- τ_j : daily crude death rate in herd j . That is, the rate at which the animals are removed from herd j (not related to the disease).
- μ_j : daily crude birth rate in herd j . It is the rate at which new animals are born in a herd.
- $\overrightarrow{\theta}_j$: daily out rate from herd j . It corresponds to $\sum_{i \neq j} \theta_{ji}$ where θ_{ji} is the daily out rate from herd j to herd i , that is, the mean number of animals going from herd j to herd i in a day, over the mean population of herd j in a day.

The stochasticity is considered in continuous time, but simulated in discrete time. The simulation in discrete time appears as more natural given available data, and is computationally more efficient. Formally, we consider a continuous time Markov chain simulated by an Euler simulation scheme that uses coupled discrete-time multinomial processes, as described in Bretó et al. 2009. In order to increase computational efficiency, we adapt this scheme by splitting it into two steps: one for the intra-herd epidemic-demographic dynamics, and one for the inter-herd dynamics.

In the first step, we use the coupled Multinomial chains defined by:

$$N_{S_j \rightarrow I_j}(t), N_{S_j \rightarrow D_j}(t), \vec{S}_j(t), S_j^*(t) \leftarrow \text{Multin} \left(S_j(t), \left[p_{SI_j}, p_{SD_j}, p_{\vec{S}_j}, 1 - (p_{SI_j} + p_{SD_j} + p_{\vec{S}_j}) \right] \right), \quad (3.4)$$

$$N_{I_j \rightarrow R_j}(t), N_{I_j \rightarrow D_j}(t), \vec{I}_j(t), I_j^*(t) \leftarrow \text{Multin} \left(I_j(t), \left[p_{IR_j}, p_{ID_j}, p_{\vec{I}_j}, 1 - (p_{IR_j} + p_{ID_j} + p_{\vec{I}_j}) \right] \right), \quad (3.5)$$

$$N_{R_j \rightarrow D_j}(t), \vec{R}_j(t), R_j^*(t) \leftarrow \text{Multin} \left(R_j(t), \left[p_{RD_j}, p_{\vec{R}_j}, 1 - (p_{RD_j} + p_{\vec{R}_j}) \right] \right), \quad (3.6)$$

$$BS_j(t) \leftarrow \text{Bin} \left(N_j(t), p_{B_j} \right), \quad (3.7)$$

where $N_{S_j \rightarrow D_j}(t)$, $N_{I_j \rightarrow D_j}(t)$, $N_{R_j \rightarrow D_j}(t)$ are the number of deaths of susceptible, infected, and recovered individuals, respectively. $\vec{S}_j(t)$, $\vec{I}_j(t)$, $\vec{R}_j(t)$ are the number of susceptible, infected and recovered animals (respectively) sold in the same period by herd j . Each term p_{XY_j} refers to the probability of an animal in herd j of going from epidemiological compartment X to the compartment Y , and $p_{\vec{X}_j}$ refers to the probability that an animal of compartment X is sold. These probabilities are defined as:

$$p_{SI_j} = \frac{\left(1 - e^{-\left(\beta_j \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j \right) \delta t} \right) \beta_j \frac{I_j(t)}{N_j(t)}}{\beta_j \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j}, p_{SD_j} = \frac{\left(1 - e^{-\left(\beta_j \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j \right) \delta t} \right) \tau_j}{\beta_j \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j}, \quad (3.8)$$

$$p_{\vec{S}_j} = \frac{\left(1 - e^{-\left(\beta_j \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j \right) \delta t} \right) \vec{\theta}_j}{\beta_j \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j},$$

$$p_{IR_j} = \frac{\left(1 - e^{-(\gamma + \tau_j + \vec{\theta}_j) \delta t} \right) \gamma}{\gamma + \tau_j + \vec{\theta}_j}, p_{ID_j} = \frac{\left(1 - e^{-(\gamma + \tau_j + \vec{\theta}_j) \delta t} \right) \tau_j}{\gamma + \tau_j + \vec{\theta}_j}, p_{\vec{I}_j} = \frac{\left(1 - e^{-(\gamma + \tau_j + \vec{\theta}_j) \delta t} \right) \vec{\theta}_j}{\gamma + \tau_j + \vec{\theta}_j}, \quad (3.9)$$

$$p_{RD_j} = \frac{\left(1 - e^{-(\tau_j + \vec{\theta}_j) \delta t} \right) \tau_j}{\tau_j + \vec{\theta}_j}, p_{\vec{R}_j} = \frac{\left(1 - e^{-(\tau_j + \vec{\theta}_j) \delta t} \right) \vec{\theta}_j}{\tau_j + \vec{\theta}_j}, \quad (3.10)$$

$$p_{B_j} = 1 - e^{-\mu_j \delta t}. \quad (3.11)$$

In the second step, we randomly assign a destination to each exiting animal of the herd $j = 1, \dots, J$. Let $Ch(j) = \{i = 1, \dots, J : \theta_{ji} \neq 0\}$, and let $j_1^*, \dots, j_{|Ch(j)|}^*$ be the elements of that set, where $|Ch(j)|$ is its cardinality. That is, the herds for which herd j is a seller according to $\theta_{ji}; i = 1, \dots, J$. We define $S_{jj^*}(t)$, $I_{jj^*}(t)$, $R_{jj^*}(t)$ as the number of susceptible, infected and recovered animals, respectively, sold by herd j to herd j^* in $[t, t + \delta t]$. So that $\vec{S}_j(t) := \sum_{j^* \in Ch(j)} S_{jj^*}(t)$, $\vec{I}_j(t) := \sum_{j^* \in Ch(j)} I_{jj^*}(t)$, $\vec{R}_j(t) := \sum_{j^* \in Ch(j)} R_{jj^*}(t)$. Then, the assignation of destination for sold animals of herds $j = 1, \dots, J$ is obtained through the following coupled multinomial chains:

$$S_{jj_1^*}(t), \dots, S_{jj_{|Ch(j)|}^*}(t) \leftarrow \text{Multin} \left(\vec{S}_j(t), \left[\vec{p}_{jj_1^*}, \dots, \vec{p}_{jj_{|Ch(j)|}^*}, 1 - \sum_{j^*=j_1^*, \dots, j_{|Ch(j)|}^*} \vec{p}_{jj^*} \right] \right), \quad (3.12)$$

$$I_{jj_1^*}(t), \dots, I_{jj_{|Ch(j)|}^*}(t) \leftarrow \text{Multin} \left(\vec{I}_j(t), \left[\vec{p}_{jj_1^*}, \dots, \vec{p}_{jj_{|Ch(j)|}^*}, 1 - \sum_{j^*=j_1^*, \dots, j_{|Ch(j)|}^*} \vec{p}_{jj^*} \right] \right), \quad (3.13)$$

$$R_{jj_1^*}(t), \dots, R_{jj_{|Ch(j)|}^*}(t) \leftarrow \text{Multin} \left(\vec{R}_j(t), \left[\vec{p}_{jj_1^*}, \dots, \vec{p}_{jj_{|Ch(j)|}^*}, 1 - \sum_{j^*=j_1^*, \dots, j_{|Ch(j)|}^*} \vec{p}_{jj^*} \right] \right), \quad (3.14)$$

where the probabilities $\vec{p}_{jj^*}; j^* = j_1^*, \dots, j_{|Ch(j)|}^*$ are defined according to the rates θ_{jj^*} analogously to the previous multinomial chains:

$$\vec{p}_{jj^*} = \frac{\left(1 - e^{-\delta t \sum_{j^* \in Ch(j)} \theta_{jj^*}}\right) \theta_{jj^*}}{\sum_{j^* \in Ch(j)} \theta_{jj^*}}. \quad (3.15)$$

Finally, the terms referring to the entries in herd j are computed as: $\overleftarrow{S}_j(t) := \sum_{i \neq j} S_{ij}(t)$, $\overleftarrow{I}_j(t) := \sum_{i \neq j} I_{ij}(t)$, and $\overleftarrow{R}_j(t) := \sum_{i \neq j} R_{ij}(t)$.

3.2.2 Farmers' decision-making model

We suppose that farmers can apply a sanitary measure that has a certain efficacy on the disease spread for a limited amount of time. Then, we assume they search to take the control decision that allows them to obtain an optimal value of an individual criterion, i.e. an expected cost resulting from the decision. To consider a simple and clear framework, we restrict ourselves to binary decisions (the measure is applied or not). Additionally, we make the assumption that decision times are synchronized, discrete, and equally spaced in time. This proves to be useful when considering the interaction of farmers' decisions, and for evaluating the effect of the time length between successive decisions on the behaviour of the integrative model. Formally, we suppose that each farmer $j = 1, \dots, J$ searches to solve:

$$\min_{a_j^t} \mathbb{E} \left[C_{a_j^t}^t(j) \right] \quad ; \quad t = \Delta_d, 2\Delta_d, 3\Delta_d, \dots, \quad (3.16)$$

where $t = \Delta_d, 2\Delta_d, 3\Delta_d, \dots$ are the decision times, so Δ_d is the duration (in days) between any two consecutive decisions. It also determines the instant at which the first decisions are taken. The term $a_j^t \in \{0, 1\}$ refers to the control decision: if $a_j^t = 1$, the control-measure is applied in herd j at decision time t , otherwise it is not. $C_{a_j^t}^t(j)$ is the cost in herd j associated with the decision taken at time t . This constitutes a dynamic decision-making problem under uncertainty, this latter affecting the cost distribution associated with each possible decision.

To define the farmer's decision-making process that attempts to solve this problem, we take an approach inspired by Fu et al. 2010; Kuga et al. 2019; Wang et al. 2012, in which farmers evaluate the result of a decision after its application, and update their preferences over time as a function of this result. In this particular context, this approach seems suitable for several reasons. First, as we mentioned earlier, costs over time not only depend on the epidemic and decision dynamics in the herd where the decision is made, but also on other herds. To exactly solve this optimization problem would imply that farmers integrate the actions and epidemic status of other herds. This is a very complex problem due to the dimensionality on the number of herds, and on the possible status of the system. Second, since we pose a dynamic decision-making problem, there is an effect of learning through repeated decisions. Indeed, we have supposed that the cost associated to a decision is observed before making the next decision. Then, it is natural to think that farmers learn from the costs they have obtained with their previous choices, to take their next decisions. Finally, through this approach we can easily consider social dynamics such as imitation effects between farmers. In our context, this consists in considering a stochastic decision mechanism where the probability of applying the measure is updated through the costs each farmer observes over time, and the costs observed by his/her neighbours.

The mechanism we propose (Algorithm 2) works by updating the probability of applying the measure, proportionally to an exponential weight that takes into account the last decision taken by the farmer and that taken by one of his/her neighbours, through a weighted sum of the associated costs. Then at each decision time, each farmer $j = 1, \dots, J$ takes a decision a_j^t using his/her current probability of applying the measure $p_1^t(j)$. We assume that this probability is initially the same for all farmers, and equal to a value p_1^{init} , and that each farmer observes the

Algorithm 2 Exponential weighting stochastic mechanism with imitation

Input: 2 options = $\{0,1\}$, $p_1^{\Delta_d}(j) := p_1^{init} \forall j$, $\kappa \geq 0$, $\rho \geq 0$, $B(j) = \{i; \theta_{ij} \neq 0 \text{ or } \theta_{ji} \neq 0\}; j = 1, \dots, J$.

- 1: **for** $t = \Delta_d, 2\Delta_d, 3\Delta_d \dots$ **do** ▷ At each decision time
- 2: **for** $j = 1, \dots, J$ **do** ▷ Each farmer
- 3: $a_j^t \leftarrow \text{Bernoulli}(p_1^t(j))$ ▷ Takes a decision using his/her current probability of applying the measure
- 4: $C_{a_j^t}^t(j)$ ▷ Observes the cost related to his/her decision
- 5: $j^* \leftarrow \text{Unif}(B(j))$ ▷ Selects one of his/her neighbours in the trade network
- 6: $(a_{j^*}^t, C_{a_{j^*}^t}^t(j^*))$ ▷ Observes the decision taken by j^* and his/her observed cost
- 7:
$$p_1^{t+\Delta_d}(j) = \frac{p_1^t(j)e^{-\kappa C_1^t(j) - \rho C_1^t(j^*)}}{p_1^t(j)e^{-\kappa C_1^t(j) - \rho C_1^t(j^*)} + (1 - p_1^t(j))e^{-\kappa C_0^t(j) - \rho C_0^t(j^*)}}$$
 (3.17)
- ▷ Updates the probability of applying the measure, where the costs of the non taken options are equal to 0, i.e. for $k = 0, 1$
 - $C_k^t(j) = C_{a_j^t}^t(j)$ if $k = a_j^t$, 0 otherwise.
 - $C_k^t(j^*) = C_{a_{j^*}^t}^t(j^*)$ if $k = a_{j^*}^t$, 0 otherwise.
- 8: **end for**
- 9: **end for**

cost related to his/her decision, and the decision and associated cost observed by one of his/her neighbours in the trade network, who is randomly chosen by the farmer. A neighbour of j in the trade network is a farmer with whom j exchanges animals according to the daily trade rates, i.e. a farmer j^* such that $\theta_{jj^*} \neq 0$, or $\theta_{j^*j} \neq 0$. In the algorithm, we note as $B(j)$ the set of neighbours of j in the trade network.

The update in the probability is then given by equation 3.17. The parameter κ represents farmer's "sensitivity" to his/her own observed costs. A κ close to zero implies that farmers are not very sensitive to their own observed costs, and therefore mostly rely on their initial probability of applying the measure, whereas a large κ represents the situation in which farmers are very sensitive to their own observed costs for updating their probability of applying the measure. For considering an imitation effect, we introduce the parameter ρ that works analogously to κ , but on the cost observed by the chosen neighbour. The parameters κ and ρ act then as weights to the farmer's and the neighbour's observed cost, respectively. In our model, farmer's next decision can be updated considering any of his/her neighbours, regardless of what the neighbour has decided in the previous step. Finally, for updating the probabilities, it is natural that these are set so that the decision with a smaller sum of weighted costs receives higher probability. Although there are many ways to turn the sum of weighted costs into probabilities, a simple and popular method is to use an exponential weighting scheme. This scheme quickly reduces the probability of the decision that has resulted to be very bad (high sum of weighted costs). This form is found in the Fermi-Pairwise rule, which has been previously used in similar contexts, as its stochastic behaviour is similar to real-life human decision-making (Nagashima et al. 2019). In this update, the cost associated to non-taken decisions are zero, i.e. either $C_1^t(j)$ or $C_0^t(j)$ is zero, and either $C_1^t(j^*)$ or $C_0^t(j^*)$ is zero. The non-zero costs define the final form of the probability update. In order to see the effect of the decisions and the associated observed costs in this update, we remark that since we consider binary decisions, equation 3.17 can be rewritten as an update on the odds of applying the measure:

$$odds_1^{t+\Delta_d}(j) = odds_1^t(j) \times e^{(1-2a_j^t)\kappa C_{a_j^t}^t(j) + (1-2a_{j^*}^t)\rho C_{a_{j^*}^t}^t(j^*)}, \quad (3.18)$$

where $odds_1^t(j) := p_1^t(j)/(1 - p_1^t(j)); \forall t = \Delta_d, 2\Delta_d, \dots$. From this we can see that the odds are reinforced or decreased as a result of the farmer's and the neighbour's decision and cost. If they

both apply the measure at time t , the term in the exponential is negative since costs are positive or zero, so the odds for j applying the measure decrease. Analogously, if neither of them applies the measure at time t , the term in the exponential is positive and the odds of applying it increase. Finally, if they do not make the same decision at time t it is the comparison between $\kappa C_{a_j^t}^t(j)$ and $\rho C_{a_{j^*}^t}^t(j^*)$ that determines the direction of the update.

Additionally, we explore an extension of the model where each farmer considers the decisions and costs observed by all of his/her neighbours at each decision time. To update his/her probability of vaccinating, he/she takes into account the costs observed by his/her neighbours who did not vaccinate at the previous decision time, and the costs observed by those who vaccinated, as described in Algorithm 3.

Algorithm 3 Exponential weighting stochastic mechanism with imitation

Input: 2 options = $\{0,1\}$, $p_1^{\Delta_d}(j) := p_1^{init} \forall j$, $\kappa \geq 0$, $\rho \geq 0$, $B(j) = \{i; \theta_{ij} \neq 0 \text{ or } \theta_{ji} \neq 0\}; j = 1, \dots, J$.

- 1: **for** $t = \Delta_d, 2\Delta_d, 3\Delta_d \dots$ **do** ▷ At each decision time
- 2: **for** $j = 1, \dots, J$ **do** ▷ Each farmer
- 3: $a_j^t \leftarrow \text{Bernoulli}(p_1^t(j))$ ▷ Takes a decision using his/her current probability of applying the measure
- 4: $C_{a_j^t}^t(j)$ ▷ Observes the cost related to his/her decision
- 5: $\overline{C_1^t}(B(j)), \overline{C_0^t}(B(j))$ ▷ Observes the mean cost observed by his/her neighbors that vaccinated, and the one observed by those that did not vaccinate. If no neighbor vaccinated/did not vaccinate, the mean associated cost of this group is zero
- 6:
$$p_1^{t+\Delta_d}(j) = \frac{p_1^t(j) e^{-\kappa C_1^t(j) - \rho n_1^t(j) \overline{C_1^t}(B(j))}}{p_1^t(j) e^{-\kappa C_1^t(j) - \rho n_1^t(j) \overline{C_1^t}(B(j))} + (1 - p_1^t(j)) e^{-\kappa C_0^t(j) - \rho n_0^t(j) \overline{C_0^t}(B(j))}} \quad (3.19)$$
▷ Updates the probability of applying the measure, where $n_1^t(j)$ and $n_0^t(j)$ are the number of neighbors of j that vaccinated, and the number that did not vaccinate, respectively. The costs of the non taken options are equal to 0, i.e. for $k = 0, 1$: $C_k^t(j) = C_{a_j^t}^t(j)$ if $k = a_j^t$, 0 otherwise.
- 7: **end for**
- 8: **end for**

3.2.3 An epidemic control measure

For the control measure that can be applied to manage the spread of the disease, we specifically consider a vaccine that can reduce the rate of disease transmission towards a susceptible vaccinated animal. We assume this is the only effect the vaccine has. We make the assumption that the vaccine maintains a constant efficacy during a certain time period, whose duration is the same as the decision time-step. Then, if the vaccine is applied on a susceptible animal in herd j at time t , the rate of transmission towards that susceptible animal during the period $]t; t + \Delta_d]$ will be $\beta_j^v = \beta_j(1 - e_v)$, where $0 \leq e_v \leq 1$ is the protection efficacy of the vaccine.

When considering such a vaccine, the epidemic-demographic model is modified as a consequence. Indeed, the susceptible compartment is divided in two sub-compartments: the non vaccinated susceptible animals (SNV_j) and the vaccinated ones (SV_j). We note as $SV_j(t)$ the susceptible animals in herd j at time t that were vaccinated at decision time t_d (where $t_d < t \leq t_d + \Delta_d$) either in herd j or elsewhere, and as $SNV_j(t)$ the susceptible animals in herd j at time t that were not vaccinated. We have then for $t \in]t_d; t_d + \Delta_d]$:

$$N_{SNV_j \rightarrow I_j}(t), N_{SNV_j \rightarrow D_j}(t), \overline{SNV_j^*}(t), SNV_j^*(t) \leftarrow \text{Multin} \left(SNV_j(t), \left[p_{SI_j}, p_{SD_j}, p_{\overline{S_j^*}}, 1 - (p_{SI_j} + p_{SD_j} + p_{\overline{S_j^*}}) \right] \right), \quad (3.20)$$

$$N_{SV_j \rightarrow I_j}(t), N_{SV_j \rightarrow D_j}(t), \overline{SV_j^*}(t), SV_j^*(t) \leftarrow \text{Multin} \left(SV_j(t), \left[p_{SVI_j}, p_{SVD_j}, p_{\overline{SV_j^*}}, 1 - (p_{SVI_j} + p_{SVD_j} + p_{\overline{SV_j^*}}) \right] \right). \quad (3.21)$$

Equations 3.20 and 3.21 replace equation 3.4 in the model. The transition probabilities of the non-vaccinated susceptible animals are exactly defined as before for the susceptible compartment in the SIR model without vaccination (equation 3.8). For the vaccinated susceptible animals, since we assume that the vaccine has only an effect on their probability to get infected after contact hence on the corresponding transmission rate, the related probabilities are modified as:

$$\begin{aligned}
 p^{SVI_j} &= \frac{\left(1 - e^{-\left(\beta_j^v \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j\right) \delta t}\right) \beta_j^v \frac{I_j}{N_j}}{\beta_j^v \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j}, & p^{SVD_j} &= \frac{\left(1 - e^{-\left(\beta_j^v \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j\right) \delta t}\right) \tau_j}{\beta_j^v \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j}, & (3.22) \\
 p_{SV_j}^{\vec{\theta}} &= \frac{\left(1 - e^{-\left(\beta_j^v \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j\right) \delta t}\right) \vec{\theta}_j}{\beta_j^v \frac{I_j(t)}{N_j(t)} + \tau_j + \vec{\theta}_j},
 \end{aligned}$$

3.2.4 An economic-epidemiological cost function

We assume that the farmers are able to assess the economic impact that their decisions have on the disease spread in their herd. Therefore, we define the costs on the basis of a simple economic cost function, related to the epidemiological consequences of the decision taken at t in herd j . We define it in particular for the considered control measure, a protective vaccine, but it can easily be modified for a control measure with a different impact on the epidemic transition rates. The cost function we considered is:

$$C_{a_j^t}^t(j) := \frac{[CF_v + CU_v N_j(t)] a_j^t + \phi r N_{S_j \rightarrow I_j}(t, t + \Delta_d)}{N_j(t, t + \Delta_d)}, \quad (3.23)$$

where in the numerator the first term refers to the cost farmers pay to apply the vaccine, and the second one to the economic impact of the epidemic consequences of the vaccine. Precisely, in the first term a_j^t equals 1 if the vaccine is applied on herd j at decision time t , and it equals 0 otherwise. CU_v is the unitary cost of the vaccine per animal, and CF_v defines a fixed cost of applying vaccination per herd. This would typically correspond to the cost of a veterinary visit. In the second term, r is the monetary value of a healthy animal, and $0 \leq \phi \leq 1$ is the rate of reduction of this value if the animal gets infected. So ϕr is the cost of an infection, that is, the loss in the monetary value of an animal if it gets infected. $N_{S_j \rightarrow I_j}(t, t + \Delta_d)$ is the number of new infections in the herd, from the moment decision is taken until the next decision time. Therefore, the benefit of having healthy animals is implicitly given by the animal not reducing its value due to an infection. We remark that we make the assumption that each farmer perfectly observes the number of new infections that occurred during the decision period, or at least the global loss in the monetary value of the herd $\phi r N_{S_j \rightarrow I_j}(t, t + \Delta_d)$ related to these new infections. However, farmers can not identify which animals are infected, which is why we assume they choose to vaccinate the whole herd if vaccination is decided. Finally, in order to account for differences in the costs that may only be related to the variation of the herd size over the period, we standardize the cost by the sum of the daily herd size during the concerned period, which we note as $N_j(t, t + \Delta_d) = N_j(t) + N_j(t + 1) + N_j(t + 2) + \dots + N_j(t + \Delta_d)$. This is equivalent to standardizing by $\Delta_d \bar{N}_j(t, t + \Delta_d)$, where $\bar{N}_j(t, t + \Delta_d)$ is the mean daily herd size over the period.

The scheme of the integrative model for vaccination can be found in figure 3.2. It shows the feed-back loop between the epidemic-demographic dynamics, and the decision dynamics. The epidemic-demographic process takes place for a period of length Δ_d , until a new decision is taken. This decision is itself a function of economic and epidemic consequences of the previous decision.

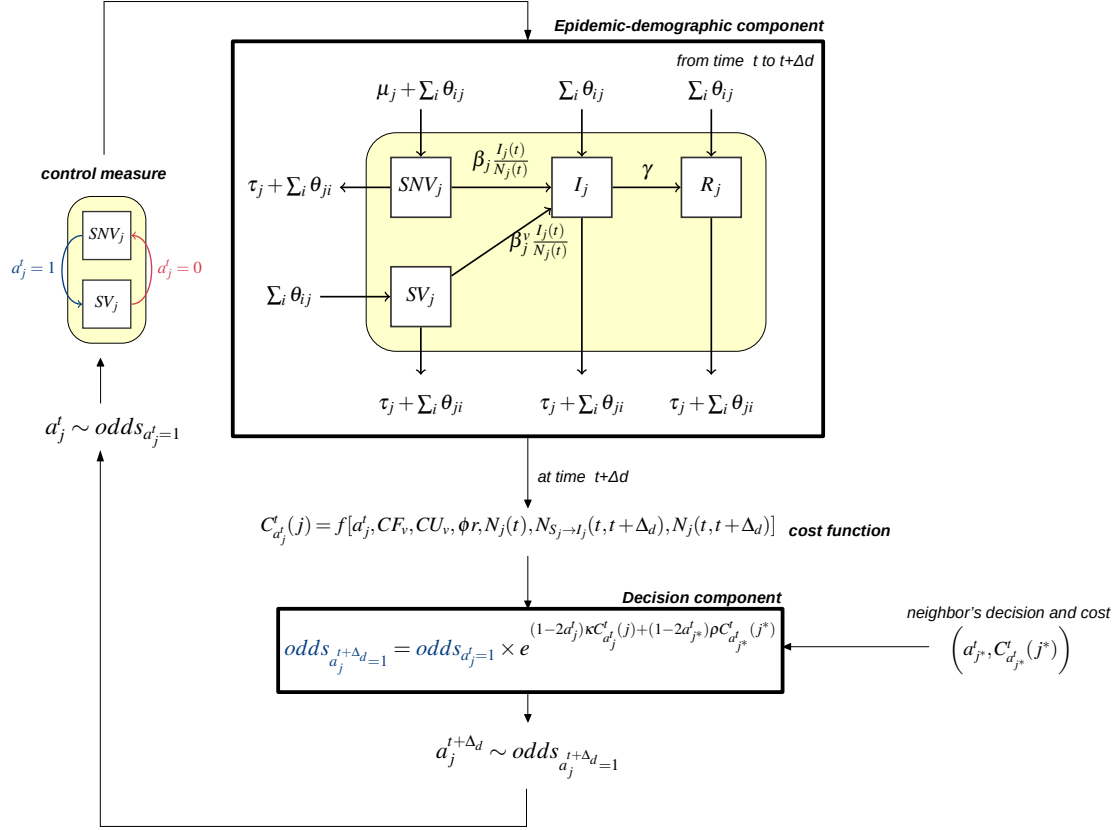


Figure 3.2: Representation of the integrative epidemic-decision dynamical model for a herd j , accounting for vaccinating decisions with a protective effect ($\beta_j^v < \beta_j$).

The flow between the non vaccinated compartment, $SNV_j(t)$, and the vaccinated one, $SV_j(t)$, is deterministic once the decision is taken. Indeed, if $a_j^t = 1$, all susceptible animals in herd j will enter the $SV_j(t)$ compartment for the next decision period. If $a_j^t = 0$, they will be in $SNV_j(t)$. If decided, vaccination is then applied only once per decision time. Indeed, in livestock diseases, as opposed to what happens for human diseases, it is not customary that farmers vaccinate newborns or the animals they buy after they have already vaccinated the herd, since each vaccination would then imply a cost for a veterinary visit. Furthermore, if the herd is vaccinated, farmers generally rely on herd immunity to indirectly protect susceptible animals in the herd.

3.3 Setting for simulations and sensitivity analyses

In this section we describe the methodology used for the simulation and analysis of the model. The simulation code is available at github.com/CristanchoLina/IntegrativeEpiDecisionModel.

3.3.1 Fixed simulation setting

For our simulation study, the population structure is set close to the one observed in real data obtained from the FCID. Furthermore, demographic parameters are fixed close to real-life values. In particular, the simulated trade network is scale-free, as the one observed in the real-life animal movements, then the in-degree and out-degree distributions follow a power law. That is, the majority of herds only buy (sell) animals to a few other herds, and very few herds buy

(sell) to many different herds, which are known as hubs. We simulate this network through the configuration model, using degree sequences generated from a power law.

The trade network is therefore simulated as a directed, weighted, scale-free network using iGraph python library calibrated to have herd's sizes, in-degree, out-degree, etc. distributions similar to the ones of Finistère's, a French region whose number of herds is close to 5000 according to the FCID. Movements of animals coming from outside this metapopulation are neglected. We specifically search to approach the distributions for 3 years (2013-2016). The followed steps are:

1. Simulate the in-degree sequence from a power-law distribution of parameter 2, chosen so as to be close to the estimated values in Dutta et al. 2014 for the parameter of the in-degree and out-degree annual distributions, 1.8 and 1.5 respectively.

That is, the fraction $P(k)$ of herds in the network having k ingoing edges for a large k_{in} , follows: $P(k_{in}) \sim k_{in}^{-\psi}$ with $\psi = 2$. We take the in-degree sequence also as the out-degree sequence, in order to build a directed non-weighted network with the given in and out degree sequences using the *Degree_Sequence* function of the *iGraph* python's module. This function uses the configuration model for generating random directed or undirected networks from given degree sequences. We then simplify the network so that it does not contain loops or multiple edges, while making sure that every node has at least an ingoing edge, and an outgoing edge.

2. Simulate the initial herd sizes from a Gamma distribution of shape parameter 9 and scale 12, in order to roughly obtain the shape of the herd size distribution in the FCID. That is $N_j(0) \sim \Gamma(9, 12)$ for $j = 1, \dots, L$.
3. Simulate the daily out-rates $\vec{\theta}_j := \sum_{j^* \in |Ch(j)|} \theta_{jj^*}$ as a random sample (with replacement) of size J , from the interval $[0.0006, 1[$, where the probabilities of selection are given by a power-like law of parameter 2. This roughly reproduces the out-strength distribution in the FCID, that is, the number of animals bought and sold over the 3 years by farm.
4. Assign the initial herd sizes and the daily out-rates proportionally to the out-degrees. This allows to reproduce the rather high correlation between the out-strength (the total number of animals sold by each herd over three years) and the out-degree, that is observed in the data.
5. Assign the daily trade-rates going from a herd j to its buyers $j_1^*, \dots, j_{|Ch(j)|}^*$ inversely proportional to the in-degree of each buyer j^* , and directly proportional to the length of the shortest path going from j^* to j . The objective is first to prevent that a herd with a lot of potential sellers receives too many animals, as herds with few sellers could end up not receiving any, and to avoid generating loops of two herds that only sell and buy to each other.

Regarding demographic parameters, we consider fixed theoretical birth and death daily rates across herds. That is $\mu_j = \mu$ and $\tau = \tau_j; \forall j = 1, \dots, J$. For the death rate the value used in all simulations is $\tau = 0.0009$. This value is chosen for animals to have approximately a mean expectancy life time of three years. The birth rate is chosen the double of the death rate, that is $\mu = 0.0018$ so that on average an animal has a calf every 1.5 years. We remark that in our simulations birth rates actually vary across herds, since we consider a soft constraint on the maximal capacity of each herd defined as $K_j = 1.5 \times N_j(0)$ for each herd j . We note as $\mathbb{1}_x$ the indicator function (i.e. it equals 1 if condition x is verified, 0 otherwise). The births and trades given by the multinomial draws (equations 3.7, 3.12 - 3.14) are modified as following:

- The births in herd j are given by $\min(BS_j(t), K_j - N_j(t)) \mathbb{1}_{K_j - N_j(t) > 0}$.
- Susceptible, infected and recovered animals going from a herd i to a herd j are equal to:
 $\min(S_{ij}, K_j - N_j(t)) \mathbb{1}_{K_j - N_j(t) > 0}$, $\min(I_{ij}, K_j - N_j(t)) \mathbb{1}_{K_j - N_j(t) > 0}$,
 $\min(R_{ij}, K_j - N_j(t)) \mathbb{1}_{K_j - N_j(t) > 0}$, respectively.

The comparison between the number of animals in each epidemic state (and the number of births) to $K_j - N_j(t)$, or simultaneously importing from different herds, can give rise to situations where the maximal capacity is exceeded at some time instants. Figure 3.3 shows however that only a minimal proportion of herds lightly exceeds the constraint on $K_j/N_j(0)$ so we allow for this excess in these herds. This is consistent over different simulations.

In figure 3.3 we present population dynamics regarding herd sizes evolution over the simulation for a single run. Figures 3.3a and 3.3b correspond to the initial and final herd size distributions of our synthetic data. As for figure 3.3c, it follows each herd size over time, where the color is given by the ratio between final and initial herd size. We see that the shape of the herd size distribution is rather stable over time. We remark that most herds quickly increase their size up to their maximal capacity. Indeed, one can have $N_j(t)$ higher than K_j , if j receives animals from more than one herd at the same time, which is why the ratio between the final and initial herd size is for some herds higher than 1.5. This dynamics is qualitatively stable over different runs.

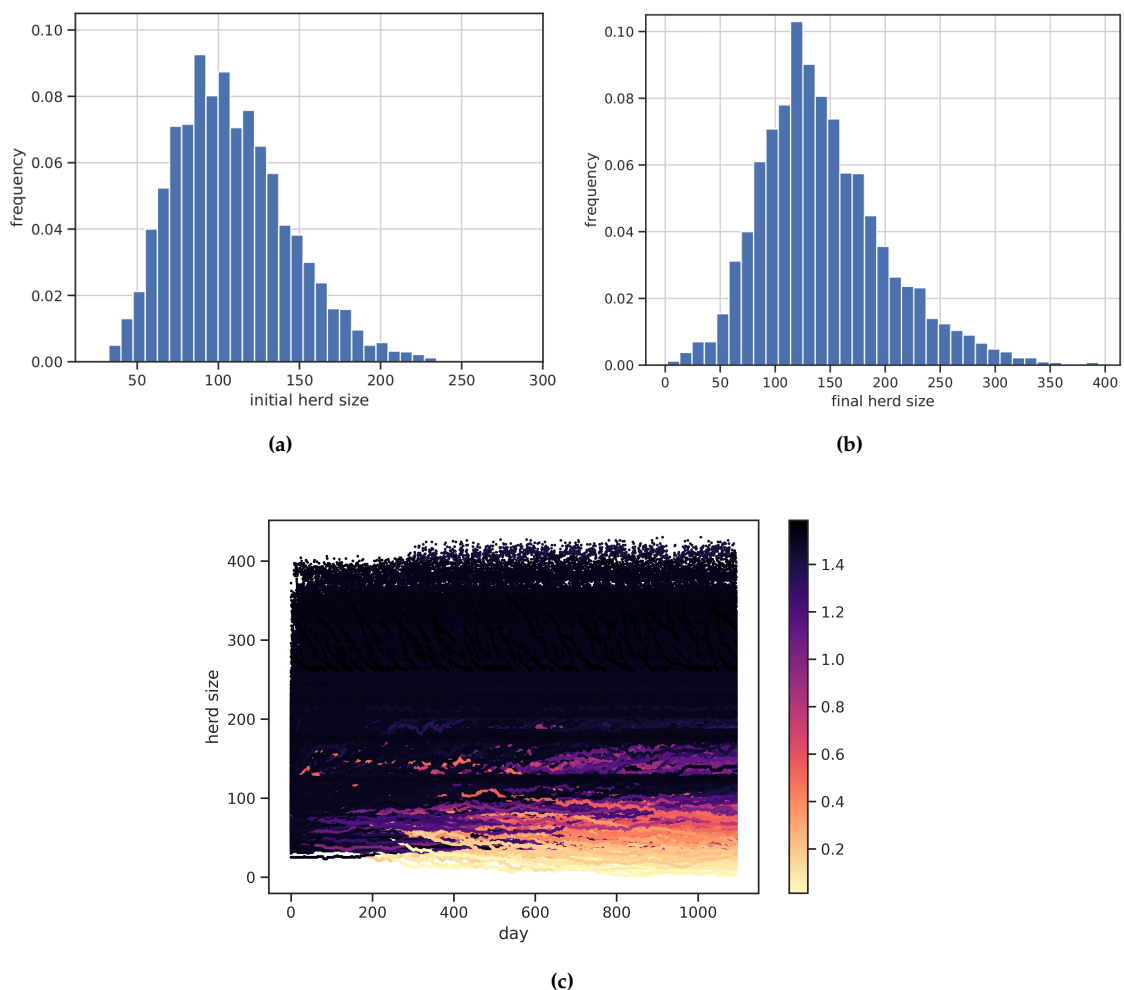


Figure 3.3: Population dynamics over 3 years. Results for a single run. (a) Initial herd size distribution, (b) Final herd size distribution (after 3 years), and (c) Herd size evolution over time. In (c) color is given by the ratio between final and initial herd size.

In figure 3.4 we present the herd size distribution in Finistère for years 2013 and 2015 according to the FCID. We roughly reproduce the range and shape of the herd size distribution, but not the high proportion of herds with very few animals (less than 5). Yet, we remark that the category

of these small herds is marginal, since it represents less than 0.1% of the total metapopulation (in number of animals).

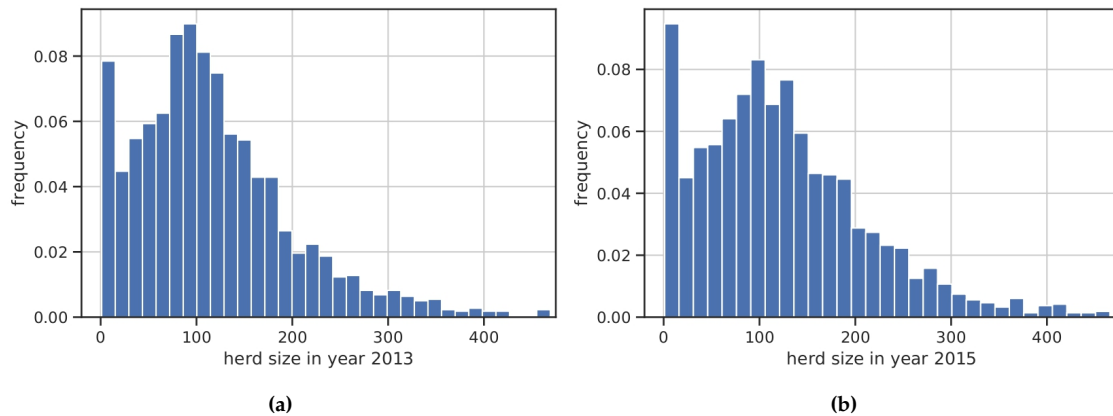


Figure 3.4: Herd size distribution in Finistère from real data (FCID) for years (a) 2013 and (b) 2015.

Figure 3.5a presents on a log scale the in(out)- degree distribution of the generated network using the previously described procedure. Figure 3.5b and 3.5c contain the respective representation of the in-strength and out-strength distribution resulting from a single run of the model (equations 3.1 - 3.22) for 3 years, again using a log scale. These two network indicators are respectively defined as the number of animals bought and sold by each farm over the 3 years simulated. The generated distributions are qualitatively stable over different runs.

Finally, figure 3.6 presents the respective distributions in real data from the FCID, for Finistère, corresponding to the 2013-2016 period. The degree distribution of the simulated network closely resembles the in-degree of the real data, the out-degree distribution being moderately different. As for the animals sold and bought over a run of our model, we roughly reproduce the shape of their distributions in the real data. In particular, the shape of the simulated out-strength distribution is quite similar to the one observed for the real network.

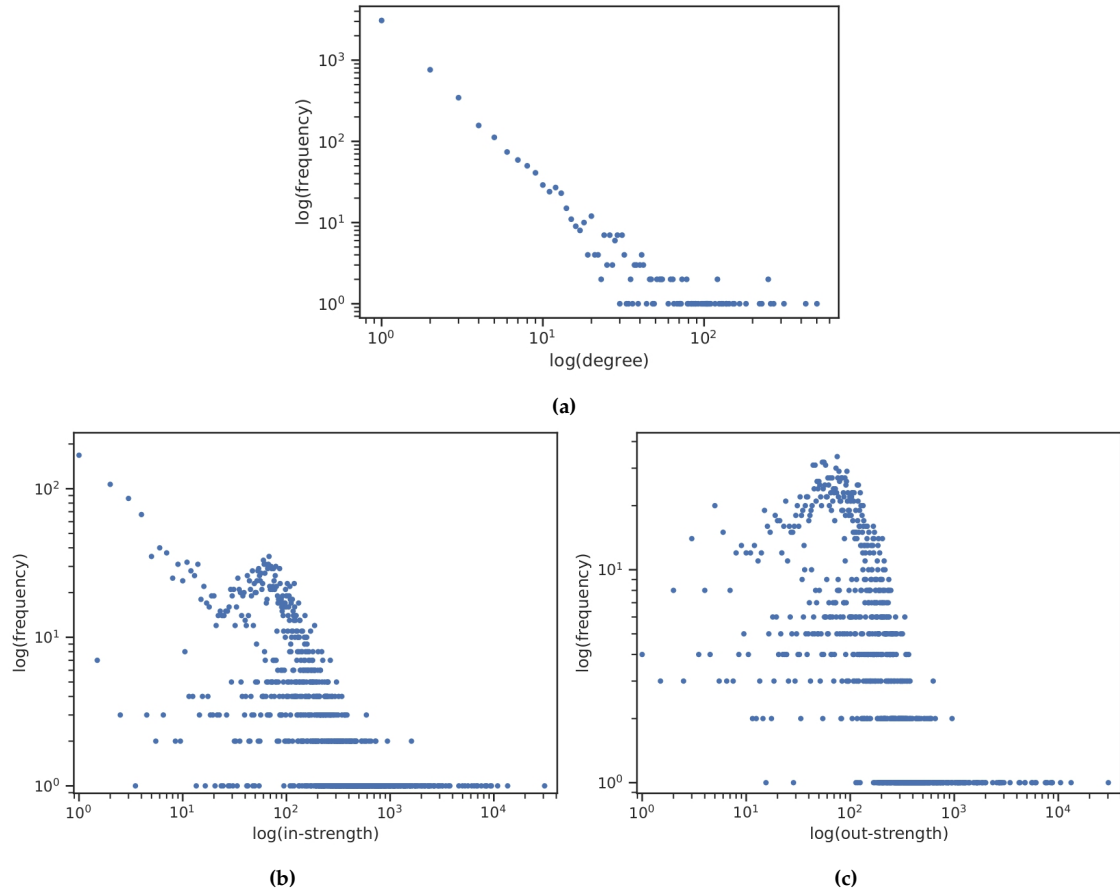


Figure 3.5: Distribution of characteristics of the simulated network in log-log scale. (a) In-degree distribution of the simulated trade network as described in the Trade Network section of this document. The out-degree sequence is the same as the in-degree sequence in our simulation. (b) In-strength distribution, and (c) out-strength distribution. These correspond respectively to the number of animals bought and sold over the 3 years for a single run of the model (equations 3.1 - 3.22).

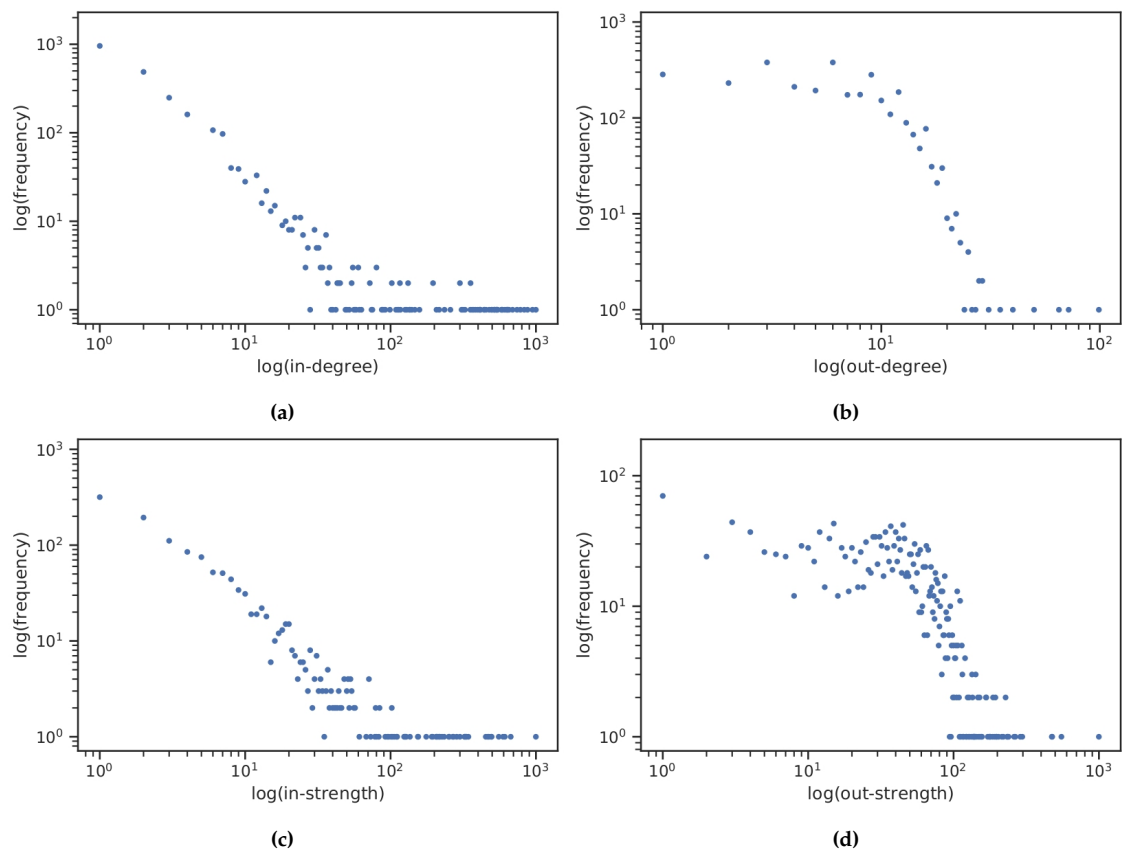


Figure 3.6: Distribution of characteristics of the real network (data from the FCID, for Finistère 2013-2016) in log-log scale. (a) In-degree distribution, (b) out-degree distribution, (c) in-strength distribution, and (d) out-strength distribution. The in-strength and out-strength correspond respectively to the number of animals bought and sold by each farm over the 3 years.

3.3.2 Parameters of the integrative model

The values of the epidemic, economic and decision related parameters used in the simulation study are given in Table 3.1.

	Parameter	Definition	Standard value	Values tested in the sensitivity analysis
Epidemic	β/γ	transmission rate per herd \times average duration of infection	2.	[1.1 , 2.07 , 3.05 , 4.02 , 5.]
	$1/\gamma$	average duration of infection (in days)	90.	[10. , 32.5 , 55. , 77.5 , 100.]
	p_{herds}^0	initial proportion of infected herds	0.10	[0.01 , 0.22 , 0.43 , 0.64 , 0.85]
	p_{anim}^0	initial proportion of infected animals in infected herds	0.15	[0.01 , 0.25 , 0.50 , 0.75 , 1.]
Economic	r	monetary value of a healthy animal (in euros)	2000.	[1000. , 1500. , 2000. , 2500. , 3000.]
	ϕ	reduction in the monetary value of an animal if it gets infected	0.8	[0.01 , 0.25 , 0.50 , 0.75 , 1.]
	CU_v	unitary cost of the vaccine per animal (in euros)	5.	[1. , 4.5 , 8. , 11.5 , 15.]
	CF_v	fixed cost of applying vaccination per herd (in euros)	50.	[1. , 25.75 , 50.5 , 75.25 , 100.]
Decision related	e_v	protection efficacy of the vaccine on susceptible animals	1.	[0.01 , 0.25 , 0.50 , 0.75 , 1.]
	Δ_d	duration of the decision (time between two consecutive decisions). It also determines the time of the first decision, and is equal to the duration efficacy of the vaccine (in days)	180.	[30. , 114 , 198 , 281. , 365.]
	p_v^{init}	farmers' initial probability of vaccinating	0.01	[0.01 , 0.25 , 0.5 , 0.74 , 0.99]
	κ	farmers' sensitivity to their own observed cost	0.5. or 12.5	[0.5 , 3.5 , 6.5 , 9.5 , 12.5]
	ρ/κ	farmers' sensitivity to a neighbour's cost / farmers' sensitivity to his/her own observed cost	0.5	[0. , 0.25 , 0.5 , 0.75 , 1.]

Table 3.1: Parameters of the integrative model: description, standard values and values tested in the full sensitivity analysis.

We remark that these are set close to realistic values, having in mind a standard SIR endemic disease. In particular, we consider the same transmission rate across herds, so $\beta_j = \beta; \forall j = 1, \dots, J$. As for the duration of the decision it is chosen to be 180 days, which is a reasonable assumption in practice. The values for κ and ρ are chosen so as to have two potentially contrasted decision scenarios.

3.3.3 Sensitivity analysis experiments

Sensitivity analysis is useful to study how much the variation in each parameter of the model contributes to the variation of the model outputs (Saltelli et al. 2008). In our sensitivity analyses we consider 13 input parameters in total. Other parameters, in particular the demographic ones, are fixed as specified earlier. We consider eight outputs corresponding to the three model components: epidemic, economic and decision related, and one additional output that combines epidemic and decision related elements. These outputs are defined in Table 3.2.

The values of the inputs used in the sensitivity analyses are chosen using Fractional Factorial design (Gunst et al. 2009) with 5 equally spaced levels, which results in 625 combinations of parameters. To obtain this design we use the R package PLANOR (Kobilinsky et al. 2020). Since the model is stochastic, we run 50 simulations for each combination, and we consider the mean and the variance of each output over runs. Table 3.1 contains the values considered for each input in the full sensitivity analysis. Since we use a IV-resolution design, we are able to estimate main effects unconfounded by two-factor interactions, while limiting the number of runs required for the analysis. With this design, we can also estimate two-factor interaction effects, even if these may be confounded, i.e. can not be estimated independently to each other (Montgomery 2017). We study the outputs individually, by groups regarding the nature of the outputs,

Group	Output	Definition
Epidemic	p_{herds}^T (final inter-herd prevalence rate)	final proportion of infected herds = $\frac{1}{J} \sum_{j=1}^J \mathbb{1}_{I_j(T)>0}$
	\bar{p}_{anim}^T (final intra-herd mean prevalence rate)	mean over final infected herds of the final proportion of infected animals = $\left(\sum_{j=1}^J \frac{I_j(T)}{N_j(T)} \right) / \left(\sum_{j=1}^J \mathbb{1}_{I_j(T)>0} \right)$
	$p_{herds}^{[0,T]}$ (inter-herd cumulative incidence rate)	cumulative proportion of newly infected herds (i.e. herds with new infections) = $\frac{1}{J} \sum_{j=1}^J \mathbb{1}_{\sum_{t=0}^T N_{S_j \rightarrow I_j}(t) > 0}$
	$\bar{p}_{anim}^{[0,T]}$ (mean cumulative intra-herd incidence rate)	mean cumulative proportion of new infected animals over susceptible animals, for newly infected herds = $\left(\sum_{j=1}^J \frac{1}{T} \sum_{t=0}^T \frac{N_{S_j \rightarrow I_j}(t)}{S_j(t)} \right) / \left(\sum_{j=1}^J \mathbb{1}_{\sum_{t=0}^T N_{S_j \rightarrow I_j}(t) > 0} \right)$
	$\bar{I}_{anim}^{[0,T]}$ (mean cumulative intra-herd incidence)	mean cumulative number of new infected animals for new infected herds = $\left(\sum_{j=1}^J \sum_{t=0}^T N_{S_j \rightarrow I_j}(t) \right) / \left(\sum_{j=1}^J \mathbb{1}_{\sum_{t=0}^T N_{S_j \rightarrow I_j}(t) > 0} \right)$
Economic	$C^{[0,T]}$ (total economic cost of the disease)	sum of the non standardized cumulative disease-related costs (costs of vaccination and costs of new infections): $\sum_{j=1}^J \left[\sum_{n=0}^{\lfloor T/\Delta_d \rfloor} C_{a_j}^{n\Delta_d}(j) \Delta_d \bar{N}_j(t, t + \Delta_d) \right]$. Counts costs even before the first decision and after the last one.
Decision related	$\bar{p}_v^{[0,T]}$ (mean vaccination proportion)	mean proportion of herds that vaccinate over the different decision times except the first one = $\left(\sum_{n=2}^{\lfloor T/\Delta_d \rfloor} \frac{1}{J} \sum_{j=1}^J \mathbb{1}_{a_j^{n\Delta_d-1}} \right) / (\lfloor T/\Delta_d \rfloor - 1)$
	aggregated vaccination patterns	vector consisting in three proportions: of herds that never vaccinate, of herds that vaccinate at least once and at most half of the time, and of herds that vaccinate more than half of the time but not always. Without taking the first decision into account.
Epidemic-decision related	mean cumulative intra-herd incidence rate by aggregated vaccination pattern	vector of the mean cumulative intra-herd incidence rate (see output $\bar{p}_{anim}^{[0,T]}$) of herds grouped by the aggregated vaccination pattern: herds that never vaccinate, herds that vaccinate at least once and at most half of the time, and herds that vaccinate more than half of the time but not always. Without taking the first decision into account.

Table 3.2: Description of the outputs of the sensitivity analyses.

and by considering all outputs together. For the multivariate analyses, we use PCA (Principal Component Analysis) to reduce the dimension of the output space, before using Analysis of variance (ANOVA) for the computation of Global Sensitivity Indices (GSI), which are weighted means of the sensitivity indices over the retained dimensions in the PCA, as described in Lamboni et al. 2011. More precisely, ANOVA is particularly suited for analyzing the outcome of a factorial design (Van Schepdael et al. 2016). For all the sensitivity analyses we use the R package multisensi (Bidot et al. 2018). In the PCAs the means are centered and scaled, and the dimension is selected as the smallest value that keeps at least 95% of the total variability. Among the many experiments, we retain the results of the three following ones:

- (i) First experiment: all 13 inputs. The means and variances of all outputs: by group, and all outputs simultaneously.
- (ii) Second experiment: all inputs except the four epidemic parameters (fixed to their standard values in Table 3.1). Means and variances of all outputs.
- (iii) Third experiment: all inputs except the two epidemic parameters (p_{herds}^0 and β/γ) and the two decision related parameters (Δ_d and p_v^{init}), fixed to their standard values in Table 3.1. Means and the variances of decision related outputs.

3.4 Results of the simulation study and sensitivity analyses

3.4.1 Model predictions for different decision scenarios

Results regarding the inter-herd prevalence, and the intra-herd prevalence for infected herds are provided for four different scenarios (figure 3.7): no farmer ever vaccinates, *never* scenario; every farmer vaccinates at every decision-time, *always* scenario; farmers vaccinate following the proposed decision-making mechanism (Algorithm 2) using $\kappa = 0.5$, *neigh-expw(0.5)* scenario; and the same mechanism using $\kappa = 12.5$, *neigh-expw(12.5)* scenario.

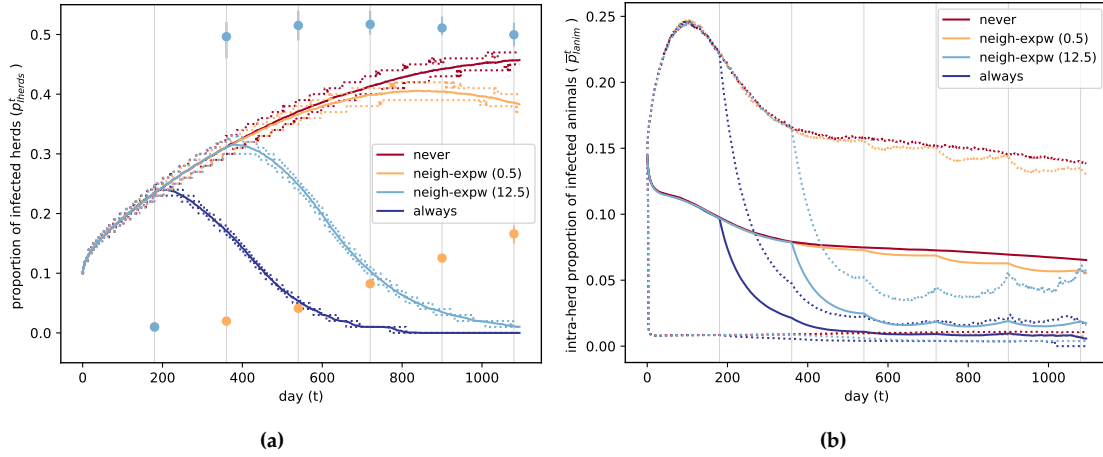


Figure 3.7: Temporal dynamics of the epidemic spread for each vaccination scenario over 50 runs. Each decision instant is represented by a vertical grey line. (a) Inter-herd prevalence. Mean over runs (solid lines), 10th and 90th percentiles over runs (dotted lines). Mean proportion of herds that vaccinate at each decision-time in each *neigh-expw* scenario (light blue and orange dots), and its variation over runs (from the 10th to the 90th percentile in light blue and orange vertical lines). (b) Intra-herd prevalence for infected herds. Mean over runs of the means over infected herds (solid lines), 10th percentile over runs of the 10th percentiles over infected herds, and 90th percentile over runs of the 90th percentiles over infected herds (dotted lines).

As expected, the worst and best case scenarios are the scenario where farmers never vaccinate, and the one where they all vaccinate at each decision time. We remark that the vaccination gain particularly affects inter-herd prevalence (figure 3.7a), but is still observable for intra-herd prevalence (figure 3.7b). In the intermediate scenarios, farmers' sensitivity to observed costs determines the changes in the proportion of herds that vaccinate over time, and therefore in the control of the pathogen spread. Indeed, in the scenario with higher farmers' sensitivities to costs (*neigh-expw(12.5)* scenario), the proportion of farmers that vaccinate quickly increases after the first decision, generating a mean inter-herd and intra-herd prevalence dynamics rather close to the best case scenario. On the contrary, the scenario with smaller farmers' sensitivities to costs (*neigh-expw(0.5)* scenario) exhibits a slow increase in the proportion of herds that vaccinate, which gives rise to a prevalence behaviour close to the one observed for the worst case scenario, even if around two years it starts to decline.

Additionally, figure 3.8 presents the temporal dynamics of the vaccination decisions of the two intermediate scenarios (for a single run as an example). In the *neigh-expw(0.5)* scenario most herds never vaccinate (67%). They are followed by herds that only vaccinate at the last decision time, which are in turn followed by those that only vaccinate at the next to last decision time, etc. Only 28 out of the 64 possible patterns (over 6 decision times) are observed in this scenario. On the other hand, in the *neigh-expw(12.5)* scenario the most frequent behaviour (39%) is to not vaccinate at the initial decision and to always vaccinate afterwards. However, this vaccination pattern is closely followed by the one where herds never vaccinate (33%). We also observe a higher variety of behaviours than in the *neigh-expw(0.5)* scenario, 44 out of the 64 possible patterns,

which translates into less frequent patterns. Nevertheless, some of them stand out: the one where herds vaccinate from the second decision time, the one where herds only vaccinate at the second decision time, and the one where herds vaccinate from the third decision time.

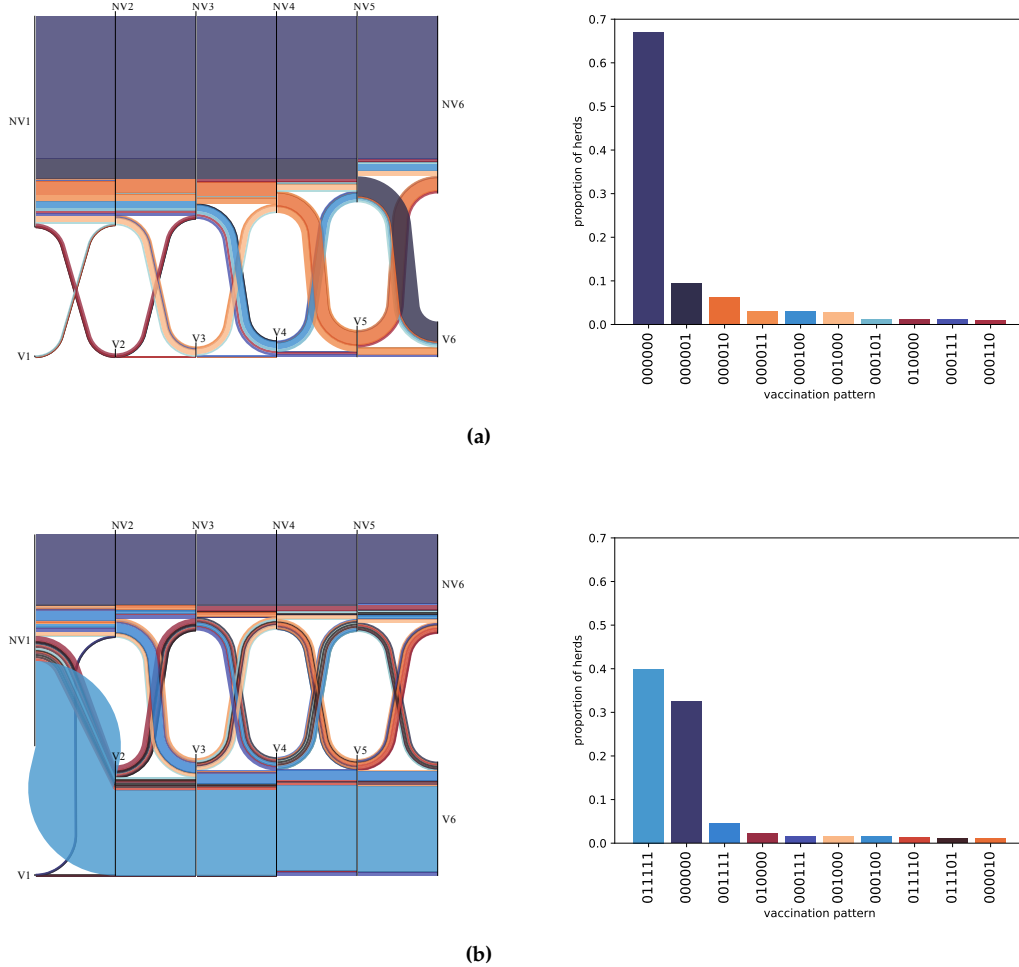


Figure 3.8: Temporal dynamics of the vaccination decisions using the decision mechanism defined in Algorithm 2 with $\kappa = 0.5$ (a), and $\kappa = 12.5$ (b). Results for one run. NV and 0 stand for not vaccinating, while V and 1 for vaccinating. Each color represents a different vaccination pattern, defined by the sequence of vaccination decisions at each of the six decision times. So the pattern 001111 (or equivalently [NV1, NV2, V3, V4, V5, V6]) concerns herds that do not vaccinate at the two first decision times, and always vaccinate afterwards. In the left plots, each vertical black line represents a decision time, and the width of the flows between decisions is proportional to the frequency of the pattern. In the right plots, the histogram of the patterns with a frequency $\geq 1\%$ is plotted. Hence, in (a), 67% of herds never vaccinate (pattern 000000). In (b), 39% of farms always vaccinate except in the first instant (pattern 011111), and 33% never vaccinate (pattern 000000).

The scenarios concerning vaccination exhibit some peaks in the intra-herd prevalence roughly at each decision time (figure 3.7b). This behaviour is firstly explained by the fact that we consider this prevalence only for infected herds at each time, so the concerned herds are not the same over the whole trajectory. Furthermore, since we consider a perfect vaccine, when a herd is vaccinated all its susceptible animals are completely protected, so that the number of animals that can actually get infected drops instantaneously to zero, until there are births or imports of non-vaccinated susceptible animals. Figures 3.9 and 3.10 present an exploration of this behaviour. In particular, these figures present the epidemic evolution of herds when grouped by their vaccination pattern (i.e. the sequence of individual decisions with respect to vaccination), as well as their main

characteristics concerning population and trade dynamics. We focus on the most frequent patterns for the two decision scenarios considered: the *neigh-expw(0.5)* scenario, where farmers have a low sensitivity to the observed costs ($\kappa = 0.5$), and the *neigh-expw(12.5)* scenario, where their sensitivity to the observed costs is higher ($\kappa = 12.5$). We recall that for both scenarios $\rho/\kappa = 0.5$.

In the *neigh-expw(0.5)* scenario (figure 3.9), the most frequent vaccination pattern (67%) is to not vaccinate at any of the six decision times: 000000. Followed by 000001, 10% of herds; 000010, 6% of herds; 000011, 3% of herds; 000100, 3% of herds; and 001000, 2.5% of herds. We observe in this figure that for each pattern, the change from not vaccinating to vaccinating (0 to 1) follows a period of strong increase in the number of infected animals in infected herds with the pattern, and allows for an abrupt reduction in the number of non vaccinated susceptible animals. This is a consequence of the fact that decision implies the vaccination of all animals in the herd. This reduction is therefore accompanied by a decline in the number of infected animals, which is easily explained by the highly effective vaccine ($e_v = 1$), causing the peak to be exactly at the time of the decision switch from 0 to 1. Furthermore, the number of non vaccinated susceptible stabilizes after the decision, which can be attributed to birth and imports of animals in this epidemic state.

In the *neigh-expw(12.5)* scenario (figure 3.10), the most frequent vaccination pattern (39%) is to vaccinate in all of the six decision times except the first one: 011111, followed by never vaccinating, 000000, 33% of herds; 001111, 4% of herds; and 010000, 2% of herds. Like for the first scenario, we see in this figure that in all patterns the change from not vaccinating to vaccinating follows a period of increase on the number of infected animals of the herds, but here the increase is smaller given the high sensitivity of farmers to costs. At the same time, we also have a sharp reduction in the number of non vaccinated susceptible animals. This reduction is again accompanied by a decline in the number of infected animals, causing the peak to be exactly at the time of the decision changes from 0 to 1. The peaks are more pronounced than in the first scenario, given the stronger reduction in the number of infected herds. Indeed, since we look only at infected animals among infected herds, and since in the *neigh-expw(12.5)* scenario there is a higher decline in the number of infected herds, over time the herds that are concerned are only those that are still infected, which corresponds to herds with not frequent vaccinating patterns.

Figure 3.11 shows that in both scenarios, herds that never vaccinate tend to be less connected than the others (smaller out-degree, smaller in-strength), contrary to herds that vaccinate more, for example, herds that only not vaccinate at the first decision time, in the *neigh-expw(12.5)* scenario. The in-degree is highly and positively correlated to the out-degree, and the out-strength is correlated with the in-strength, so their plots are not shown.

In order to analyse model predictions on the long term, we simulated each of the four scenarios over a trajectory of 9 years (figure 3.12). We observe that for the inter-herd prevalence (figure 3.12a), in the *never* scenario the pathogen is indeed endemic in the meta-population, reaching a stable level after 6 years at 38% percent of infected herds. The *always* and *neigh-expw(12.5)* scenario do not change their behaviour after the initial three years. As for the *neigh-expw(0.5)* scenario, even if in the first three years it is rather similar to the *never* scenario, in the following years the inter-herd prevalence declines to be close to zero at the end of the year 9. The intra-herd prevalence (figure 3.12b) does not really change in any of the scenarios with respect to the behaviour observed in the 3-year simulation. There is only an increased variation among infected herds, for the two intermediate scenarios, that is explained by the fact that infected herds change over time, and that in the long run the concerned herds are only those with not frequent vaccination patterns. The dynamics of the total number of infected animals (figure 3.12c) is an alternative quantity to study. Yet, as evidenced by the figure, it is highly correlated to the proportion of infected herds.

Concerning decisions, in figure 3.12a we observe that for the two intermediate scenarios, the proportion of herds that vaccinate over time follows the same behaviour as in the three year simulations, arriving to around 35% at nine years, for both scenarios. This indicates that using the model, a high proportion of farmers keeps vaccinating even if there are no more infected herds.

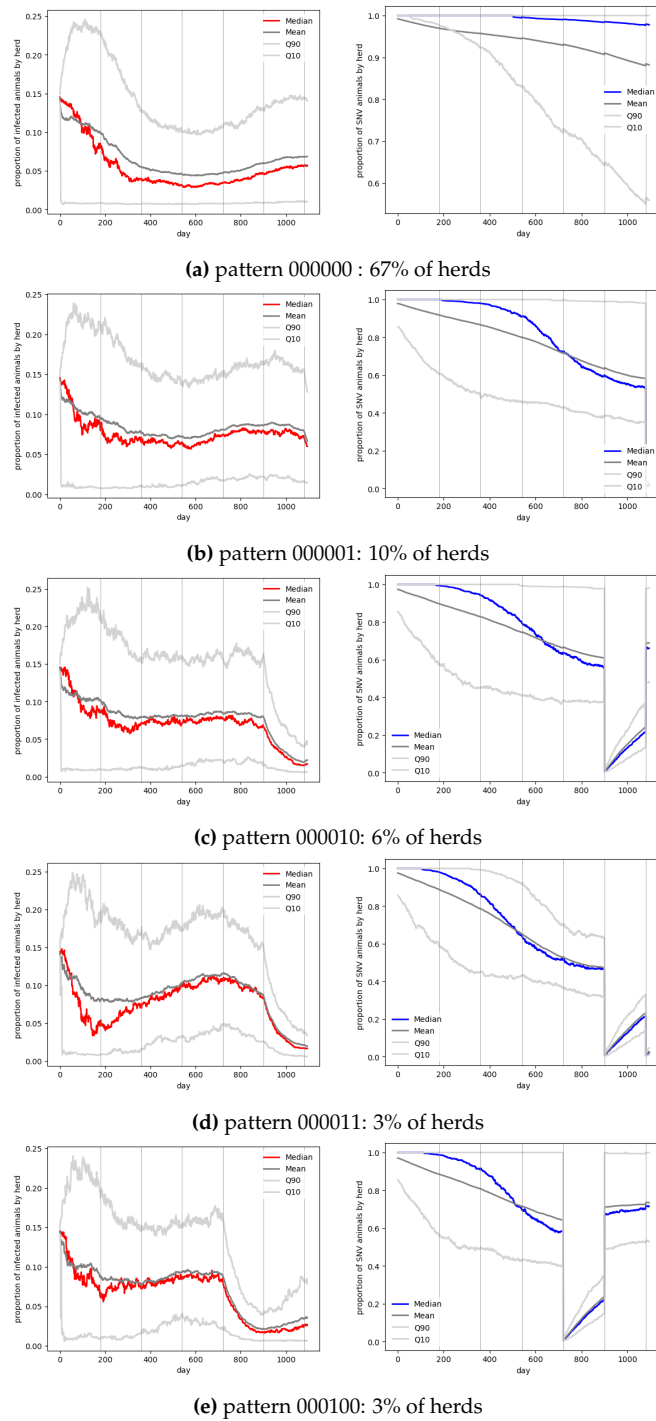


Figure 3.9: Epidemic spread for each vaccination pattern for the most frequent patterns in scenario *neigh-expw(0.5)*, based on a single run. The plots in the left show intra-herd prevalence among infected herds of the given pattern (mean, median, 10th and 90th percentiles). The plots in the right show the proportion of non vaccinated animals among infected herds of the given pattern (mean, median, 10th and 90th percentiles). Decisions instants are represented by the vertical grey lines. 0 stands for not vaccinating, while 1 stands for vaccinating.

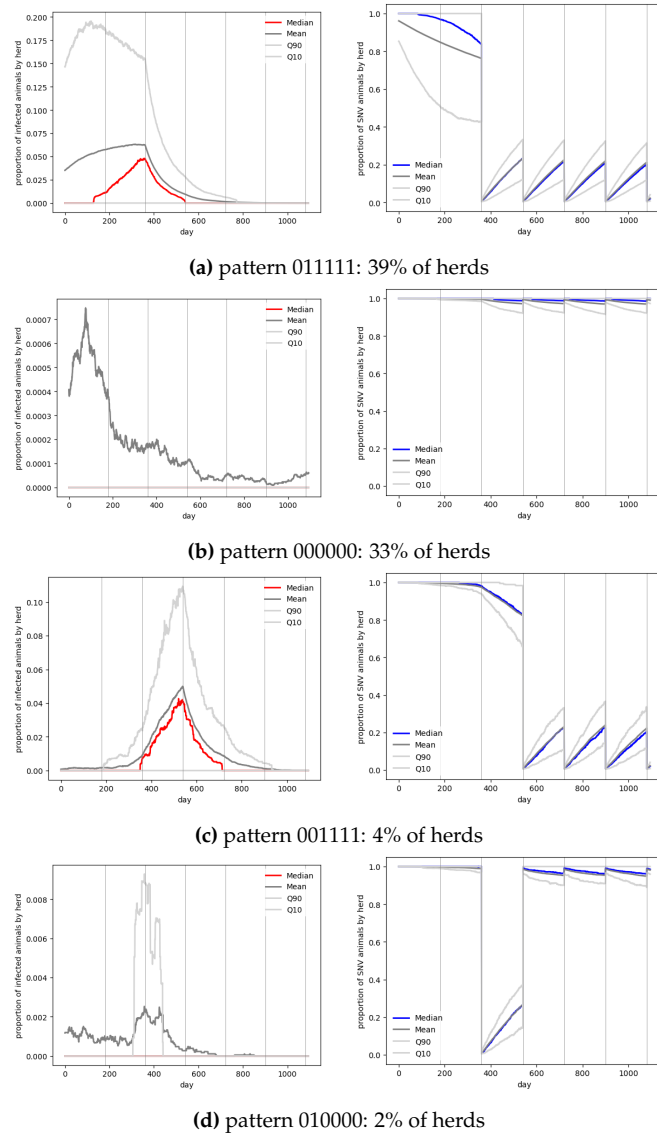


Figure 3.10: Epidemic spread for each vaccination pattern for the most frequent patterns in scenario *neigh-expw(12.5)*, based on a single run. The plots in the left show intra-herd prevalence among infected herds of the given pattern (mean, median, 10th and 90th percentiles). The plots in the right show the proportion of non vaccinated animals among infected herds of the given pattern (mean, median, 10th and 90th percentiles). Decisions instants are represented by the vertical grey lines. 0 stands for not vaccinating, while 1 stands for vaccinating.

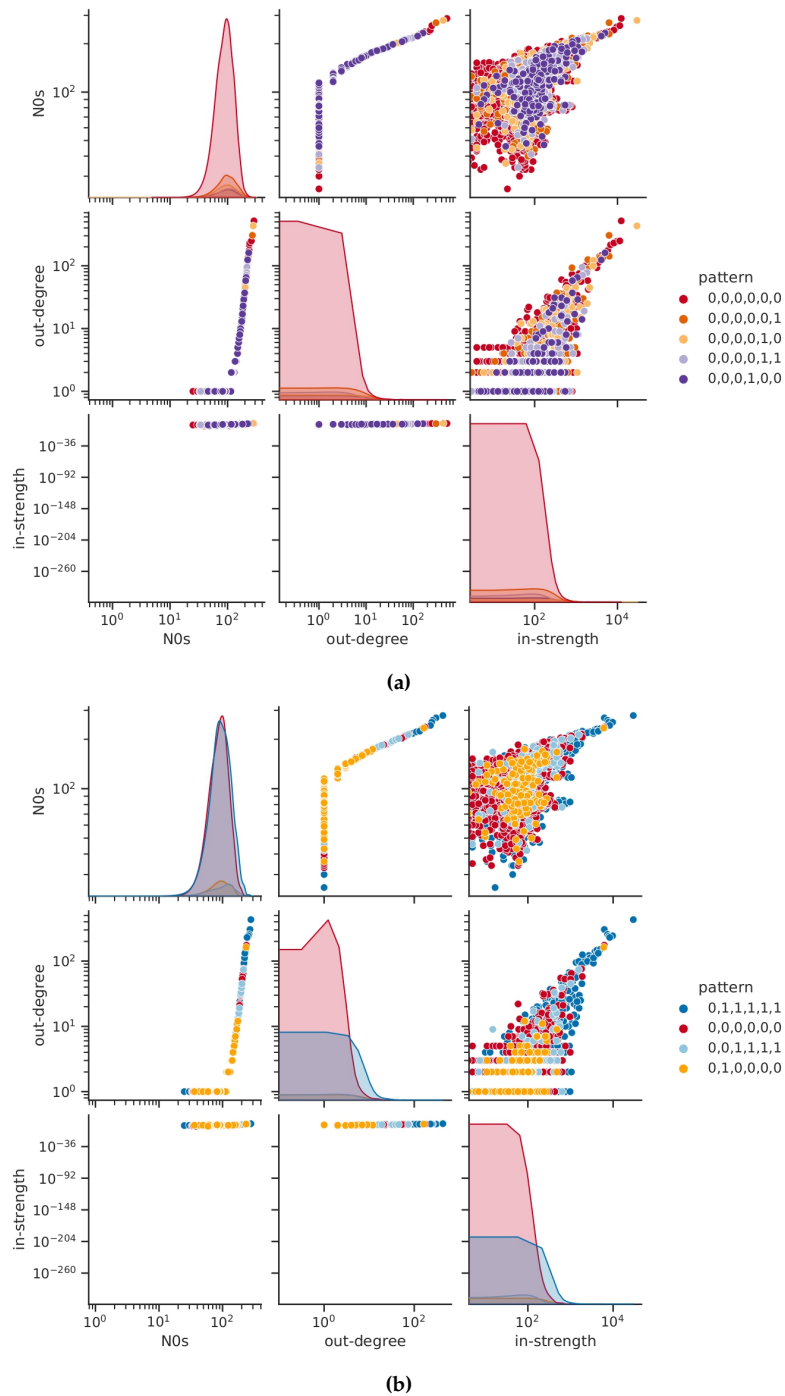


Figure 3.11: Exploration of network characteristics (initial herd size, out-degree and in-strength) of herds in the most frequent vaccination patterns in scenarios using the decision mechanism defined in Algorithm 2 with $\kappa = 0.5$ (a), and $\kappa = 12.5$ (b). 0 stands for not vaccinating, while 1 stands for vaccinating. Results for a single run.

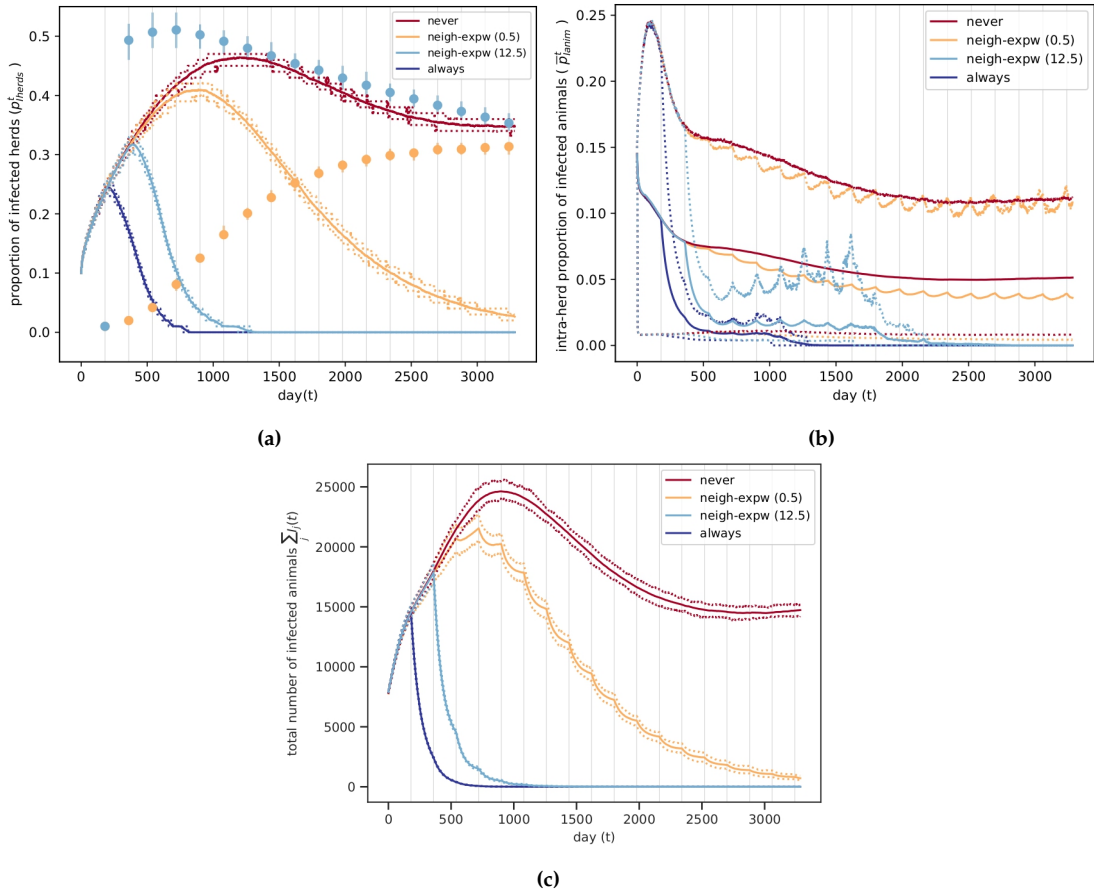


Figure 3.12: Model predictions over 9 years. Temporal dynamics of the epidemic spread for each vaccination scenario over 50 runs. Each decision instant is represented by a vertical grey line. (a) Inter-herd prevalence. Mean over runs (solid lines), 10th and 90th percentiles over runs (dotted lines). Mean proportion of herds that vaccinate at each decision-time in each *neigh-expw* scenario (light blue and orange dots), and its variation over runs (from the 10th to the 90th percentile in light blue and orange vertical lines). (b) Intra-herd prevalence for infected herds. Mean over runs of the means over infected herds (solid lines), 10th percentile over runs of the 10th percentiles over infected herds, and 90th percentile over runs of the 90th percentiles over infected herds (dotted lines). (c) Total number of infected animals in the metapopulation. Mean over runs (solid lines), 10th and 90th percentiles over runs (dotted lines). Inter-herd and intra-herd prevalence rates are defined in Table 3.2.

We now present the model predictions according to Algorithm 3, where each farmer considers the decisions and costs observed by all of his/her neighbors in the trade network at each decision time. Two numerical scenarios are considered based on this model: the *all-neigh-expw(0.5)* scenario, where $\kappa = 0.5$, and the *all-neigh-expw(12.5)* scenario, where $\kappa = 12.5$. Figures 3.13a and 3.13b show the model predictions over 9 years according to these experiments, as well as for the scenario where no farmer ever vaccinates, *never* scenario, and the scenario where every farmer vaccinates at every decision-time, *always* scenario. A focus on the epidemic and decision dynamics over 3 years can be found in figures 3.13c, 3.13d and 3.14.

We observe that results are not very different for both scenarios, *all-neigh-expw(0.5)* and *all-neigh-expw(12.5)*, with respect to the *neigh-expw(0.5)* and *neigh-expw(12.5)* scenarios (figures 3.8 and 3.12). In both scenarios there are slightly less vaccination patterns, and patterns that are more frequent, such as 011111 in the *all-neigh-expw(0.5)* scenario, or 011110 in the *all-neigh-expw(12.5)* scenario (figure 3.14). In the *all-neigh-expw(0.5)* scenario, the proportion of herds that vaccinate increases and stabilizes more rapidly to a smaller value (20% versus 30% using only one neighbor). The epidemic dynamics is therefore different (figure 3.13): the highest proportion of infected herds is 30% (versus 40% using only one neighbor), but afterwards the prevalence decreases less

rapidly. While in the *all-neigh-expw(12.5)* scenario the proportion of herds that vaccinate increases more rapidly at the second decision time, and then it continues to decrease. We remark that in this scenario the proportion of herds that never vaccinate is about 15% smaller, and the proportion of herds that always vaccinate from the second decision time is almost 25% higher than in the scenario considering only one neighbor. Yet, the prevalence of the disease decreases only slightly faster.

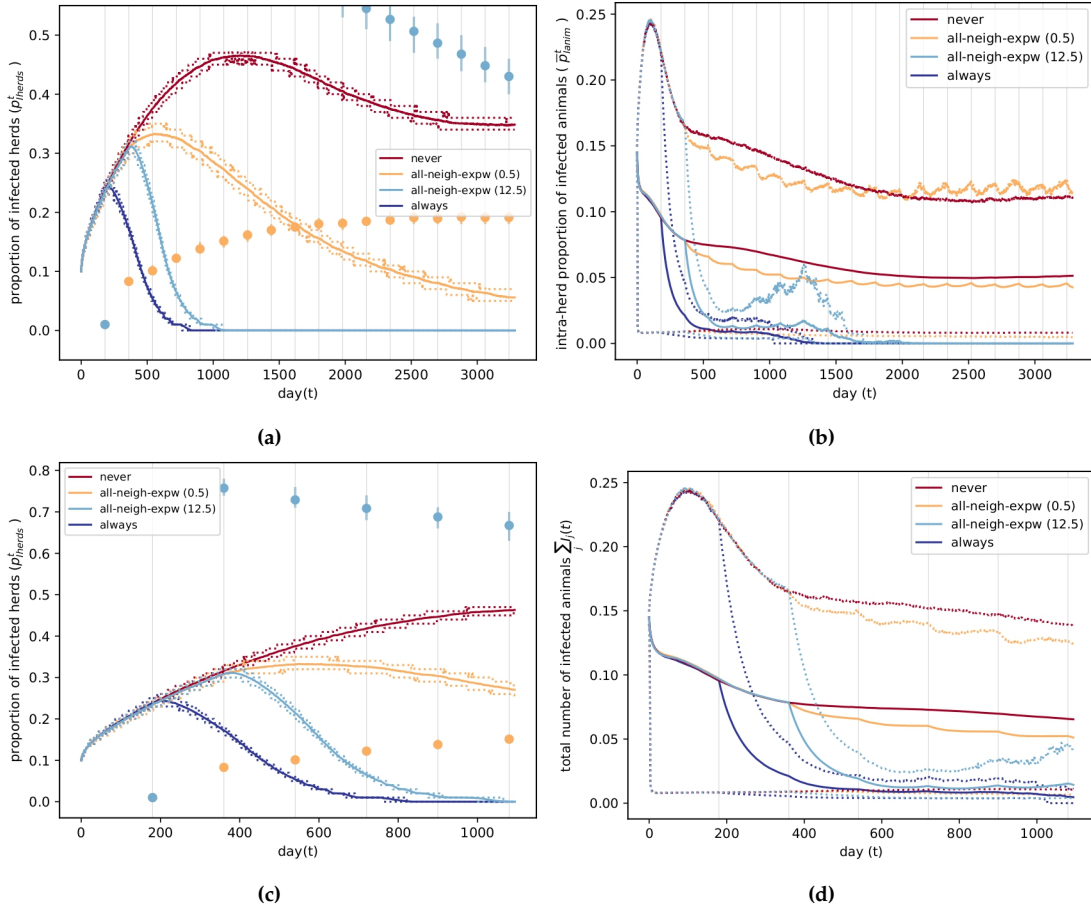


Figure 3.13: Model predictions considering information from all neighbors in the decision, as described in Algorithm 3. Temporal dynamics of the epidemic spread for each vaccination scenario over 50 runs. Each decision instant is represented by a vertical grey line. (a) Inter-herd prevalence over 9 years, and (c) over 3 years. Mean over runs (solid lines), 10th and 90th percentiles over runs (dotted lines). Mean proportion of herds that vaccinate at each decision-time in each *neigh-expw* scenario (light blue and orange dots), and its variation over runs (from the 10th to the 90th percentile in light blue and orange vertical lines). (b) Intra-herd prevalence for infected herds over 9 years, and (d) over 3 years. Mean over runs of the means over infected herds (solid lines), 10th percentile over runs of the 10th percentiles over infected herds, and 90th percentile over runs of the 90th percentiles over infected herds (dotted lines). Inter-herd and intra-herd prevalence rates are defined in Table 3.2.

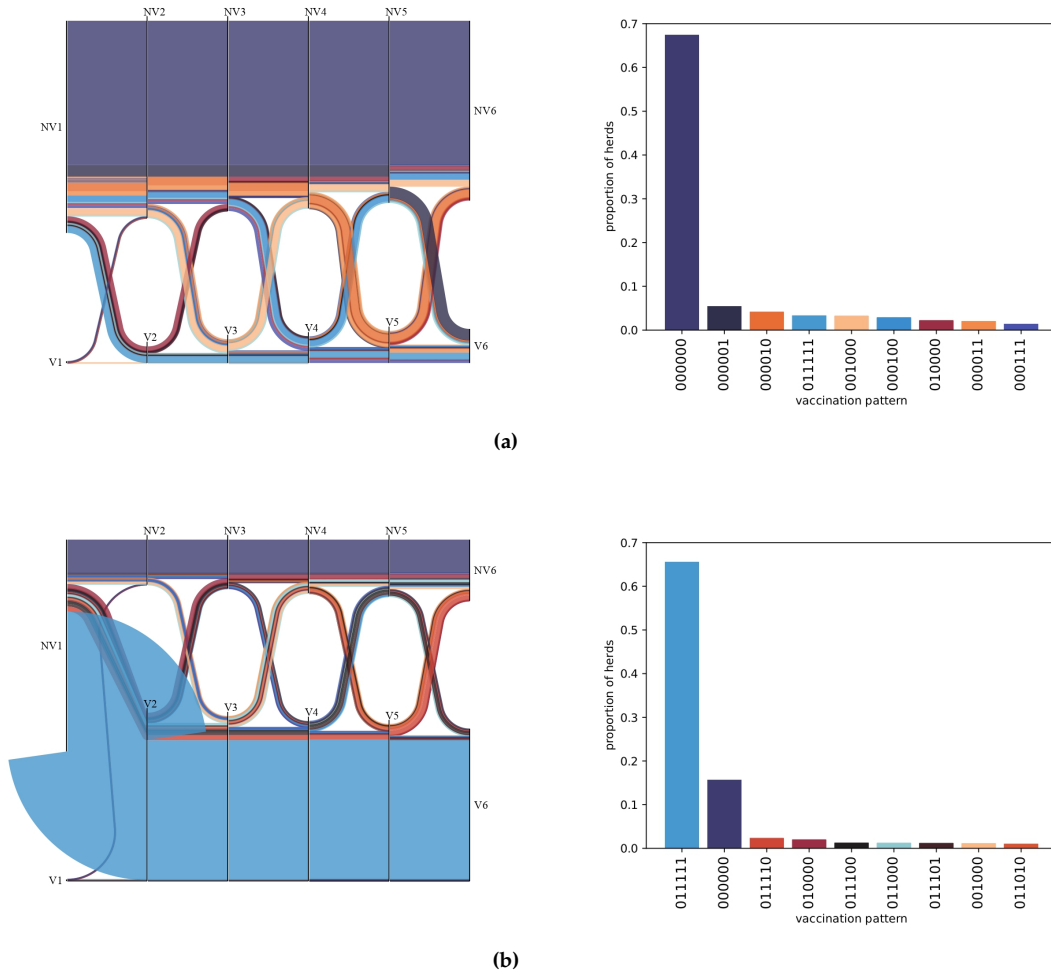


Figure 3.14: Temporal dynamics of the vaccination decisions using the decision mechanism defined in Algorithm 3 with (a) $\kappa = 0.5$, and (b) $\kappa = 12.5$. Results for one run. NV and 0 stand for not vaccinating, while V and 1 for vaccinating. Each color represents a different vaccination pattern, defined by the sequence of vaccination decisions at each of the six decision times. So the pattern 001111 (or equivalently [NV1, NV2, V3, V4, V5, V6]) concerns herds that do not vaccinate at the two first decision times, and always vaccinate afterwards. In the left plots, each vertical black line represents a decision time, and the width of the flows between decisions is proportional to the frequency of the pattern. In the right plots, the histogram of the patterns with a frequency $\geq 1\%$ is plotted.

3.4.2 Key determinant parameters of decision-making and epidemiological dynamics

We present the results for the sensitivity analyses on the means over runs for the concerned outputs in each experiment in figure 3.15. Overall, in figure 3.15a we see that according to experiment (i), the most influential parameters of the model are the epidemic parameter β/γ , which contributes to 25% of the variation of the means, and the decision related parameter Δ_d , which contributes to 14%. So only these two parameters account for more than 38% of the variation. They are followed by other epidemic and decision related parameters, as well as by an economic parameter. Furthermore, for each group of outputs, the parameters with the highest main effects on the means are of the same nature as the outputs (epidemic parameters have the greatest influence on epidemic outputs, economic parameters on the economic output, etc). For the epidemic outputs, the most influential parameter, β/γ , has a contribution of 61% to the variation

of the means. As expected, the exploration of simulation results evidences that this contribution translates into an increase in the propagation of the pathogen.

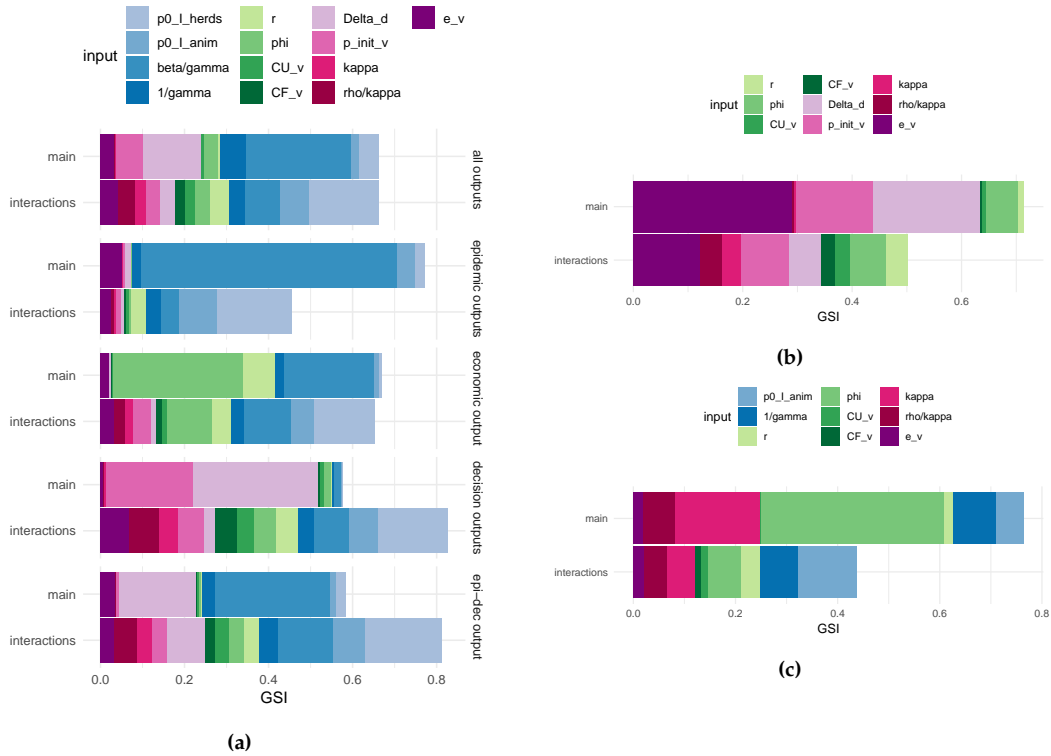


Figure 3.15: Global Sensitivity Indices (GSI) for the means over runs of the outputs considered in each experiment. Sensitivities are split in main effect and two-factor interactions. Blue colors correspond to epidemic parameters, green colors to economic parameters, and pink colors to decision related parameters. (a) GSI for the means of all outputs, and by group of outputs in experiment (i). (b) GSI for the means of all outputs in experiment (ii). (c) GSI for the means of decision outputs in experiment (iii). See Table 3.1 for parameters definition, and Table 3.2 for outputs definition.

When focusing on the mean of decision related outputs, even if interactions have the strongest effect, the most influential main effect is Δ_d , i.e. the duration between two consecutive decisions, which contributes 30% of the variation. It is followed by the initial probability of vaccinating, which contributes 21% to the variation. We remark that Δ_d has an overall negative influence on vaccination of herds, as it determines if control decisions are taken at early stages of the epidemic, and is therefore associated with a higher spread of the pathogen. The initial probability of vaccinating has, on the contrary, a positive effect on the vaccination and on the limitation of the epidemic spread. Concerning the interaction effects, epidemic parameters have the highest influence on the means of each group of outputs, and when considering the means of all outputs together. In particular, $p_{I_{herds}}^0$ is for each group the most influential parameter through its interaction effects. It mostly interacts with other epidemic parameters such as $p_{I_{anim}}^0$, but it has smaller interactions with other parameters as well.

For experiment (ii), figure 3.15b shows that when fixing epidemic parameters, overall the greatest main effects are those of the vaccine efficacy e_v (29% contribution), the duration of the decision Δ_d (19% contribution), and the initial probability of vaccinating p_v^{init} (14% contribution). They are followed by the main effect of the economic parameter: ϕ (6% contribution). Concerning higher order effects, we mainly observe interactions between the first three parameters: e_v , Δ_d , and p_v^{init} . Overall, e_v has the greatest interaction influence. Finally, figure 3.15c shows that in experiment (iii) the parameters ϕ and κ manage to explain about 50% of the variability of the means through their main effects, having a 35% and a 16% contribution, respectively. Each of the other parameters explains less than 10% of the variation.

Figure 3.16 presents the results of the three sensitivity analysis experiments regarding the variance of outputs over runs. We remark that in each experiments and for every group of outputs that we consider, the variation in the variances over runs is mostly explained by interaction effects between the input parameters.

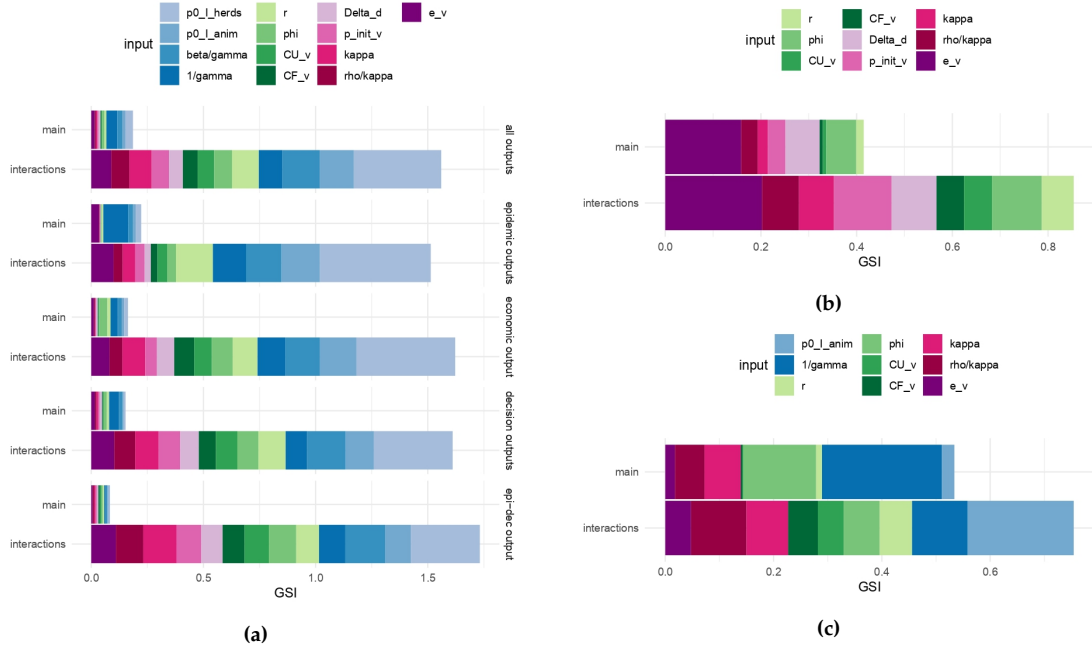


Figure 3.16: Global Sensitivity Indices (GSI) for the variances over runs of the outputs considered in each experiment. Sensitivities are split in main effect and two-factor interactions. Blue colors correspond to epidemic parameters, green colors to economic parameters, and pink colors to decision related parameters. (a) GSI for the variances of all outputs, and by group of outputs in experiment (i). (b) GSI for the variances of all outputs in experiment (ii). (c) GSI for the variances of decision outputs in experiment (iii). See Table 3.1 for parameters definition, and Table 3.2 for outputs definition.

3.5 Discussion

In this work we present a new integrative model for the epidemic spread of a livestock disease on a trade network, accounting for farmers' dynamic decisions concerning the adoption of a control measure in their herd. The model consists of an epidemic-demographic and a decision-making components that are interlinked through a feed-back loop. On the one hand, control decisions have consequences on the epidemic spread, both at the intra-herd and the inter-herd levels. On the other hand, the epidemic spread has an impact on the control decisions that farmers subsequently take. For the epidemic-demographic component we use a stochastic compartmental model with demography on a trade network, that accounts for intra-herd population changes, in particular those that concern animal transfers. For the decision-making component we assume the same dynamic decision problem for each farmer, and we propose a mechanism that represents their decision-making strategy.

Whereas most epidemiological models found in literature do not consider the voluntary adoption of a control measure for the spread of a disease (Wang et al. 2016), or consider an exogenous probability of applying the measure in order to only study the observed effects of decisions (Rat-Aspert et al. 2010), we propose a decision model that considers strategic interactions and cognitive considerations in the decision-making process. Our model can therefore be considered as a game-theoretical or a psychological model, according to the conceptual classification of behavioural epidemiological models found in Wang et al. 2016. The decision-mechanism we

propose takes into account different phenomena such as learning, stochastic behaviour, and imitation dynamics. To our knowledge, these elements are not present in the few existing models that have aimed at dynamically integrating the epidemic and decision-making processes of a livestock unregulated disease (Krebs et al. 2018). We remark that the basic structure of the decision-problem and the decision-mechanism can be found in different fields, particularly in the field of online optimization (such as multi-armed bandits (Auer et al. 2002a)). However, we do not seek to find an asymptotically optimal algorithm, which is often the goal in that area, but rather to describe farmers' decision-making process for the application of a control measure such as vaccination. More precisely, we consider an update of the probability of a farmer applying the measure, that is based on self-obtained results, and on neighbours' results.

In our model, farmer's next decision is based on a neighbour regardless of what the neighbour has decided in the previous step. This is not the case in similar models focused on human diseases (Fu et al. 2010; Kuga et al. 2019; Wang et al. 2012), in which a person only considers other people's observations if they have taken the opposite decision. In particular, this allows to always decrease the odds of a farmer vaccinating if both the farmer and his/her neighbour have previously vaccinated. Together with the use of the trade network as the information network in our model, this can amplify the emergence of strategic behaviours, as the farmer can search to benefit from the vaccination of one of their neighbours, while avoiding the cost of the vaccine. The behaviour where individuals (consciously or not) benefit from the actions of others without having to bear the cost, is known as free-riding, and has been previously addressed within vaccination decision-making models for human diseases (Wang et al. 2016). In particular, Ibuka et al. 2014 shows it is possible that individuals will consciously free-ride when making vaccinating decision.

Overall, our integrative model can be considered as an SIR model with pulsed vaccination (Choisy et al. 2007) in a metapopulation (Terry 2010), but where the pulse vaccination is asynchronous among sub-populations, and non equally spaced in time for each population, since the decision to vaccinate is not made at each decision time by each farmer. Our formalization of the integrative model is presented as general as possible so it can potentially be adapted to more complex epidemiological models or to other decision-making mechanisms that may be more relevant for specific contexts. Similar models have been previously proposed for human diseases (Fu et al. 2010; Kuga et al. 2019; Wang et al. 2012), yet none truly establishes the model in a generic manner in order to facilitate its adaptation for other diseases, or control measures. Even if the economic cost we propose is associated with vaccination and the consequences of an SIR model, its basic structure could take into account the epidemiological and economic consequences of other measures, for example a treatment that would increase the recovery rate of infected animals. In particular, if the epidemic model was aged-structured, the cost on which farmers base their decisions could be refined to take into account the age of the animals. The real-life farmers decision-making being undoubtedly complex, the decision model we propose is reductive. Yet it provides a complete and adaptive framework with respect to state-of-the-art methods in veterinary epidemiology. In the presence of detailed information on farmer's real-life behaviour, our model could be run with other parameter values, or it could be modified to stick closer to reality if observations on farmers' decisions denote a different decision-making process.

Among the methodological extensions to consider, we believe that the model could mostly benefit from a relaxation of some hypotheses in the decision-making mechanism. First, we consider that farmers perfectly observe the costs associated to their control decisions, as well as the decisions and costs of their neighbours, which is not completely realistic. Actually, farmers may observe costs with some error, or neighbours may not communicate their true actions or costs. Second, we assume that the trade network is the information network through which farmers share their observations. But farmers may be informed about other herds control practices in a more aggregated way, or only from geographical neighbours. Furthermore, from an economic point of view, in our decision-model we consider only the financial results of the farmer's decision, which is in principle a good indicator of what interests him/her. We remark however that farmers may have social, personal or environmental motivations for taking decisions related to animal welfare (Hansson et al. 2015). For example, farmers may have a non-use value for their

animals, that is, a value related to the animal well-being independently of the use the farmer may make of the animal (Lagerkvist et al. 2011). Even if some refinement could be made in this direction, this does not seem straightforward from a mathematical modelling perspective. However, our decision model can implicitly integrate this information through the values of the initial probability of vaccinating and the parameters κ and ρ . In addition, considering other types of farmers' behaviour can be of interest in this context. For example, the adaptation of the exchange network as a function of other farmers' health state. This intervention is known as network rewiring (Kiss et al. 2017), and is generally appropriate for regulated diseases for which there is aggregated information on the health status of each herd, i.e. neglecting the intra-herd epidemic dynamics. Yet, even with this aggregated information, the adaptation of the network can be quite complex. Lastly, an exploration on the emergence of collaborative behaviours (Cardillo et al. 2013; Karlsson et al. 2020; Wang et al. 2017), in particular through network reciprocity (Nowak 2006), can be an interesting perspective for a deeper understanding of the observed decision dynamics.

Regarding model's predictions, simulations evidence the retroactive effect between the dynamics of the epidemic spread, and the dynamics of the vaccination decisions. A deeper examination of the model through sensitivity analysis confirms that decision-parameters play a role in the model's behaviour. Apart from the epidemic parameters, the time between two consecutive decisions has the highest impact overall, and is the main driver in decision related outputs. Indeed, the shorter the time between decisions, the more frequently farmers evaluate their information on epidemic spread, and the fastest they start vaccinating if necessary. A constantly updated local information on the disease spread regularly helps updating farmers' vaccination decisions from the beginning of the epidemic, and is therefore crucial for limiting the disease spread. This is consistent with observations from models for human diseases, where health information can produce the eradication of the disease if there is a rapid diffusion of this information and if individuals decide to act based on this information (Kiss et al. 2010). Furthermore, it has been documented that the impact of locally spreading information is amplified if information and disease transmission networks overlap (Funk et al. 2009), as it is the case in our model. Finally, an extension of the model where each farmer considers all of his/her neighbours decisions in the decision-making process, evidenced small differences with respect to the model considering only one neighbour per decision instant. In particular, when farmers have a small sensitivity to costs, taking into account all of their neighbours seems to be slightly better in the short term for controlling the epidemic diffusion, but not in the long term. When farmers have a very high sensitivity to costs, considering all of their neighbours does not significantly change the course of the epidemics with respect to the scenario where they consider only one neighbour.

Overall, we conclude that our model effectively integrates the dynamics of the decision process regarding the voluntary adoption of a sanitary measure in each herd, and the dynamics of the epidemic spread over a structured population of herds in a trade network. Hence, we make a significant step towards accounting for human decision-making in mechanistic epidemiological models, in particular for endemic animal diseases. Given its integrative structure, its flexibility and stability in results, our model can be well adapted for simulation studies concerning specific real-life diseases or other control measures.

Chapter 4

Dynamic resource allocation for controlling pathogen spread on a large metapopulation network

Cristancho-Fajardo L., Ezanno P. and Vergu E. (2022) "Dynamic resource allocation for controlling pathogen spread on a large metapopulation network" In: *Journal of the Royal Society Interface* 19. DOI: 10.1098/rsif.2021.0744

Contents

4.1	Introduction	80
4.2	Dynamic resource allocation in the metapopulation framework	81
4.3	Score-based strategies	82
4.3.1	Greedy scores	82
4.3.2	Heuristic scores	84
4.4	Simulation setting and numerical explorations	84
4.4.1	Setting for the exploration of infection-related dynamics with score-based resource allocation	84
4.4.2	Setting for percolation analysis	86
4.5	Results	87
4.5.1	Greedy scoring functions	87
4.5.2	Results of numerical explorations	88
4.6	Discussion	102
4.7	Appendix: Analytical derivation of the greedy scoring functions	105
4.7.1	Vaccination	106
4.7.2	Treatment	110

To control the spread of an infectious disease over a large network, the optimal allocation by a social planner of a limited resource is a fundamental and difficult problem (section 4.1). We address this problem for a livestock disease that propagates on an animal trade network according to an epidemiological-demographic model based on animal demographics and trade data (section 3.2.1). We assume that the resource is dynamically allocated following a certain score, up to the limit of resource availability. We adapt a greedy approach to the metapopulation framework, obtaining new scores that minimise approximations of two different objective functions (sections 4.3.1, 4.5.1 and 4.7), for two control measures: vaccination and treatment (section 4.3). Through intensive simulations, we compare the greedy scores with several heuristics (section 4.4). Although topology-based scores can limit the spread of the disease, information on herd

health status seems crucial to eradicating the disease. In particular, greedy scores are among the most effective in reducing disease prevalence, even though they do not always perform the best (section 4.5.2). However, some scores may be preferred in real life because they are easier to calculate or because they use a smaller amount of resources. The developed approach could be adapted to other epidemiological models or to other control measures in the metapopulation setting (detailed in section 4.6).

The modelling choices in this chapter were motivated by the state of the art in section 2.1.1.1.2, section 2.1.1.1.4 and section 2.1.3.4, regarding epidemiological modelling, and in section 2.2.3 and section 2.3, regarding centralised decision-making.

4.1 Introduction

Infectious disease spread is a problem that can have important social, sanitary, and economic consequences. Like for human diseases, this is a major public health concern for animal diseases, for guaranteeing animal welfare and food security (Tomley et al. 2009). In this context, epidemiological models, together with other relevant mathematical approaches, can help in the description and understanding of the mechanisms involved in the disease propagation, as well as in assessing the effectiveness of control measures (Heesterbeek et al. 2015). An approach for controlling a disease spreading on a population, from a social planner's point of view, is the allocation of a resource that has an effect on this spread (Preciado et al. 2013). Many questions can arise in this context: how much resource is needed to restrain the disease propagation to a certain level (Enayati et al. 2020; Holme et al. 2017; Tanner et al. 2008), when should it be allocated (Thompson et al. 2018), and where. In this work we are interested in the third question. More specifically, we are concerned with the problem of dynamically deciding where to allocate a limited available resource in an optimal manner, in order to minimise disease spread on a large animal metapopulation network.

On the one hand, most of the research addressing the issue of resource optimal allocation on a large network (Lorch et al. 2018; Scaman et al. 2016; Wijayanto et al. 2019; Zhan et al. 2017; Zhang et al. 2015) does not focus on metapopulation networks, i.e. does not account for infection-related dynamics within each sub-population represented by a node of the network. Even more, works relying on mean-field theory (e.g. Forster et al. 2007), do not consider structured populations, in particular as a network.

On the other hand, the existing studies addressing the resource allocation problem on metapopulation networks are based on techniques that lack of scalability. For example, optimal control (Lenhart et al. 2007) and reinforcement learning (Sutton et al. 2018), which would theoretically give an optimal strategy, cannot be used in the context of very large networks due to the dimensionality in the space state (Andriotis et al. 2019; Dvijotham et al. 2013). Even if we consider a sub-population as healthy or infected (only two possible states per sub-population), a network with J sub-populations would have 2^J possible health states, which yields an asymptotically intractable optimal allocation when the network is very large. Hence, authors that study resource allocation on metapopulation networks generally build and evaluate their approach in a small number of sub-populations, usually less than 50 (Bussell et al. 2019; Carli et al. 2020; Chernov et al. 2020; Duijzer et al. 2018; Long et al. 2018; Matrajt et al. 2013; Mbah et al. 2011; Rao et al. 2021). Recently, Lemaitre et al. 2021 proposed a framework built upon optimal control theory that is able to deal with the dynamic allocation problem in a network of hundreds of sub-populations, thanks to several simplifications, among which considering only a subset of edges for the optimisation. Yet, such a scale does not allow to capture the structural characteristics of complex networks, such as the animal trade network we consider here. In particular, animal trade networks are in general scale-free (Rautureau et al. 2011) (most herds have few trading partners while a few herds have many) and dynamic, as the amount of exchanged animals can vary over time (Brooks-Pollock et al. 2015).

Finally, published studies assume the resource is in general distributed only once, before or at

the beginning of an outbreak (e.g. Dangerfield et al. 2019; Rao et al. 2021). Therefore, the resource allocation problem is static. However, the allocation problem can be intrinsically dynamic if it is studied in the long run. For example, if the available resource is a vaccine, these can have a limited effect in time, so there is need for several vaccination campaigns.

Given the intractability of the optimal strategy, in this work we restrict ourselves to score-based strategies, i.e. strategies that consist in allocating the resource according to a scoring function (or indicator), up to the limit given by the available quantity of the resource. Furthermore, given the practical importance of dynamical aspects of the allocation problem, we include this view in the present study, i.e. we assume the resource allocation is dynamic.

The contribution of this work is twofold. First, from a methodological perspective, by adapting the greedy approach in Scaman et al. 2015, we obtain analytic scores for controlling disease spread on a large animal trade network, where the disease propagation is represented by a stochastic SIR model that accounts for demography and trade (model introduced in section 3.2.1). The approach consists in finding the scoring function that minimises a short-term approximation of a given objective function (sections 4.2, 4.3.1 and 4.5.1). Our generalisation is mainly driven by the metapopulation framework, which implies that a herd is not only infected or healthy, but that it has an internal infection and demographic dynamics. In particular, this allows for the possibility of needing different amounts of resources for different herds. Furthermore, we extend this approach for two different types of resources (vaccination and treatment) and two different objective functions (the number of infected animals and the number of infected herds). Second, regarding real-life disease control for livestock diseases, in section 4.5.2 we evaluate the performances of the analytically obtained strategies along with the one of several heuristic strategies (introduced in section 4.3.2) that can be relevant for this context. Finally, in section 4.6 we extend the interpretation of the analytically found scores and discuss their suitability in a metapopulation context in the light of simulation-based results. We also consider several perspectives, either based on the development of other greedy scores or on new simulation studies that could use the explorations performed in this work.

4.2 Dynamic resource allocation in the metapopulation framework

We supposed that there is a central social planner seeking to minimise the disease propagation on the animal trade metapopulation network detailed in section 3.2.1, by distributing a limited amount of a resource among the herds in the network, dynamically with a given decision time-step. This dynamic resource allocation problem was formulated as:

$$\min_A F(A) \quad \text{subject to} \quad \sum_{j=1}^J b_j(t) A_j(t) \leq b_{fix}, \quad \forall t = \Delta_d, 2\Delta_d, \dots \quad (4.1)$$

$F(A)$ in equation 4.1 is the function that the planner has to minimise, which depends on the allocation strategy A , i.e. the function that determines the decisions $A_j(t)$, $\forall j = 1, \dots, J, \forall t = \Delta_d, 2\Delta_d, \dots$, where j denotes the herd, and Δ_d is the decision time-step. We assumed binary allocation decisions for each herd, i.e. $A_j(t) = 1$ if the resource is allocated to herd j at time t , $A_j(t) = 0$ otherwise. In the condition of equation 4.1, b_{fix} is the quantity of resource that is available at each decision time and $b_j(t)$ is the quantity of the resource that would be needed for herd j at time t if this herd was selected for the allocation.

We supposed that the resource to allocate could either be a vaccine or a treatment. For the vaccine, we assumed that if applied to a susceptible animal at time t , the disease transmission rate towards this susceptible becomes $\beta^v = \beta(1 - e_v)$ during the period $]t; t + \Delta_d]$, where $0 \leq e_v \leq 1$ is the protective efficacy of the vaccine. That is, Δ_d is also the duration of the vaccine's efficacy. We underline that even if all susceptible animals in herd j are vaccinated at time t , infections can

occur within j from time t to $t + \Delta_d$ if new susceptible animals enter the herd, through births or imports of susceptible animals from other herds. For the treatment, we supposed it increases the recovery rate of treated infected animals by an additional factor γ' , i.e. reducing the mean duration of the infectious period for these animals to $(\gamma + \gamma')^{-1}$.

Regarding the resource constraint, for vaccination we assumed it concerned the number of available doses, and if a herd j was selected at time t all the animals in the herd would be vaccinated. So in the condition in equation 4.1, $b_j(t) = N_j(t)$ (the size of herd j at time t). For the treatment, the resource constraint was on the number of herds in which animals were treated at each decision time, so $b_j(t) = 1$. The choice to consider that the constraint for the treatment involves the number of herds was mainly motivated by analytical considerations discussed in section 4.3.1.

4.3 Score-based strategies

Score-based strategies consist in ordering herds according to a certain scoring function Ξ and selecting the top herds, up to the limit given by the condition in equation 4.1. Let $\Omega_\Xi(t)$ be the set that contains the selected herds according to $\Xi(t)$ (the score values at time t) and b_{fix} (the available quantity of resource per decision time-step). Then, $A_j(t) = 1$ if $j \in \Omega_\Xi(t)$, 0 otherwise. Figure 4.1 represents the modelling and optimisation framework of the dynamic resource allocation of vaccines under a score-based strategy. The treatment allocation differs only in the infection and demographic dynamics component, and in the control measure component. That is, for any herd j , there is an additional compartment T_j for treated infected animals, where animals go from I_j to T_j at decision time t , if $A_j(t) = 1$. Unlike vaccination, the transition from T_j to I_j is not possible even if $A_j(t) = 0$, i.e. treated individuals can only recover.

Regarding the scoring function, it can either be optimised or heuristic. In the following we address the two possibilities.

4.3.1 Greedy scores

First, following the approach in Scaman et al. 2015 we searched for optimised scoring functions. This approach consists in finding a scoring function that minimises a short-term approximation of the objective function:

$$F(A) := \int_0^\infty e^{-at} \mathbb{E}[J_I^A(t)] dt,$$

where $J_I^A(t)$ is a function of the infection-related state of the network at time t . The objective establishes an infinite time horizon, and a parameter $a \geq 0$ which reduces the long term impact of $J_I^A(t)$. This means that the larger a , the more we are interested in the short-term behaviour of the infection dynamics, i.e. the more weight we put on the initial stochasticity of the disease diffusion.

Let us note $X(t)$ the infection-related state of the network at time t . In the SIR metapopulation framework, $X(t)$ is a $J \times 3$ matrix where each row $j \in \{1, \dots, J\}$ contains the values of $S_j(t)$, $I_j(t)$ and $R_j(t)$ at time t . Since the allocation strategy has no impact on the initial state, and since the process is Markovian, the minimisation problem (equation 4.1) is equivalent to:

$$\min_A \int_{u=0}^\infty e^{-au} \mathbb{E} [J_I^A(t+u) | X(t) = X] du, \quad (4.2)$$

for all decision times t and for all network possible states X , under the same constrain as in equation 4.1.

Then, focusing on the short-term behaviour of the system, it is possible to obtain an approximation of the objective function, which is based on a Taylor expansion of $\mathbb{E} [J_I^A(t+u) | X(t) = X]$

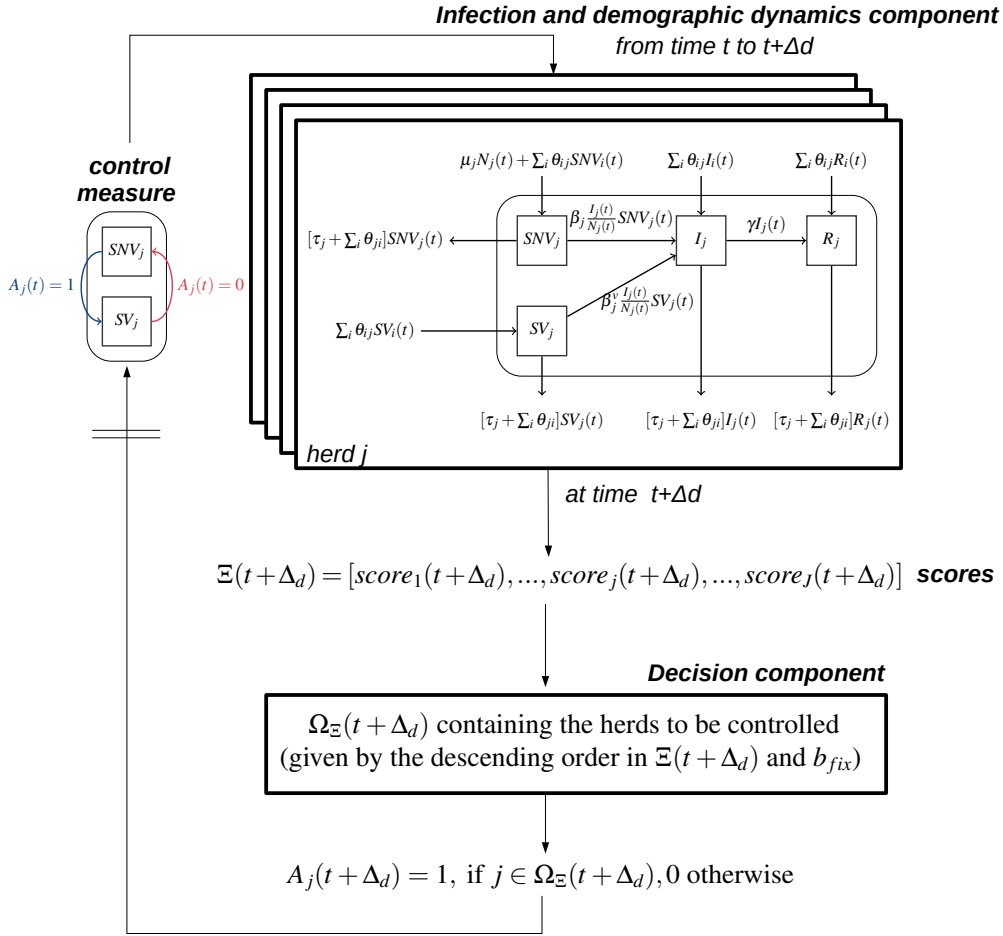


Figure 4.1: Representation of the modelling and optimisation framework of the score-based dynamic resource allocation under constraint, applied to a metapopulation and vaccine allocation. See main text in sections 3.2.1 and 4.2 for parameter definitions.

for a small value of u . Finally, one must find the scoring function that minimises the obtained approximation. This approach is therefore called greedy, as it yields locally (i.e. short-term) optimal allocation decisions. We stress that the form of the scoring function will not depend on the parameter a in equation 4.2, i.e. we do not need to fix a value for this parameter when exhibiting the scoring function. Yet, the higher is a , the lower is the impact of the approximation accuracy. See section 4.7 for more details on the approach.

We adapted this greedy approach to the metapopulation framework by considering, for each of the two types of resource (vaccination and treatment), the minimisation of two different objective functions in equation 4.2. A function on the number of infected animals:

$$J_I^A(t) = \sum_{j=1}^J I_j(t), \quad (4.3)$$

and a function on the number of infected herds:

$$J_I^A(t) = \sum_{j=1}^J \mathbb{1}_{I_j(t) > 0}. \quad (4.4)$$

That is, we treated four different cases, depending on the type of resource (vaccine or treatment) and the objective to be minimised: the number of infected animals (equation 4.3) or the number

of infected herds (equation 4.4). In particular, the derivation of the score for the allocation of a treatment under the objective of equation 4.4 required to consider that the resource constraint of equation 4.1 was expressed in terms of the number of attainable herds ($b_j(t) = 1$). This allowed the total number of treated herds to be formulated as the minimum between b_{fix} and the number of herds that have exactly one infected animal (see section 4.7.2.1 for details).

4.3.2 Heuristic scores

In addition, we considered three types of heuristic scores based on: the topology of the static aggregated network; the demographic changes in the network; the dynamic infection-related state of the network. Table 4.1 contains the list of the 16 heuristic scoring functions we tested for the metapopulation framework: 5 topological ones (in-strength, out-strength, closeness, betweenness, pagerank), 5 demographic ones ($N_j(t)$, $\text{purchases}_j(0, t)$, $\text{purchases}_j(t - \Delta_d, t)$, $\text{sales}_j(0, t)$, $\text{sales}_j(t - \Delta_d, t)$), 5 epidemiological ones ($s_j(t)$, $i_j(t)$, $r_j(t)$, $i_j(t - \Delta_d, t)$, $r_j(t - \Delta_d, t)$), and a random scoring function. All the topological scoring functions are classical centrality measures in networks (Newman 2018).

4.4 Simulation setting and numerical explorations

The simulation setting was described in detail in section 3.3.1 of chapter 3. The metapopulation structure was set close to real data, extracted from the French Cattle Identification Database (FCID): animal movements correspond to the Finistère administrative area in Western France, which is densely populated with cattle. The trade parameters ($\theta_{ji}, \forall j, i = 1, \dots, J$) underlying the network structure were set to values based on this data. It is scale-free and consists of $J = 5000$ herds.

Without loss of generality, the values of demographic and epidemiological parameters were set to the same value for all herds. In particular, death and birth rates were $\tau_j = \tau, \mu_j = \mu, \forall j = 1, \dots, J$, where $\tau = 0.0009 \text{ days}^{-1}$ and $\mu = 2\tau$. So, based on field reality, animals have a mean life time of approximately three years, and on average an animal gives birth to a calf every 1.5 years. As herd sizes could vary not only through births and deaths, but also through animal movements, we established a constraint on the capacity of each herd, $N_j(t) \leq 1.5N_j(0)$, so that these remain relatively stable over time. Figure 3.3b in chapter 3 shows the final herd sizes distribution for a run of the epidemiological-demographic model (without resource allocation) on the simulated trade network. Regarding the epidemiological parameters, we set $\beta_j = \beta, \forall j = 1, \dots, J$, and considered a disease with moderate immediate impact and long-lasting development ($\beta/\gamma = 2$ and $1/\gamma = 90$ days). A second numerical setting, corresponding to a disease with higher early peak and smaller infection duration ($\beta/\gamma = 4$, $1/\gamma = 30$ days), was also explored. Finally, for the available resource, we assumed it could either be a perfectly effective protective vaccine, i.e. $e_v = 1$, or a treatment that greatly reduces the infectious period of infected animals (but which is not perfect in order to avoid instantaneous recovery, an unrealistic assumption). More specifically, the duration of the infectious period with treatment was assumed to be 3% of the duration without treatment, i.e. the mean duration of the infectious period for a treated infected animal is $(\gamma + \gamma')^{-1} = 0.03(1/\gamma) = 2.7$ days.

The simulation code is available at github.com/CristanchoLina/DRAAnimalMetapop.

4.4.1 Setting for the exploration of infection-related dynamics with score-based resource allocation

Given these parameter values, we simulated the infection-related dynamics of the metapopulation during 3 years in 74 ($= ([16+3] + [16+2]) \times 2$) cases characterised by the type of resource, the score according to which it is allocated (16×3 for vaccination and 16×2 for treatment) and the

	Scoring function for herd j	Description
Topological	in-strength_j	weighted in-degree, i.e. the sum of the weights of all edges going to j . Weights \equiv daily mean trade rates. Measures how much j purchases from other herds.
	out-strength_j	weighted out-degree, i.e. the sum of the weights of all edges going from j . Weights \equiv daily mean trade rates. Measures how much j sells to other herds.
	closeness_j	inverse of the sum of distances to other herds, based on the weighted shortest paths going from/to j . Weights \equiv daily mean trade rates. Measures how 'close' j is to other herds.
	betweenness_j	ratio of the weighted shortest paths that pass through j over all such paths in the network. Weights \equiv daily mean trade rates. Measures the influence of j on the animal exchanges of the network.
	pagerank_j	stationary density of a discrete-time random walk. Weights \equiv daily mean trade rates. Measures the accessibility to j .
Demographic	$N_j(t)$	current herd size.
	$\text{purchases}_j(0, t)$	} number of animals bought by j over the period $]0, t]$ and $]t - \Delta_d, t]$, respectively.
	$\text{purchases}_j(t - \Delta_d, t)$	
	$\text{sales}_j(0, t)$	} number of animals sold by j over the period $]0, t]$ and $]t - \Delta_d, t]$, respectively.
$\text{sales}_j(t - \Delta_d, t)$		
Epidemiological	$s_j(t)$	current proportion of susceptible animals in j : $\frac{S_j(t)}{N_j(t)}$.
	$i_j(t)$	current proportion of infected animals in j : $\frac{I_j(t)}{N_j(t)}$.
	$r_j(t)$	current proportion of recovered animals in j : $\frac{R_j(t)}{N_j(t)}$.
	$i_j(t - \Delta_d, t)$	increment of the proportion of infected animals in j between times $t - \Delta_d$ and t : $\frac{I_j(t)}{N_j(t)} - \frac{I_j(t - \Delta_d)}{N_j(t - \Delta_d)}$.
	$r_j(t - \Delta_d, t)$	increment of the proportion of recovered animals in j between times $t - \Delta_d$ and t : $\frac{R_j(t)}{N_j(t)} - \frac{R_j(t - \Delta_d)}{N_j(t - \Delta_d)}$.
	$\text{random}_j(t)$	assigns a random score for j at each decision time t .

Table 4.1: Heuristic scoring functions for herd j at time t . Dependence on t means the score is dynamic in time, otherwise it is static.

scenario. This last one can be: an *epidemic* scenario, where initially 10% of the herds (chosen completely at random) had a random subset of 15% of their animals infected; or an *endemic* scenario, where the initial state was given by the state at roughly 3 years (1080 days) without resource allocation departing from the epidemic scenario. Indeed, figure 3.12 in chapter 3 shows that if the simulations are extended beyond 3 years, the total proportion of infected animals remains rather stable, and that there is only a 10% reduction in the proportion of infected herds between levels attained at 3 and 9 years. Hence, although the infection dynamics after 3 years of simulation did not reach a steady state state rigorously speaking, this date was chosen as the initial point of the endemic scenario. Indeed, on the one hand, at this date the pathogen had widely spread in the metapopulation (see figure 3.12 in chapter 3), and on the other hand, considering 3 years limits the simulation cost. In each case, we explored the dynamics of the proportion of infected herds and of the total number of infected animals for a fixed value of the available quantity of resource, b_{fix} . The values of b_{fix} and Δ_d parameters can be found in table 4.2. We supposed vaccination decisions were more spaced in time since vaccines have a preventive nature and tend to

have long-lasting effect. Meanwhile, we supposed that treatment decisions were more frequent as they are more prone to being applied in a critic situation.

Resource	Parameter	Definition	Values
Vaccine	b_{fix}	number of available doses at each decision time (as a % of the initial total number of animals).	$(25\% \times \sum_{j=1}^J N_j(0))$ ¹ and $([5\% - 90\%] \times \sum_{j=1}^J N_j(0))$ ²
	Δ_d	decision step (in days).	180
Treatment	b_{fix}	number of attainable herds at each decision time.	25 ¹ and [5 - 100] ³
	Δ_d	decision step (in days).	15

Table 4.2: Parameter values in the allocation problem depending on the type of resource. ¹Values for infection-related dynamics explorations. ² Values for percolation analysis of vaccination. From 5% to 30%: by 0.5%. From 40% to 90%: by 10%. ³ Values for percolation analysis of treatment. From 5 to 75: by 5. From 80 to 100: by 10.

We stress that in all simulations for the vaccine allocation, only herds with at least one susceptible animal were eligible, and that the treatment allocation only considered herds with at least one infected animal.

Furthermore, we evaluated how other quantities of interest, raw proxies of ‘costs’, varied with the chosen scoring function and the amount of available resource. We stress that the quantity of resource that is actually used may be less than the available quantity (b_{fix}). For vaccination, we explored the mean proportion of wasted doses over time, i.e. the total number of vaccine doses allocated to infected or recovered animals, out of the number of available doses. For the treatment, where we assumed that only infected animals were treated, we evaluated the mean number of used doses over time, and we explored the mean size of the target population, i.e. the total number of animals in treated herds.

4.4.2 Setting for percolation analysis

We performed percolation experiments Li et al. 2021; Stauffer et al. 2018 by assessing the final (after 3 years) proportion of infected herds when the allocation was done using a given scoring function, and a fixed quantity of resource. We also included in this analysis an exploration of the raw proxies of costs. The range of values tested for b_{fix} (cf. table 4.2) was chosen so as to exacerbate differences in the infection-related dynamics by scoring function. The objective was to analyse the dependence of the relative performances of the scoring functions on limiting disease spread with respect to the available quantity of the resource. To avoid increasing computation time, we chose to run the percolation analysis only for a subset of scoring functions. This subset was determined from the results of the complete exploration setting described in section 4.4.1.

Finally, we investigated the sensibility of the infection-related results of percolation analysis to the value of the vaccine efficacy e_v . More specifically, we evaluated through simulations two cases where vaccine efficacy is not perfect: $e_v = 0.9$ and $e_v = 0.7$. Lower values for e_v were not considered, since it is very unlikely that a vaccine for a livestock disease with an efficacy below 0.7 is even considered in the field, as its implementation could induce greater economic costs than epidemiological benefits. With each of these values we ran the percolation analysis described above, for three values of the available number of doses: [25%, 40%, 70%] of the initial total number of animals in the metapopulation.

4.5 Results

4.5.1 Greedy scoring functions

We obtained four different analytic scoring functions by considering the minimisation problem in equation 4.2 for each resource and for each objective with the greedy approach. The scoring functions for an objective on the number of infected herds (greedyV_infherds and greedyT_infherds) were obtained through a second-order Taylor expansion, while the other two scoring functions (greedyV_infanimals and greedyT_infanimals) were obtained using a first-order approximation. Details on how the greedy scoring functions were computed in each case can be found as an appendix in section 4.7.

As mentioned in the previous section, the values of β_j, τ_j, μ_j were set equal for all herds in our simulations, so we present in table 4.3 the scores in this setting. The greedy scoring functions found for the generic framework, where these values can be different across herds, can be found in table 4.4. In simulations, five scoring functions (directly or indirectly issued from the greedy approach) were used. Table 4.3 includes in total six scoring functions: the four obtained by optimisation, an additional scoring function for vaccination (greedyV_infherds_threshold) and the scoring function greedyT_infherds_threshold for the treatment, which replaced greedyT_infherds in our simulations.

Resource	$J_j^A(t)$	Scoring function for herd j	Scoring function name
Vaccine	inf. animals (eq. 4.3)	$\frac{I_j(t)}{N_j(t)} S_j(t)$	greedyV_infanimals
	inf. herds (eq. 4.4)	$\frac{I_j(t)}{N_j(t)} S_j(t) \left(\mathbb{1}_{I_j(t)=1} (\gamma + \tau + \sum_{i \neq j}^J \theta_{ji}) + \sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i(t)=0} \right)$	greedyV_infherds
	inf. herds (eq. 4.4)	$\frac{I_j(t)}{N_j(t)} S_j(t) \left(\mathbb{1}_{0 < I_j(t) < 20} (\gamma + \tau + \sum_{i \neq j}^J \theta_{ji}) + \sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i(t)=0} \right)$	greedyV_infherds_threshold
Treatment	inf. animals (eq. 4.3)	$I_j(t)$	greedyT_infanimals
	inf. herds (eq. 4.4)	$\left[- \sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i(t) > 0} \right]_{j: I_j(t)=1}$	greedyT_infherds
	inf. herds (eq. 4.4)	$\left[\left(- \sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i(t) > 0} \right) \mathbb{1}_{0 < I_j(t) < 20} \right]$	greedyT_infherds_threshold

Table 4.3: Greedy scoring functions studied in the numerical explorations. All the greedy scores are dynamic.

To minimise the number of infected animals by distributing a vaccine, the greedyV_infanimals scoring function privileges herds with a large within-herd incidence rate ($\beta I_j(t) S_j(t) / N_j(t)$), i.e. many infected animals and a large proportion of susceptible animals. For the treatment, with greedyT_infanimals the allocation would be made only as a function of the number of infected animals by herd ($I_j(t)$).

Regarding the minimisation of the number of infected herds, for vaccination it led to a scoring function, greedyV_infherds, favouring two types of herds: either herds that have a large within-herd incidence rate, and that send to many healthy herds (large $\sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i(t)=0}$); or herds with only one infected animal, also presenting a large proportion of susceptible animals and which sell many animals (large $\sum_{i \neq j}^J \theta_{ji}$). With the same objective for treatment, the greedyT_infherds scoring function only concerns herds with exactly one infected animal, and among these, the priority is on herds that send the smallest flows to infected buyers (small $\sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i(t) > 0}$).

We remark that the minimisation of the function on the number of infected animals for vaccination and treatment led to scoring functions that only depend on the epidemiological state of herd j , but not on the states of other herds, and in particular not on the topology of the network. This is due to the use of a first-order Taylor development for approximating the objective function (see sections 4.7.1.2 and 4.7.2.2 for details).

The two additional scoring functions considered on the basis of the analytically obtained scores were built in the following way.

For vaccination, the additional scoring function greedyV_infherds_threshold consists in replacing $\mathbb{1}_{I_j(t)=1}$ by $\mathbb{1}_{0 < I_j(t) < 20}$ in greedyV_infherds. This intends to avoid that the first term of the sum in the scoring function becomes 0 for herds that have few infected animals but not neces-

sarily just one. Similarly, for the treatment, we replaced `greedyT_infherds` by a scoring function with a softer condition on the number of infected animals. The condition in `greedyT_infherds` on having exactly one infected animal for a herd to be eligible appeared to be too restrictive. Indeed, if the quantity of available treatment exceeded the number of herds that satisfy this condition, the rest of the treatment would not be allocated to any herd. Yet, allocating the exceeding treatment to herds with more than one infected animal could only be beneficial for limiting disease spread and would satisfy the constraint on the quantity of available treatment. Hence, the `greedyT_infherds_threshold` scoring function considers all herds that are potentially eligible with a non-negligible probability, yet favours herds that have few infected animals ($\mathbb{1}_{0 < I_j(t) < 20}$).

4.5.2 Results of numerical explorations

4.5.2.1 Infection-related dynamics following score-based resource allocation

Figure 4.2 presents the results for the dynamics of the proportion of infected herds, under the setting described in section 4.4.1 for a subset of the scoring functions. In particular, since the results for the topological scores were very similar, we present only the results of the `pagerankj` score. This is also the case for demographic scores, so we chose the sales over the decision period, $\text{sales}_j(t - \Delta_d, t)$, as the representative score for this group. For the epidemiological scores, we present only the results for the best performing score, $i_j(t)$. In addition, we included as the best and worst reference cases, results for cases where there is sufficient resource for all herds (*full_budget*), and where there is no resource to allocate (*no_budget*). Figure 4.3 presents the complete results for cases by scenario (epidemic or endemic), type of available resource (vaccine or treatment) and heuristic or greedy score according to which allocation is performed. We also included in figure 4.3 results for the dynamics of the total number of infected animals, yet we remark they were similar to the ones found for the proportion of infected herds.

One of the main remarks emerging from the analysis of figures 4.2 and 4.3, is that topological and demographic scoring functions were the groups that performed less well for limiting the disease spread, both for vaccination and for the treatment. This is in comparison with the group formed by the scoring functions that take into account the health statuses of the herds, i.e. the greedy scoring functions, and the best performing epidemiological functions, in particular $i_j(t)$.

Furthermore, figure 4.3 shows that for vaccination, the scoring function $s_j(t)$ (the proportion of susceptible animals) performed the worst for allocating the available quantity of vaccines (equal to 25% of the initial number of animals in the metapopulation) in both scenarios. The group of topological scoring functions performed better than the random scoring function in the epidemic scenario, but was not very different from it in the endemic scenario. Demographic scoring functions performed in general better than the topological ones, in particular $\text{sales}_j(0, t)$ and $\text{sales}_j(t - \Delta_d, t)$. Apart from $s_j(t)$, the epidemiological scoring functions had a good performance, except for the proportion of recovered animals ($r_j(t)$) in the endemic scenario, where it was just as good as the random score.

The second important remark (figure 4.2) is that the best performing scores were the proportion of infected animals ($i_j(t)$), and `greedyV_infanimals`. Figure 4.3 shows that this applies both for reducing the total number of infected animals and the proportion of infected herds. They were closely followed by `greedyV_infherds_threshold` in both scenarios. Meanwhile, `greedyV_infherds` was not as good as these three, particularly in the endemic scenario.

Figures 4.2 and 4.3 also present results when the resource was a treatment and b_{fix} was equal to 25 herds. In particular, figure 4.3 shows that the epidemic scenario, $s_j(t)$ performed badly when compared to the other scoring functions. Yet $r_j(t)$ arrived to perform worse at the end of the three years in this scenario, and was the worst-performing scoring function in the endemic scenario. Even more, these two scoring functions, the topological and demographic ones, and the difference in the proportion of recovered animals ($r_j(t - \Delta_d, t)$), all performed worse than the random score in both scenarios. Regarding the other scoring functions, `greedyT_infanimals` (allocating according to the number of infected animals by herd, $I_j(t)$) had the best performance

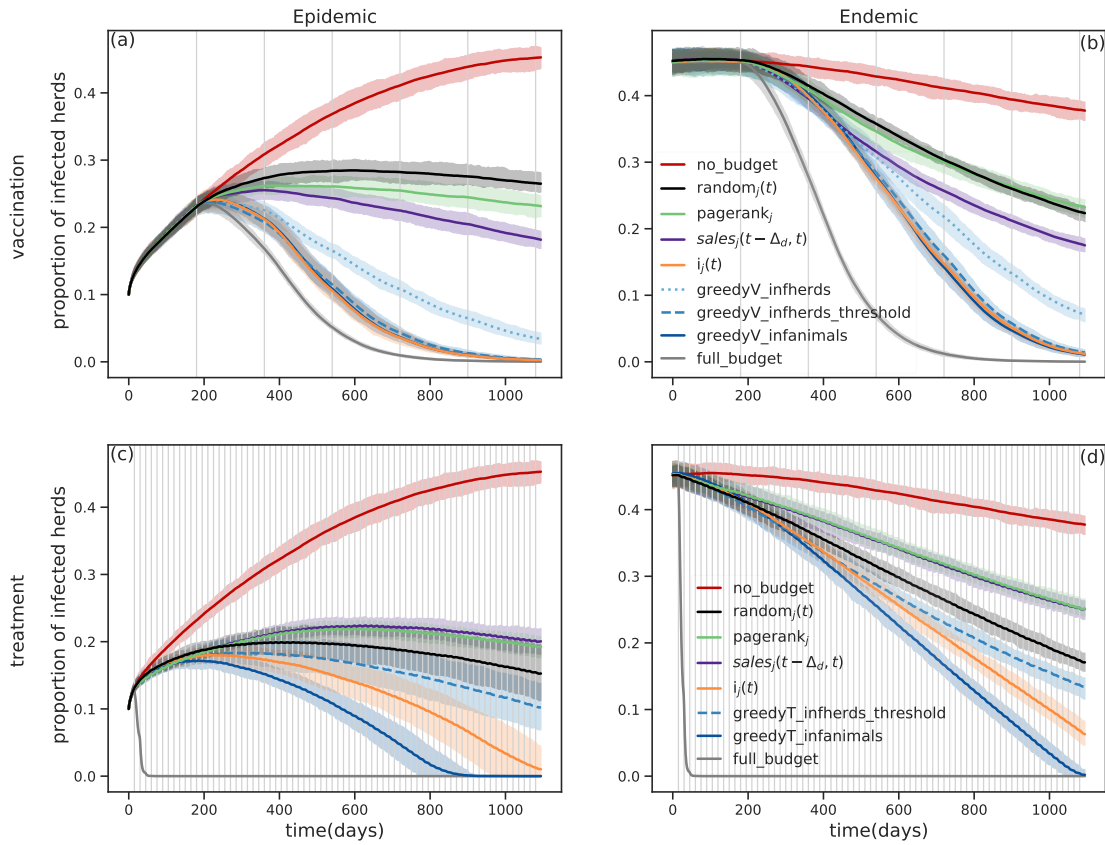
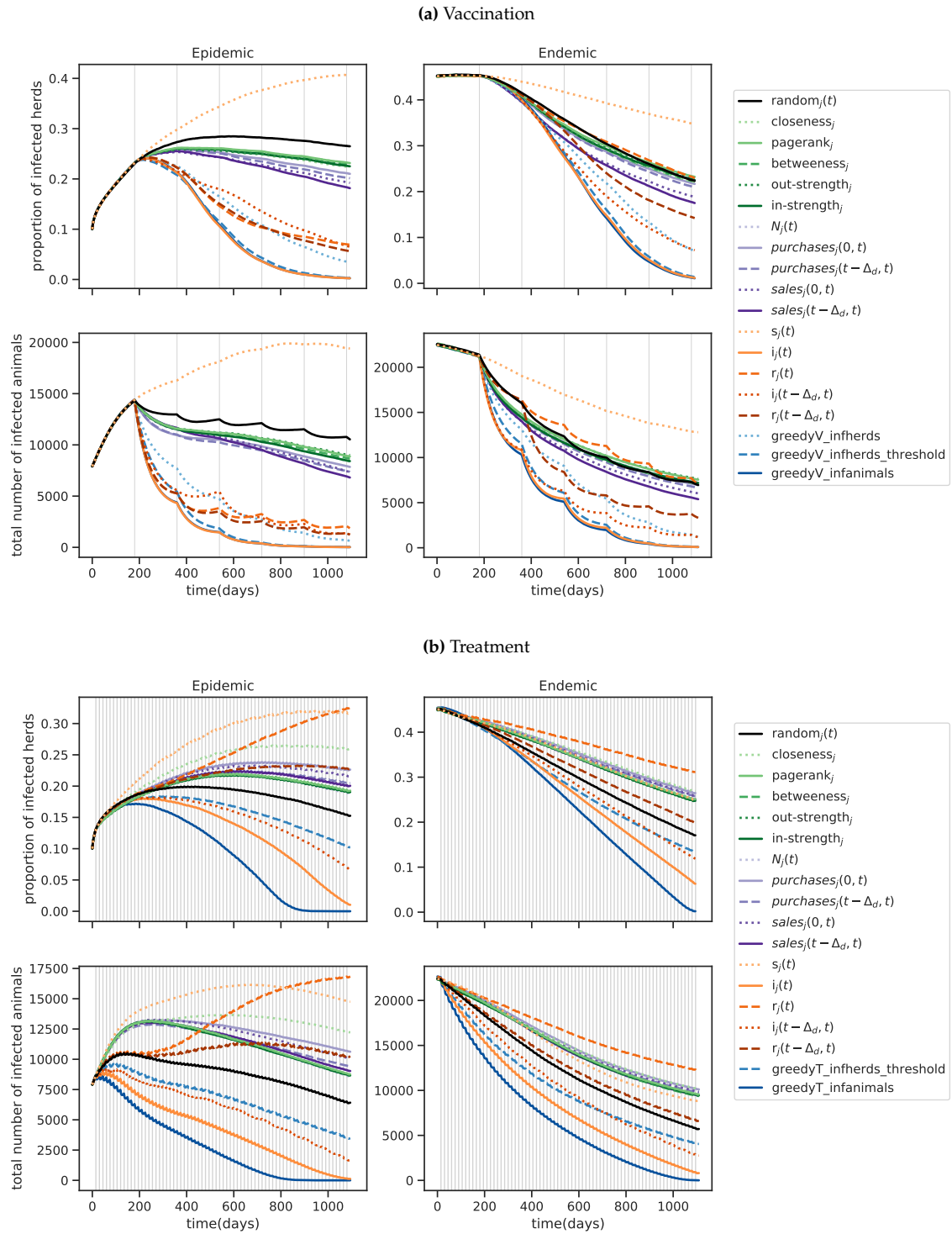


Figure 4.2: Dynamics of infection spread (proportion of infected herds) under resource allocation based on different scoring functions. Epidemic (left) and endemic (right) scenarios, for vaccination (a and b) and treatment (c and d). `no_budget` (red) represents the case where there is no resource allocated, and `full_budget` (grey) the case where the resource is not limited. For other curves, each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange), and greedy scoring functions (blue tones). For vaccination, the amount of available doses by decision time was $b_{fix} = 25\%$ of the initial total number of animals in the metapopulation. For treatment, the number of attainable herds by decision time was $b_{fix} = 25$ herds. Gray vertical lines represent resource allocation times. Mean results and 95 confidence bands over 50 runs.

in both scenarios. It was followed by the score $i_j(t)$, i.e. the proportion of infected animals by herd, though this last one did not manage to eradicate the disease before the 3 years. The next best performance was given by $i_j(t - \Delta_d, t)$, which was followed by `greedyT_inferherds_threshold`. To sum up, for the treatment allocation, only the greedy scores and two epidemiological scores ($i_j(t)$ and $i_j(t - \Delta_d, t)$) performed better than the random score. In particular `greedyT_infanimals` was the only one that eradicated the disease within the 3 years (figure 4.2).

The dynamics of raw proxies of costs can be found in figure 4.4. In particular, it is shown that the proportion of herds that were vaccinated varied according to the allocation scoring function. Indeed, topological and demographic scoring functions led to vaccinating slightly less herds than the epidemiological and the greedy scores (excluding $s_j(t)$). Unsurprisingly, $s_j(t)$ led to the highest proportion of vaccinated herds and did not waste any doses, i.e. only vaccinated herds without infected or recovered animals, while $r_j(t)$ wasted the highest proportion of available doses. For the treatment, the topological and demographic scoring functions led to the smallest number of used doses and to the highest size of target population, contrary to the epidemiological and greedy scores.



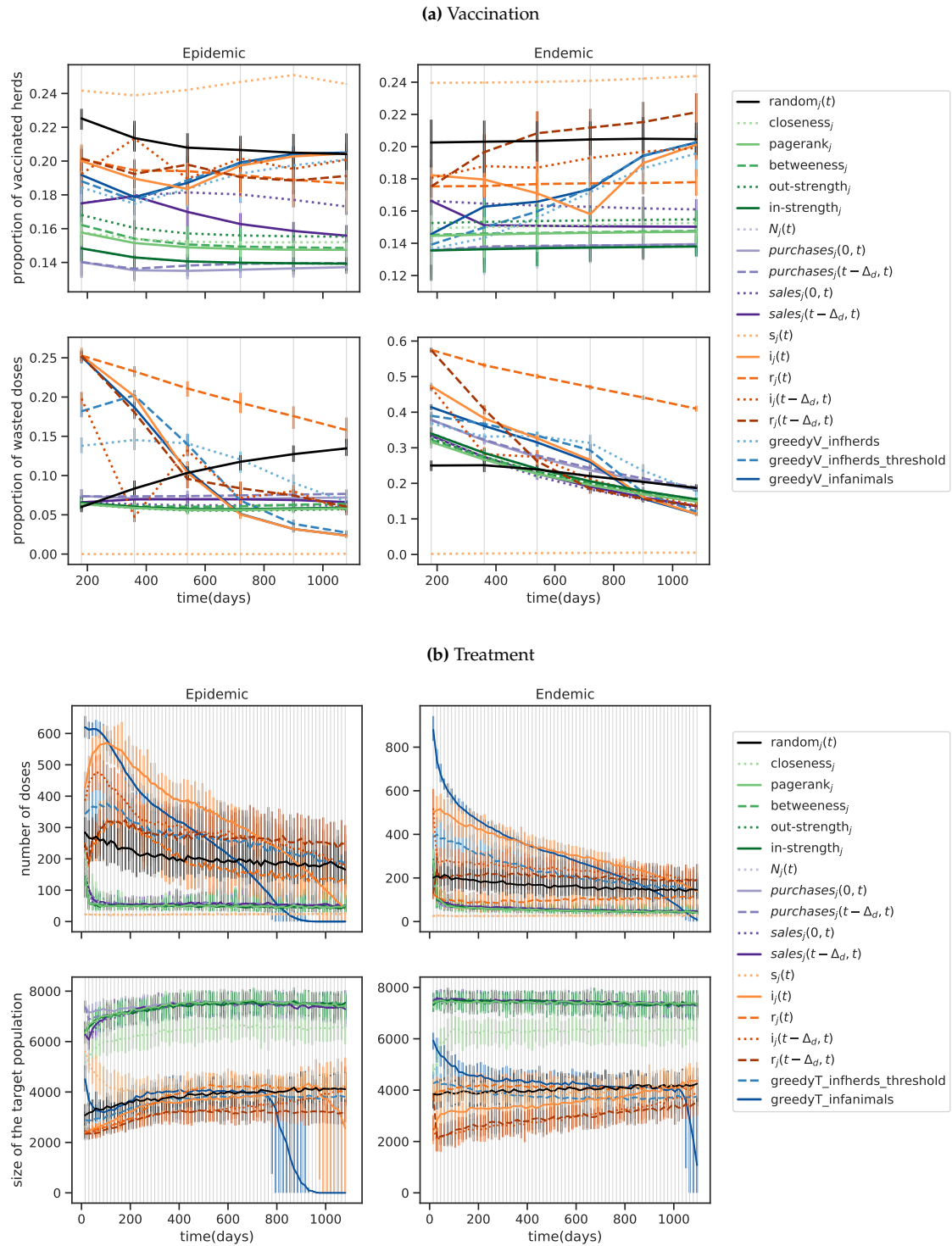


Figure 4.4: Exploration of the resource allocation by scoring function (when $\beta/\gamma = 2, 1/\gamma = 90$). Epidemic (left) and endemic (right) scenarios. Each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green tones), demographic (purple tones), epidemiological (orange tones), and greedy scoring functions (blue tones). For vaccination (a), mean over runs for the proportion of vaccinated herds and wasted doses through time, when $b_{fix} = 25\%$ of the initial number of animals in the metapopulation. For treatment (b), mean over runs for the number of used doses and size of the target population through time, when $b_{fix} = 25$ herds. Gray vertical lines represent resource allocation times. Coloured vertical bars represent the variation between the 5th and the 95th percentiles over runs. Results for 50 runs.

The Jaccard index (Levandowsky et al. 1971) measures similarity between two sets as the size of the intersection divided by the size of the union of the sets. It ranges between 0 and 1. If it is close to 1, the two sets are similar. If close to 0 they are not. Figure 4.5 shows the relationship between allocation decisions among different scoring functions at a given decision time. Both for vaccination and treatment, decisions according to the topological and demographic indicators were very similar (in terms of targeted herds) according to the Jaccard index. For vaccination, decisions according to epidemiological and greedy scoring functions were similar at the first decision time (6 months), yet this similarity diminished over time.

In figure 4.6 we measure similarity for different scoring functions between decisions taken over time according to the given scoring function, i.e. how much the allocation decisions taken according to a fixed scoring function are similar over decision times. Regarding decisions through time for a given scoring function, topological functions tended to allocate the resource to the same herds over time. This was also the case for demographic functions, except for the sales scoring functions in vaccination, for which the first decisions were less and less similar to decisions at the following decision times. On the contrary, for $s_j(t)$ the similarity between consecutive vaccination decisions seemed to increase over time. And for each of the other epidemiological and greedy scoring functions, vaccination decisions were in general less similar over time. Treatment decisions according to each epidemiological and greedy scoring function were very different over time as long as the disease was not eradicated, except for decisions according to $r_j(t)$.

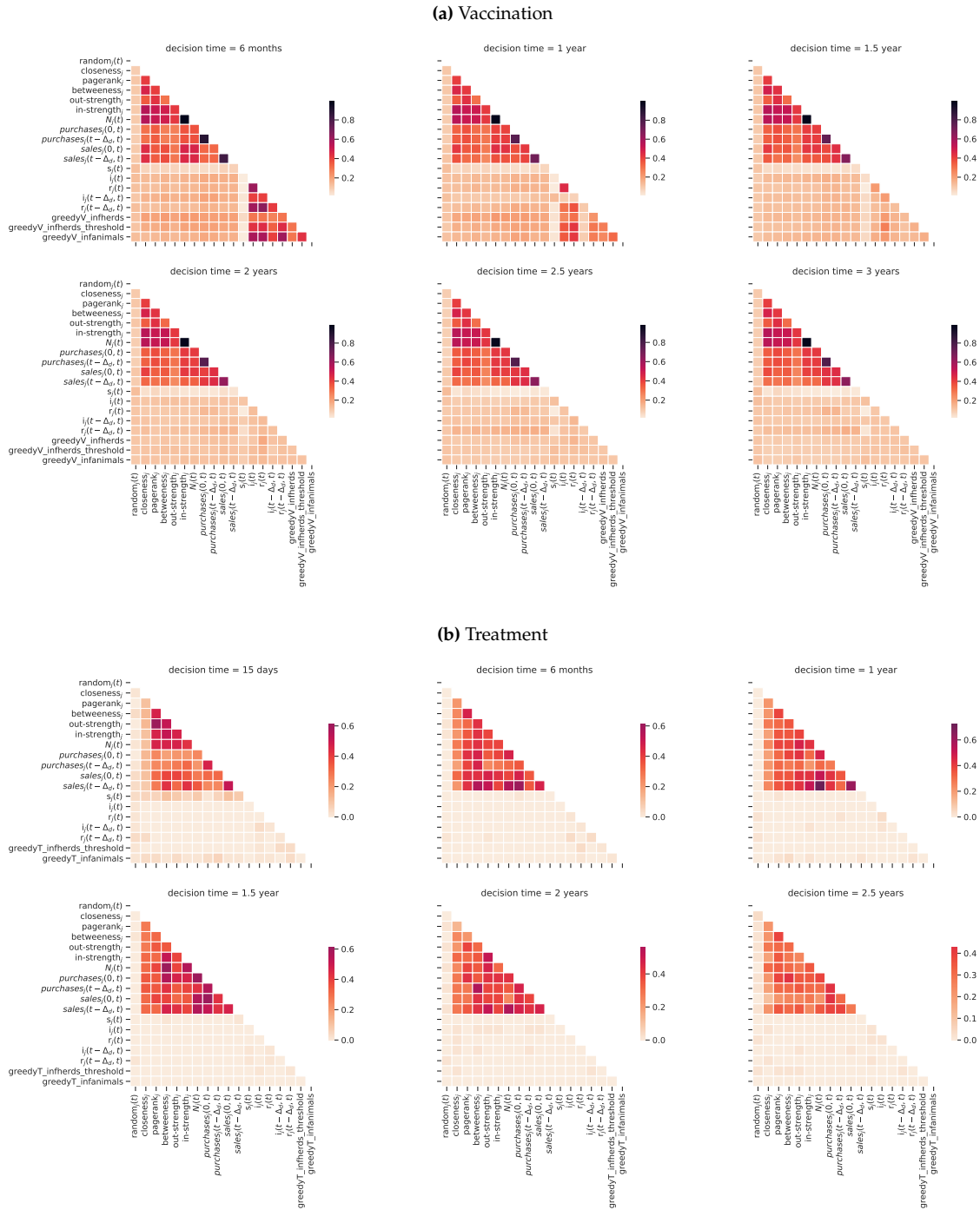


Figure 4.5: Jaccard index of similarity at different decision times between decisions taken according to each pair of scoring functions when $\beta/\gamma = 2, 1/\gamma = 90$. Results for one run (for illustrative purposes) of the epidemic scenario. For (a) vaccination with $b_{fix} = 25\%$ of the initial number of animals in the metapopulation, and (b) treatment with $b_{fix} = 25$ herds. The colour scale is the same for all sub-panels (centered at 0.5).

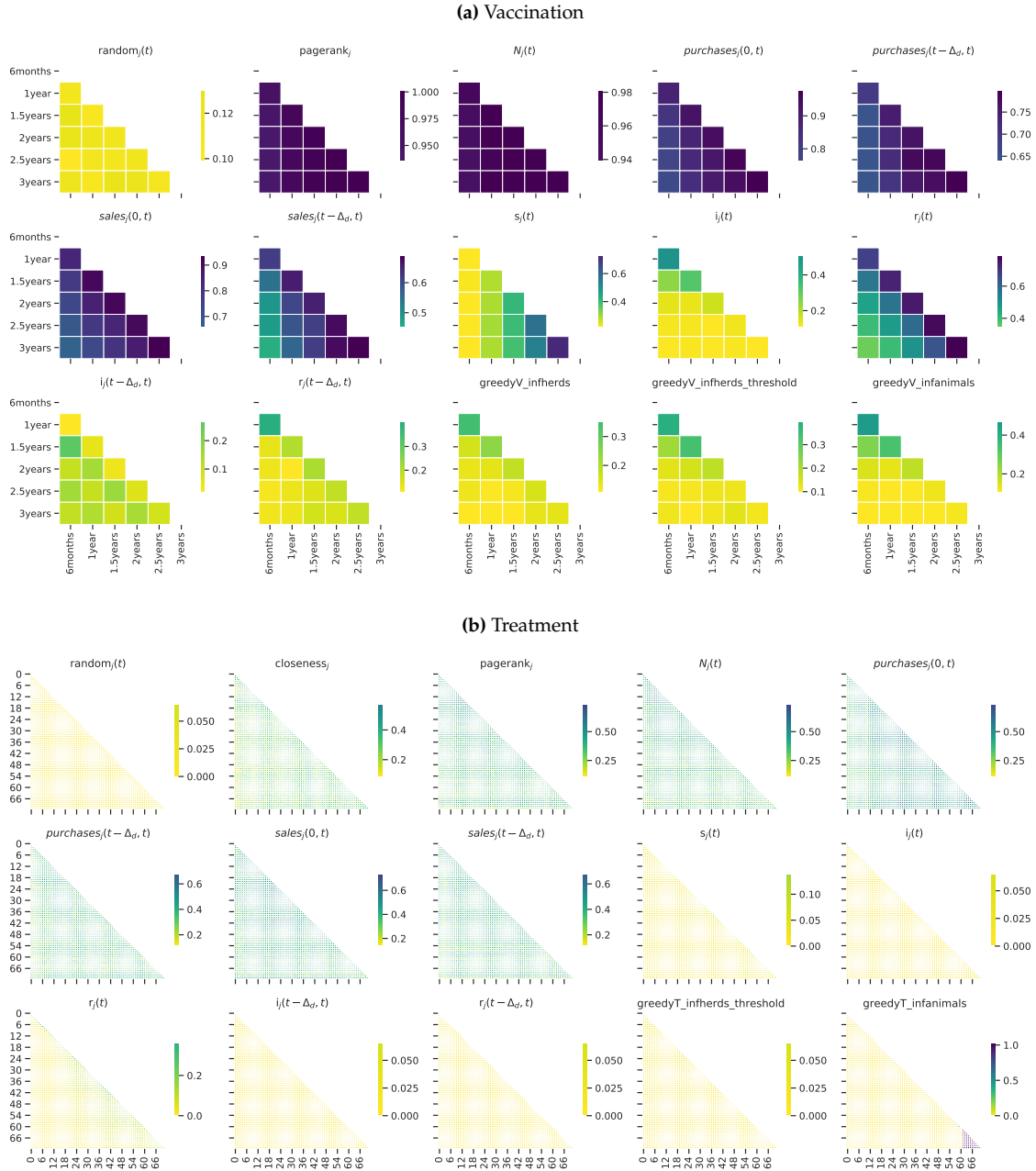


Figure 4.6: Jaccard index of similarity for different scoring functions between decisions taken over time according to the given scoring function when $\beta/\gamma = 2, 1/\gamma = 90$. Results for for one run (for illustrative purposes) of the epidemic scenario. For (a) vaccination with $b_{fix} = 25\%$ of the initial number of animals in the metapopulation, and (b) treatment with $b_{fix} = 25$ herds. The colour scale is the same for all sub-panels (centered at 0.5).

4.5.2.2 Percolation analysis results

Figure 4.7 shows results of the percolation analysis in the endemic scenario for each type of resource, using a selected subset of scoring functions. Results in the epidemic scenario can be found in figure 4.8. We ran this analysis using all the epidemiological and greedy scores, since the infection-related dynamics results in section 4.5.2.1 were quite different for the scores within each of these groups. In contrast, because the results for the topological scores were very similar, as were the results of the demographic scores, we considered only one of each type: $\text{sales}_j(t - \Delta_d, t)$ for the demographic scoring functions and pagerank_j for the topological scores. This figure (also figure 4.7) confirms the main observations made in section 4.5.2.1: certainly, the best-performing scoring-functions for reducing disease prevalence, for almost every quantity of resource that we tested, were the greedy scores along with some epidemiological scores.

For vaccination, figure 4.7(a) shows that when the proportion of vaccinated herds was at least 0.10 of the initial number of animals in the metapopulation, the scoring functions that led to the lowest final proportion of infected herds were $\text{greedyV_inf_animals}$, $\text{greedyV_infherds_threshold}$ and $i_j(t)$. They were followed by the greedyV_infherds scoring function. The topological and demographic scoring functions pagerank_j and $\text{sales}_j(t - \Delta_d, t)$ were just as good as the greedy scores only when the available quantity of resource was very low (particularly when the proportion of vaccinated herds was less than 0.05), but did not perform well as this proportion was higher. The $r_j(t - \Delta_d, t)$ and $i_j(t - \Delta_d, t)$ scoring functions had a bad performance for such low levels of vaccinated herds, performed better for medium levels of vaccination, and performed worse when the proportion of vaccinated herds was higher than 0.2. In particular, $i_j(t - \Delta_d, t)$ was better than greedyV_infherds when the proportion of vaccinated herds was less than 0.2. On the contrary, as the proportion of vaccinated herds was higher, $r_j(t)$ performs better. Indeed, when less than 25% of herds were vaccinated, it was just as bad as the random scoring function, which was the second worst score after $s_j(t)$. But for a proportion of vaccinated herds higher than 0.45, its performance was close to the one of the best-performing greedy scores. Even when this proportion was more than 0.6, only $r_j(t)$, $i_j(t)$ and the greedy scores arrived to eradicate the disease.

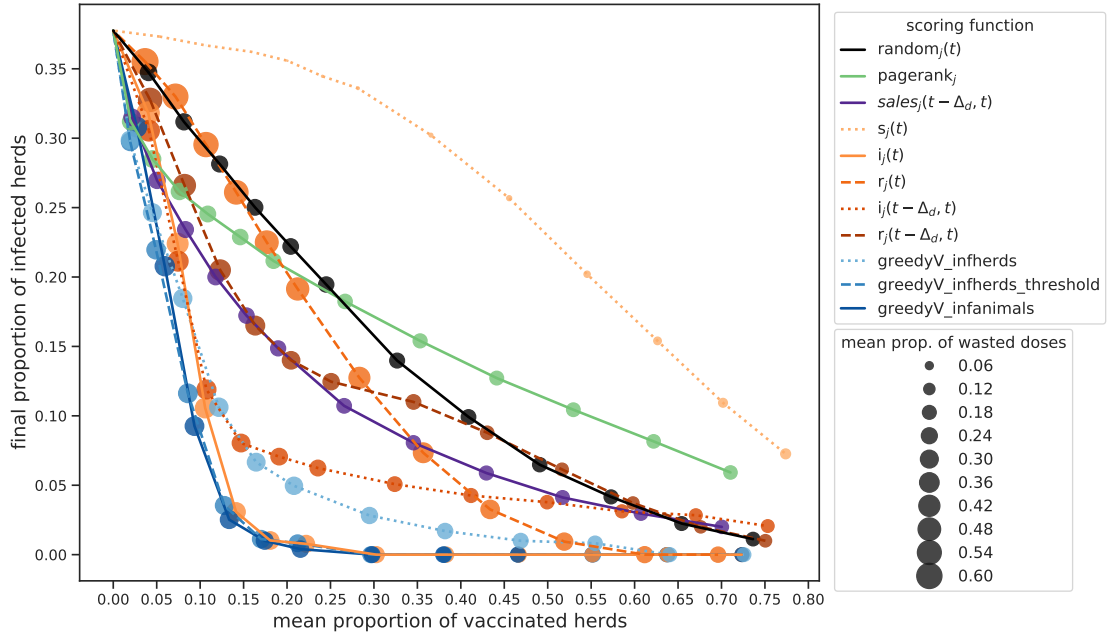
Regarding the mean proportion of wasted doses (over time and over runs), it was surprisingly almost zero when the vaccine allocation was done using $s_j(t)$ as criterion, irrespective of the proportion of vaccinated herds. The other scoring functions led to quite similar levels of wasted doses, except for $r_j(t)$ and $r_j(t - \Delta_d, t)$, which led to the highest proportion of wasted doses, particularly when the proportion of vaccinated herds was lower than 0.15.

In figure 4.7(b), we observe that when the resource to allocate was a treatment, $\text{greedyT_infanimals}$ managed to eradicate the disease when it was possible to treat at least 25 herds per decision time-step, i.e. 0.5% of the total number of herds. As for $i_j(t)$, it could eradicate the disease when this percentage was higher than 0.6%, and $i_j(t - \Delta_d, t)$ when it was at least 0.7% herds. When b_{fix} was equal to 0.9% of the total number of herds, $\text{greedyT_infherds_threshold}$, the random scoring function and $r_j(t - \Delta_d, t)$ also eradicated the disease, in that order. The other epidemiological scoring functions and the topological and demographic ones performed worse than the random score for all values of b_{fix} . In particular, $\text{sales}_j(t - \Delta_d, t)$ only eradicated the disease if b_{fix} was higher than 1% of the total number of herds, $s_j(t)$ could only do it when this percentage was higher than 1.2%, and the other two scoring functions ($r_j(t)$ and pagerank_j) eradicated it when it was possible to treat at least 1.3% and 1.4% of the total number of herds, respectively.

As for the number of used doses of treatment, they were in general lower for bad-performing scoring functions, and higher for those that performed the best. The exception was $r_j(t)$, which performed badly and led to a high number of used doses. When b_{fix} was high, it used even more doses in average than the best scoring functions for reducing disease prevalence.

Results in the epidemic scenario (figure 4.8) were quite similar. Yet, for the treatment, only $\text{greedyT_infanimals}$, $i_j(t)$ and $i_j(t - \Delta_d, t)$ were always better than the random score.

(a) Vaccination



(b) Treatment

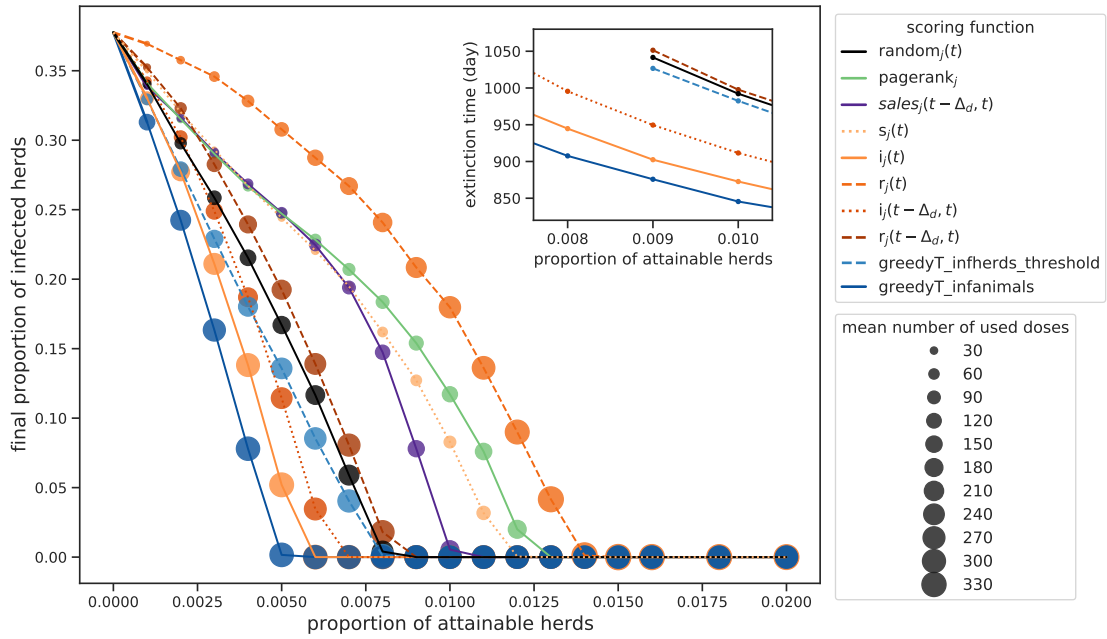


Figure 4.7: Percolation experiments results in the endemic scenario. Each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange tones), and greedy scoring functions (blue tones). For (a) vaccination, final (after 3 years) mean (over runs) proportion of infected herds as a function of the mean (over time and over runs) proportion of vaccinated herds. Each point corresponds to a value of b_{fix} , and its size represents the mean (over time and over runs) proportion of wasted doses. For (b) treatment, final (after 3 years) mean (over runs) proportion of infected herds as a function of the proportion of attainable herds at each decision time. Each point corresponds to a value of b_{fix} , and its size represents the mean (over time and over runs) number of used doses. The inset shows the mean (over runs) extinction time with each scoring function, for a restricted range of the proportion of attainable herds. Results for 50 runs.

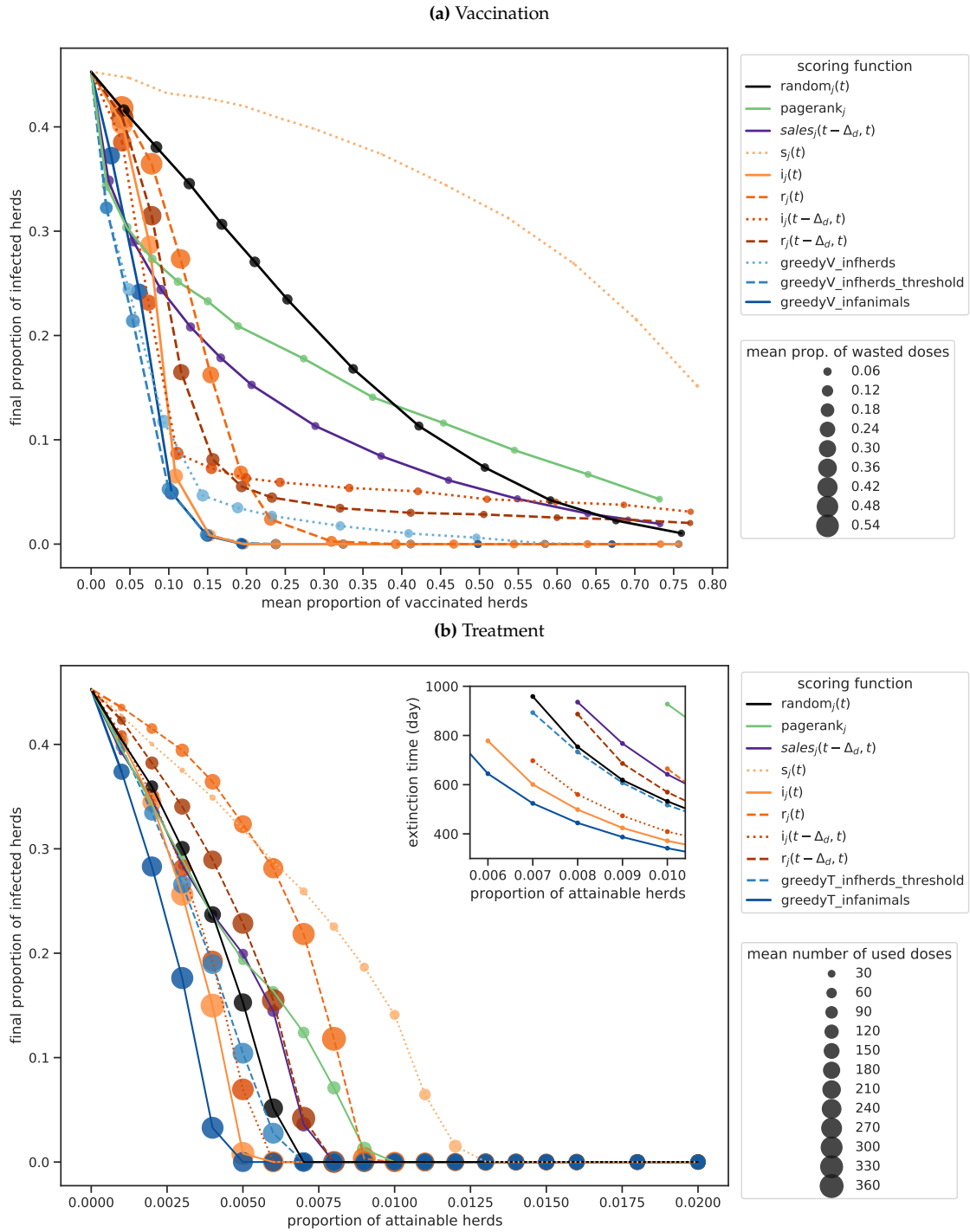


Figure 4.8: Percolation experiments results (when $\beta/\gamma = 2, 1/\gamma = 90$) in the epidemic scenario. Each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange tones), and greedy scoring functions (blue tones). For (a) vaccination, final (after 3 years) mean (over runs) proportion of infected herds as a function of the mean (over time and over runs) proportion of vaccinated herds. Each point corresponds to a value of b_{fix} , and its size represents the mean (over time and over runs) proportion of wasted doses. For (b) treatment, final (after 3 years) mean (over runs) proportion of infected herds as a function of the proportion of attainable herds at each decision time. Each point corresponds to a value of b_{fix} , and its size represents the mean (over time and over runs) number of used doses. The inset shows the mean (over runs) extinction time with each scoring function, for a restricted range of the proportion of attainable herds. Results for 50 runs.

The greedyT_inferherds_threshold indicator was better than the random score only when b_{fix} was less than 0.8% of the total number of herds. Additionally, figure 4.9 shows that the variability of the percolation results were moderate between runs, with no or little overlapping 90% confidence intervals.

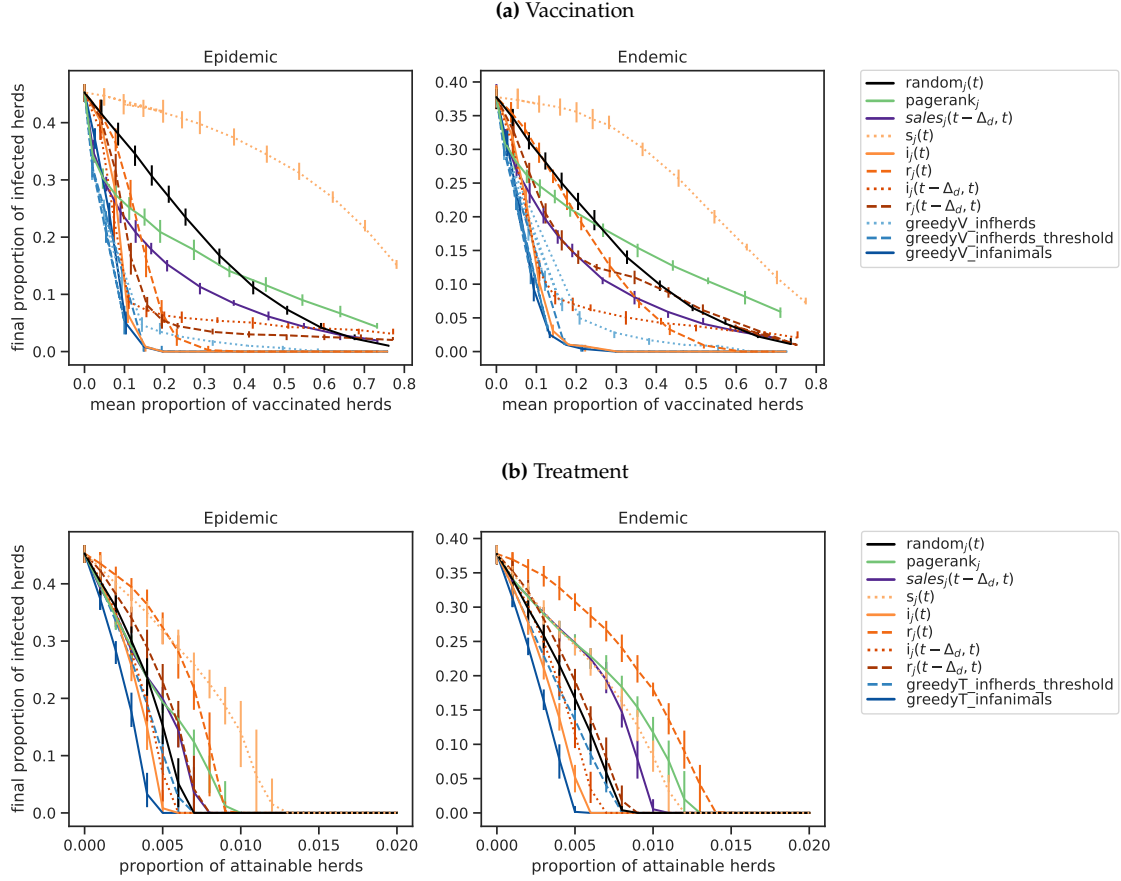


Figure 4.9: Variation due to intrinsic stochasticity of the model in the percolation experiments results when $\beta/\gamma = 2, 1/\gamma = 90$. Epidemic (left) and endemic (right) scenarios. Each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange tones), and greedy scoring functions (blue tones). For (a) vaccination, final (after 3 years) mean (over runs) proportion of infected herds as a function of the mean (over time and over runs) proportion of vaccinated herds. For (b) treatment, final (after 3 years) mean (over runs) proportion of infected herds as a function of the proportion of attainable herds at each decision time. Coloured vertical bars represent the variation between the 5th and the 95th percentiles over runs. Results for 50 runs.

In addition, figure 4.10 shows that when vaccine efficacy is decreased, the loss of performance is not the same for all scores. In particular, the scores that are not the best but perform rather well (epidemiological scores and greedyV_inferherds) are the ones for which performance is most depreciated. This results in $sales_j(t - \Delta_d, t)$ performing better than $i_j(t - \Delta_d, t)$ when the average proportion of vaccinated herds was 55% and vaccine efficacy was 0.9 or less. Yet, despite the overall loss in efficiency, two of the greedy scores and one epidemiological score ($i_j(t)$) still performed the best.

Finally, figures 4.11 and 4.12 show that a disease with higher early peak and smaller infection duration ($\beta/\gamma = 4, 1/\gamma = 30$ days) spreads and fades out rapidly at the intra-herd level. At the same time, recurrent outbreaks are observed which are likely due to reintroduction of infected animals by trade. Figure 4.13 shows that, in such a case, allocating vaccines according to the historic sales, in particular according to $sales_j(t - \Delta_d, t)$, was among the best strategies in both the epidemic and endemic scenarios.

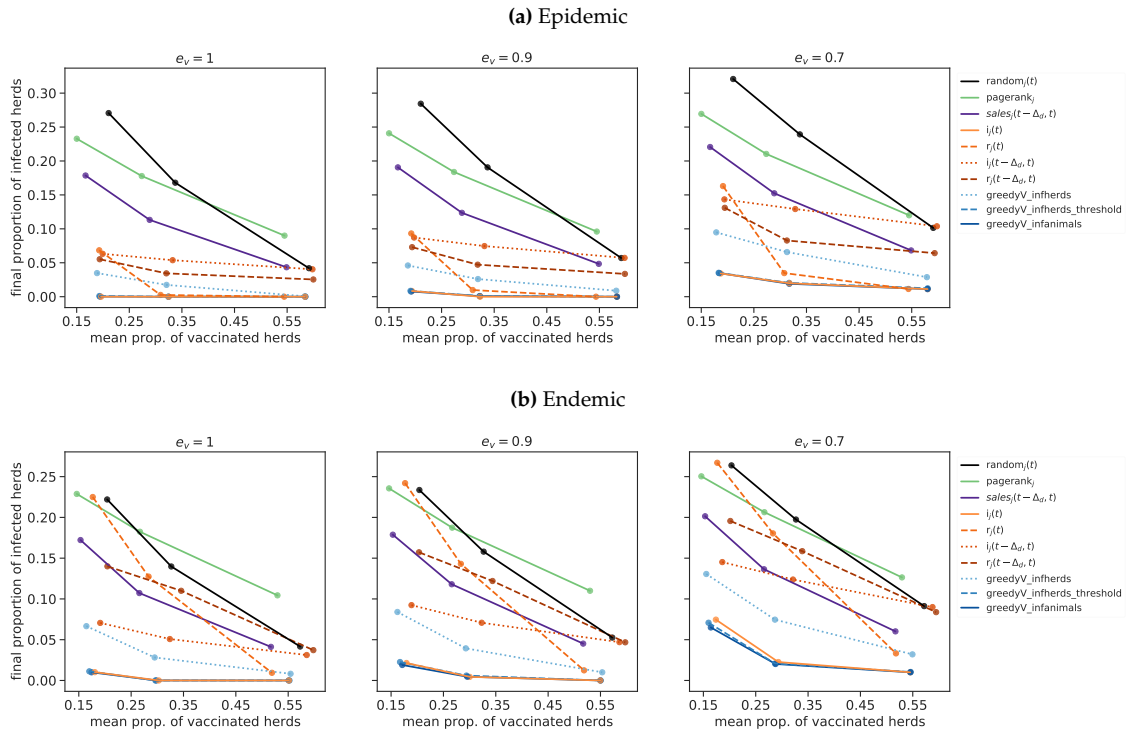


Figure 4.10: Percolation experiments results (when $\beta/\gamma = 2, 1/\gamma = 90$) for vaccination in the epidemic (a) and the endemic (b) scenario, with three values of vaccine efficacy. Each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange tones), and greedy scoring functions (blue tones). Final (after 3 years) mean (over runs) proportion of infected herds as a function of the mean (over time and over runs) proportion of vaccinated herds. Each point corresponds to a value of b_{fix} . Results for 50 runs.

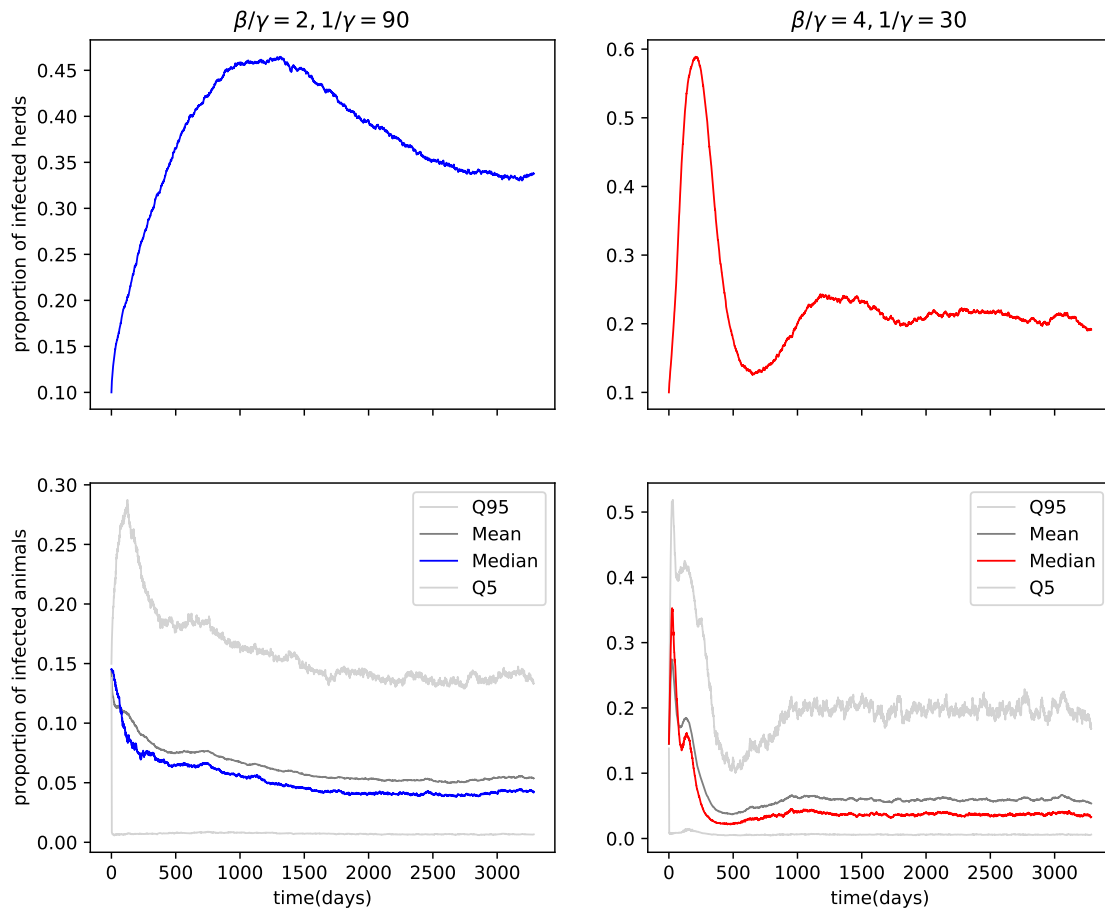


Figure 4.11: Temporal epidemic dynamics over 9 years for two sets of epidemic parameters: $\beta/\gamma = 2, 1/\gamma = 90$, (left), and $\beta/\gamma = 4, 1/\gamma = 30$, (right). For each set, inter-herd prevalence (above), and intra-herd prevalence for infected herds (below): mean and median, 5th and 95th percentiles over infected herds. Results for one run for illustrative purposes.

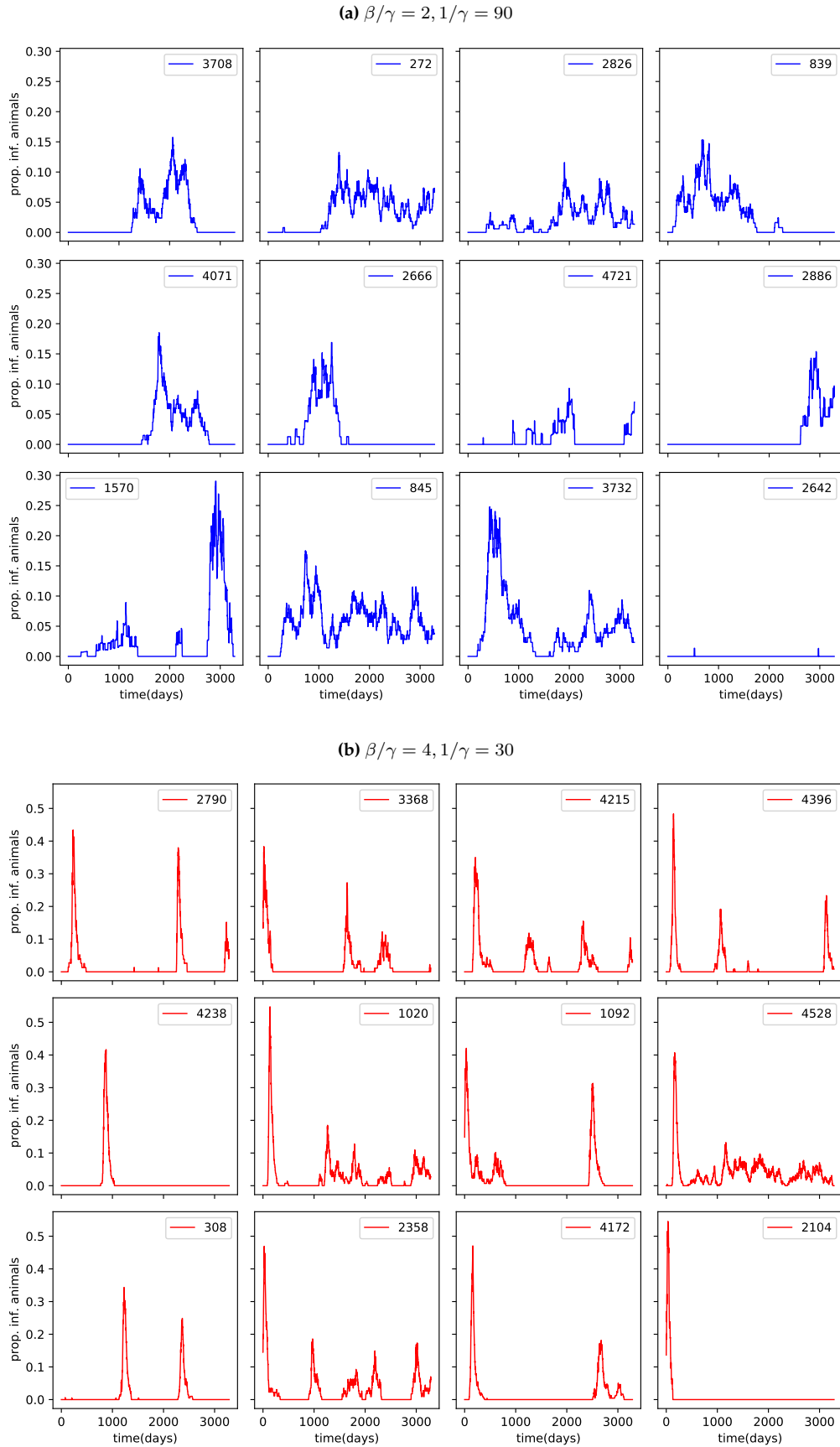


Figure 4.12: Proportion of infected animals by herd, for a subset of 36 herds infected at least once over 9 years, when (a) $\beta/\gamma = 2, 1/\gamma = 90$, and (b) $\beta/\gamma = 4, 1/\gamma = 30$. Results for one run (for illustrative purposes).

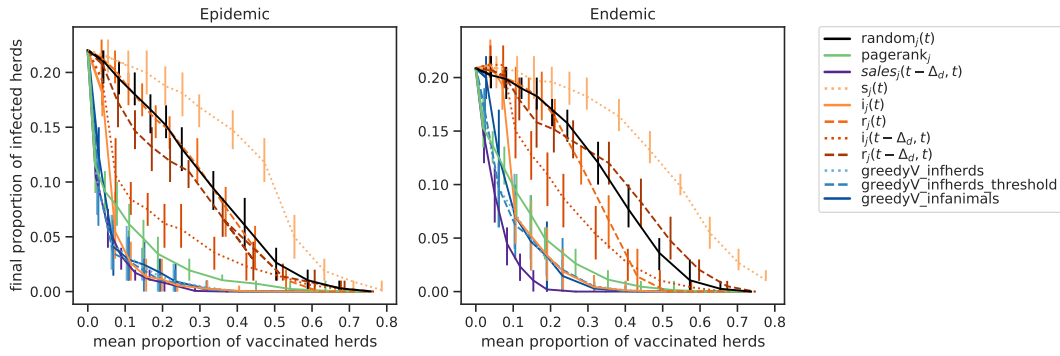


Figure 4.13: Percolation experiments results (when $\beta/\gamma = 4, 1/\gamma = 30$) for vaccination, with variation due to intrinsic stochasticity of the model. Epidemic (left) and endemic (right) scenarios. Each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange tones), and greedy scoring functions (blue tones). Final (after 3 years) mean (over runs) proportion of infected herds as a function of the mean (over time and over runs) proportion of vaccinated herds. Coloured vertical bars represent the variation between the 5th and the 95th percentiles over runs. Results for 50 runs.

4.6 Discussion

To control an infectious disease that spreads in a metapopulation network, allocating a limited resource is a fundamental yet difficult question, especially for large networks. In this study, we considered this resource allocation problem for a livestock disease that spreads over a large animal trade network, where the intra-herd infection and demographic dynamics was specified as an SIR stochastic model taking into account animal movements and demography.

The problem of resource allocation in networks had been previously addressed from several perspectives such as optimal control (Bussell et al. 2019) and reinforcement learning (Probert et al. 2019), but mostly for networks where each node is an individual (Wijayanto et al. 2019), or where the network is rather small (Rowthorn et al. 2009). Yet, in the context of a very large network these methods lack of scalability for tracking the optimal solution (Scaman et al. 2015). In this work, we chose to concentrate on strategies based on scoring functions, heuristics and optimised, which consist in ordering the nodes of the network according to their score and allocating the resource to the top of the ranking, up to the limit given by the available resource.

First, following the greedy approach in Scaman et al. 2015, we provided new analytic scoring functions for controlling the disease spread over the animal metapopulation network by optimising approximated objective functions. The scoring functions we derived depend on the infection-related state of the herd, and some are also dependent on the topology of the metapopulation network. They differ according to the objective of the control (minimising the number of infected animals vs minimising the number of infected herds) and the type of available resource (a protective vaccine or a treatment that reduces the infectious period). Meanwhile, most similar existing approaches for other population structures derive strategies solely for distributing a vaccine (Tanner et al. 2008; Zhan et al. 2017), or are concerned with only one objective to be optimised (Forster et al. 2007; Wijayanto et al. 2019).

Through intensive simulations, we observed that these analytically obtained scoring functions can be optimal for reducing disease prevalence in the metapopulation, though this is not always the case. For example, even if `greedyV_inanimals` (the score for greedily minimising a function on the total number of infected animals in the metapopulation) showed the greatest reduction in disease prevalence through vaccination, allocating vaccines according to the proportion of infected animals by herd, $i_j(t)$, can be just as good when the number of available vaccine doses equalled 25% of the initial number of animals in the metapopulation (figure 4.2). Even more, we observed that this was the case as long as the number of available doses was more than 15% of

the initial number of animals in the metapopulation (figure 4.7(a)).

For the treatment, most of the scoring functions, in particular topological and demographic ones, were counterproductive in the sense that they performed worse than randomly allocating the resource among the infected herds (figures 4.2 and 4.7(b)). We explain this by the fact that infected herds which were central in the network were not the most infected ones (in terms of the proportion of infected animals). Indeed, figure 4.4 shows that the random allocation among infected herds also targeted herds with many infected animals, while the scoring functions that performed badly only targeted high sized infected herds but generally with few infected animals.

Furthermore, we noticed that irrespective of the resource type, the optimised scoring function for an objective on the number of infected herds was outperformed by the optimised scoring function for an objective on the number of infected animals. Even if for vaccination, a slightly modified version of greedyV_inherds provided results almost as good as the ones of greedyV_inanimals (figures 4.2 and 4.7(a)), this was not the case for the treatment. Indeed, the allocation implemented using the greedyT_inanimals scoring function, i.e. the number of infected animals per herd, yielded undoubtedly the best results (figures 4.2 and 4.7(b)). This is probably due to the fact that the scoring functions for minimising an objective on the number of infected herds only focus on the fast recovery of slightly infected herds ($0 < I_j(t) < 20$), for vaccination, or on avoiding that completely healthy herds receive infectious animals ($I_j(t) > 0$), for the treatment. Although this is the best way to have a small incremental number of infected herds from one instant to another according to these scoring functions, it does not take into account new animal infections which only occur at the intra-herd level once the herd is infected. Our interpretation is that a scoring function obtained with the greedy approach (which consists in focusing on the short-term behaviour of the objective function) performs better for limiting the disease spread if the objective function it is built on directly captures the intra-herd aspect of the disease dynamics. Hence, minimising a first-order approximation for an objective that directly concerns the number of infected animals can provide more performing scoring functions than that obtained by minimising a higher-order approximation for an objective that does not.

Numerical investigations also allowed to evidence that intra-herd health information can be crucial for optimally controlling the disease spread of a slowly-spreading disease such as the one considered in this work. In most combinations, given by the resource's type and available quantity, even if topology-based scoring functions managed to limit disease spread, scoring functions based on the infection-related state of herds performed better (figures 4.7 and 4.8). This observation can give some light into why most control strategies implemented in real systems might fail to eradicate livestock diseases in areas that lack this kind of information.

A final interesting remark is that the best scoring functions for reducing disease prevalence can induce a higher number of wasted vaccine doses or a higher number of used treatment doses compared to other scores that performed less well for controlling the disease spread (figures 4.7, 4.8 and 4.9). However, for vaccination, they are not necessarily those that vaccinate the highest proportion of herds.

To our knowledge, our work is one of the few studies that explores dynamic resource allocation in a metapopulation network, for many allocation scoring functions (16 heuristics and at least 2 optimised scores by measure) while varying the available quantity of the resource in two different scenarios. Despite the fact that the performances of the scores could have been different if the network had been static or not scale-free, score-based resource allocation can be a relevant approach for controlling pathogen spread in other cases, as the complexity of the problem is mainly due to the large dimension of the network. Furthermore, we stress that the scores we found by optimisation using the greedy approach are the same irrespective of the network topology as they are based on fixed trade parameters representing the static aggregated network. Indeed, they are built on a static view of the network, irrespective if it is actually static or dynamic (as it is in our study). We believe it would be possible to use the same approach to obtain optimised scores that take into account the dynamic nature of the network, i.e. scores that are function of time-dependent trade parameters, although this requires a new formal analysis. Of course, these new scores could be of a different form than the ones we found and hence their performance might also be different.

We assessed the robustness of the results regarding vaccine allocation by considering realistic values for the vaccine efficacy. This showed a limited impact on the relative performance of the different scoring functions (figure 4.10). We note that although, in reality, vaccines are rarely perfect effective, and also take some time to be effective, it did not seem straightforward to determine an appropriate time frame for the vaccine to have an effect. More importantly, it seems unlikely that a slight delay in the effect of the vaccine would have a significant impact on our results, given that we were considering a pathogen that spreads rather slowly.

Regarding the limitations of our work, we emphasise that in the context of a fast-spreading pathogen, the current framework is not really appropriate. In such a case, other decision factors should be taken into account. For example, if the disease is zoonotic or has a strong economic impact, the social planner may consider more radical options, such as mass culling. In this case, the question of resource allocation thus becomes irrelevant. In particular, we showed that for a disease with higher early peak and shorter infection duration ($\beta/\gamma = 4$, $1/\gamma = 30$ days) an intra-herd epidemic extinguishes before a new disease introduction occurs (figures 4.11 and 4.12). Therefore, it is not surprising that vaccinating herds that sell many animals appears as a good strategy for limiting disease propagation (figure 4.13). So, the resource allocation problem seems more straightforward in such a scenario and does not necessarily require an optimisation procedure.

Additionally, a parameter that could impact our conclusions is the decision step, Δ_d , for which we considered a fixed heuristic value. Although the assumption of regular vaccination decisions defined by the duration of the protection conferred by the vaccine appears to be a realistic hypothesis relative to field practice, a more versatile assumption could be considered to determine the frequency of allocation decisions. Indeed, the decision step could be determined in an adaptive manner by the social planner, for example by taking into account the stability in disease prevalence, or some external input such as the farmers' demand for accelerating resource allocation. A second option would be to determine the decision step by optimisation. Yet, these are essentially different problems from the one we addressed in this chapter: determining when to allocate instead of where to allocate a limited resource. Optimising both aspects at the same time is a more complex problem that to the best of our knowledge has only been addressed by heuristic approaches (e.g. Venkatramanan et al. 2019). In particular, it does not seem straightforward to address with the approach of this study.

Finally, we stress that the performance of the epidemiological and the greedy scoring functions can be counterbalanced by their difficulty of access. Indeed, having updated knowledge on the epidemiological state of all the herds of the network is a strong hypothesis in real-life, as this kind of information can be hard to gather for most livestock diseases (Carslake et al. 2011). For example, a scoring function calculated as the increment in the proportion of infected animals in a herd over a certain period can be observed through changes in the herd's seroprevalence between two time points, which incurs into increased logistics, can be observed with error and not in real time. Furthermore, having such updated and detailed health-related information can be costly, and this cost should be taken into account in the constrained optimisation problem for the allocation. Among the possible perspectives of this work, the previous point opens an important one: combining scoring functions for improved performances, and above all for yielding a scoring function that can be useful in practice. This could be achieved, for example, through the (linear) combination of scoring functions, or through the selection of herds at the top of the ranking given by several scoring functions that do not allocate the resource in a similar way. Additionally, for cases when the value of the score is the same for many herds, the allocation could be done using a second scoring function that would take different allocation decisions. As a second, more methodological perspective, the greedy scoring functions built on first-order approximations could eventually be constructed using higher-order approximations. This could lead to analytic scores that also depend on network topology, and could improve their performance. Although it has to be stressed that they were already among the best-performing scores for reducing disease prevalence in all cases. Finally, despite the fact that we focused on a protective vaccine and a treatment that increases the recovery rate, other types of resources could be studied with the same approach. However, it might not be straightforward to derive the analytic

expression of the greedy scoring function in such cases. For example, the effect of the restriction of animal movements, which is a relevant control measure in this context, lies on the connections of the herd rather than on the intra-herd level, which could further complicate the derivation of the scoring function.

4.7 Appendix: Analytical derivation of the greedy scoring functions

We adapted the greedy approach proposed in Scaman et al. 2015, which consists in minimising a short-term approximation of the objective function, for the metapopulation framework. This approach finds locally (in time) optimal decisions. Let us note $X(t)$ the epidemic state of the network at time t . In the metapopulation framework, $X(t)$ is a $J \times 3$ matrix, where each row contains the number of susceptible, infected and recovered animals by each herd at time t . Since the strategy has no impact on the initial state, and since the process is Markovian, the problem (under the same constraint) is equivalent to:

$$\begin{aligned} & \min_A \int_{t=0}^{\infty} e^{-at} \mathbb{E} [J_I^A(t) | X(0) = X] dt \quad \forall X \\ \stackrel{(Markov)}{\iff} & \min_A \int_{u=0}^{\infty} e^{-au} \mathbb{E} [J_I^A(t+u) | X(t) = X] du \quad \forall X, \forall t \geq 0 \\ \iff & \min_A \int_{u=0}^{\infty} e^{-au} \phi_{t,X}(u) du \quad \forall X, \forall t \geq 0 \end{aligned}$$

where $X = ((S, I, R))_{J \times 3}$ is a possible epidemic state matrix of the network, and $\phi_{t,X}(u) := \mathbb{E} [J_I^A(t+u) | X(t) = X]$. The idea of the greedy approach is to optimise an approximation of the short term behaviour of $\phi_{t,X}(u)$ in order to find a scoring function which at each time gives a locally optimal decision. For this we use Taylor expansion evaluated in 0, to obtain an approximation of the function to minimise.

Using a second-order Taylor expansion of $\phi_{t,X}(u)$ in 0 yields:

$$\begin{aligned} & \int_{u=0}^{\infty} e^{-au} \phi_{t,X}(u) du \\ &= \int_{u=0}^{\infty} e^{-au} \left[\phi_{t,X}(0) + u\phi'_{t,X}(0) + \frac{u^2}{2}\phi''_{t,X}(0) + R(u) \right] du \\ &= \phi_{t,X}(0) \int_{u=0}^{\infty} e^{-au} du + \phi'_{t,X}(0) \int_{u=0}^{\infty} ue^{-au} du + \frac{\phi''_{t,X}(0)}{2} \int_{u=0}^{\infty} u^2 e^{-au} du + \int_{u=0}^{\infty} R(u) e^{-au} du \end{aligned}$$

where $R(u) = O(u^3)$. Since $R(u) = O(u^3)$, for a positive constant C :

$$\begin{aligned} \int_{u=0}^{\infty} -Cu^3 e^{-au} du &\leq \int_{u=0}^{\infty} R(u) e^{-au} du \leq \int_{u=0}^{\infty} Cu^3 e^{-au} du \\ -C \frac{1}{a^4} &\leq \int_{u=0}^{\infty} R(u) e^{-au} du \leq C \frac{1}{a^4} \end{aligned}$$

using successive integration by parts. So $\int_{u=0}^{\infty} R(u) e^{-au} du = O(\frac{1}{a^4})$. Therefore:

$$\begin{aligned}
& \int_{u=0}^{\infty} e^{-au} \phi_{t,X}(u) du \\
&= \phi_{t,X}(0) \int_{u=0}^{\infty} e^{-au} du + \phi'_{t,X}(0) \int_{u=0}^{\infty} u e^{-au} du + \frac{\phi''_{t,X}(0)}{2} \int_{u=0}^{\infty} u^2 e^{-au} du + O\left(\frac{1}{a^4}\right) \\
&= \frac{1}{a} \phi_{t,X}(0) + \frac{1}{a^2} \phi'_{t,X}(0) + \frac{1}{a^3} \phi''_{t,X}(0) + O\left(\frac{1}{a^4}\right)
\end{aligned}$$

Then, we must minimise

$$\frac{1}{a} \phi_{t,X}(0) + \frac{1}{a^2} \phi'_{t,X}(0) + \frac{1}{a^3} \phi''_{t,X}(0) \quad (4.5)$$

with respect to $A(t)$.

In some cases, it is not obvious to find $\phi''_{t,X}(0)$. In that case, we have to limit ourselves to using a first-order Taylor expansion of $\phi_{t,X}(u)$ in 0, which yields:

$$\int_{u=0}^{\infty} e^{-au} \phi_{t,X}(u) du = \frac{1}{a} \phi_{t,X}(0) + \frac{1}{a^2} \phi'_{t,X}(0) + O\left(\frac{1}{a^3}\right)$$

In that case, we have to find the scoring function that minimises the approximation:

$$\frac{1}{a} \phi_{t,X}(0) + \frac{1}{a^2} \phi'_{t,X}(0) \quad (4.6)$$

Table 4.4 contains the scoring functions obtained for each resource (vaccination and treatment) and for each objective (defined by the form of $J_I^A(t)$) with the greedy approach. In the following, we detail how these scores were obtained.

Resource	$J_I^A(t)$	Scoring function for herd j	Scoring function name
Vaccine	$\sum_{j=1}^J I_j(t)$	$\beta_j \frac{I_j(t)}{N_j(t)} S_j(t) \left(\mathbb{1}_{I_j(t)=1} (\gamma + \tau_j + \sum_{i \neq j} \theta_{ji}) + \sum_{i \neq j} \theta_{ji} \mathbb{1}_{I_i(t)=0} \right)$	<i>greedyV_infherds</i>
	$\sum_{j=1}^J \mathbb{1}_{I_j(t)>0}$	$\beta_j \frac{I_j(t)}{N_j(t)} S_j(t)$	<i>greedyV_infanimals</i>
Treatment	$\sum_{j=1}^J I_j(t)$	$\left[-\tau_j - \sum_{i \neq j} \theta_{ji} \mathbb{1}_{I_i(t)>0} \right]_{j: I_j(t)=1}$	<i>greedyT_infherds</i>
	$\sum_{j=1}^J \mathbb{1}_{I_j(t)>0}$	$I_j(t)$	<i>greedyT_infanimals</i>

Table 4.4: Greedy scoring functions for herd j , found for both types of resource, and for an objective that depends either on the number of infected herds, or on the total number of infected animals.

4.7.1 Vaccination

For vaccination, we assumed that b_{fix} is an available number of doses. The constraint in the optimisation problem is therefore by $\sum_{j=1}^J N_j(t) A_j(t) \leq b_{fix}$. We assume that if applied to a susceptible animal at time t , the disease transmission rate towards this susceptible becomes $\beta^v = \beta(1 - e_v)$ during the period $]t; t + \Delta_d]$, where $0 \leq e_v \leq 1$ is the protective effectiveness of the vaccine. That is, Δ_d is also the duration of the effectiveness of the vaccine.

4.7.1.1 Minimise a function of the number of infected herds

In the optimisation problem, we define $J_I^A(t) = \sum_{j=1}^J \mathbb{1}_{I_j(t)>0}$.

Here, we use a second-order approximation of $\phi_{t,X}(u)$ evaluated at 0. So we minimise expression 4.5 with respect to $A(t)$.

- Minimise $\phi_{t,X}(0)$

$$\phi_{t,X}(0) = \mathbb{E} [J_I^A(t) | X(t) = X] = \mathbb{E} \left[\sum_{j=1}^J \mathbb{1}_{I_j(t) > 0} | X(t) = X \right] = \sum_{j=1}^J \mathbb{1}_{I_j > 0} \quad (4.7)$$

which does not depend on $A(t)$

- Minimise $\phi'_{t,X}(0)$

We note that $\phi_{t,X}(u) = \phi_{0,X}(u) = \mathbb{E}[J_I^A(u)]$ if $X(0) = X$:

$$\begin{aligned} \phi_{t,X}(u) &= \mathbb{E} [J_I^A(t+u) | X(t) = X] \\ &= \mathbb{E} [J_I^A(u) | X(0) = X] \quad (\text{Markov}) \\ &= \mathbb{E}[J_I^A(u)] \text{ if } X(0) = X \end{aligned}$$

Let $Q = \{Q_{(x,x')}, x, x' \in \mathbb{N}^{3J}\}$ be the infinitesimal generator matrix of the Markov process $\{X(t)\}_t = \{(S(t), I(t), R(t))\}_t$.

Let the canonical basis of \mathbb{N}^{3J} be written as $(e_j^S, \dots, e_j^S, e_1^I, \dots, e_j^I, e_1^R, \dots, e_j^R)$.

In the present case, where the resource to be allocated is a vaccine that if applied in a herd decreases the infectious rate towards susceptible animals in the herd, the infinitesimal transition rates of the process are given by:

$$Q_{(x,x')} = \left\{ \begin{array}{lll} \beta_j \frac{I_j}{N_j} S_j (1 - e_v A_j(t)) & \text{if} & x' = x + e_j^I - e_j^S \quad (\text{infection in } j) \\ \gamma I_j & \text{if} & x' = x + e_j^R - e_j^I \quad (\text{recovery in } j) \\ \theta_{ji} S_j & \text{if} & x' = x + e_i^S - e_j^S \quad (\text{susceptible animal going from } j \text{ to } i) \\ \theta_{ji} I_j & \text{if} & x' = x + e_i^I - e_j^I \quad (\text{infected animal going from } j \text{ to } i) \\ \theta_{ji} R_j & \text{if} & x' = x + e_i^R - e_j^R \quad (\text{recovered animal going from } j \text{ to } i) \\ \tau_j S_j & \text{if} & x' = x - e_j^S \quad (\text{death of a susceptible animal in } j) \\ \tau_j I_j & \text{if} & x' = x - e_j^I \quad (\text{death of an infected animal in } j) \\ \tau_j R_j & \text{if} & x' = x - e_j^R \quad (\text{death of a recovered animal in } j) \\ \mu_j N_j & \text{if} & x' = x + e_j^S \quad (\text{birth of a susceptible animal in } j) \\ 0 & \text{otherwise} & \end{array} \right.$$

From now on, we note $\vec{\theta}_j = \sum_{i \neq j}^J \theta_{ji}$. We use Q as the derivation operator of the law of the process, as described in Méléard 2016, to obtain $\frac{d}{dt} \mathbb{E}[f(X(t))] = \mathbb{E}[g(X(t))]$, where $g(x) = \sum_{x \neq x'} Q_{x,x'} [f(x') - f(x)]$, for the function $f(X(t)) = \mathbb{1}_{I_j(t) > 0}$:

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t) > 0}] &= \mathbb{E} \left[(\mathbb{1}_{I_j(t)+1 > 0} - \mathbb{1}_{I_j(t) > 0}) \left(\beta_j (1 - e_v A_j(t)) \frac{I_j(t)}{N_j(t)} S_j(t) + \sum_{i \neq j}^J \theta_{ij} I_i(t) \right) \right] \\ &\quad + (\gamma + \tau_j + \vec{\theta}_j) \mathbb{E} [(\mathbb{1}_{I_j(t)-1 > 0} - \mathbb{1}_{I_j(t) > 0}) I_j(t)] \\ &= \mathbb{E} \left[\mathbb{1}_{I_j(t)=0} \left(\beta_j (1 - e_v A_j(t)) \frac{I_j(t)}{N_j(t)} S_j(t) + \sum_{i \neq j}^J \theta_{ij} I_i(t) \right) \right] \\ &\quad - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{E}[\mathbb{1}_{I_j(t)=1} I_j(t)] \\ &= \sum_{i \neq j}^J \theta_{ij} \mathbb{E}[\mathbb{1}_{I_j(t)=0} I_i(t)] - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{E}[\mathbb{1}_{I_j(t)=1}] \end{aligned} \quad (4.8)$$

Then

$$\frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t)>0}] \Big|_{t=0} = \sum_{i \neq j}^J \theta_{ij} \mathbb{1}_{I_j=0} I_i - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j=1}$$

Thus:

$$\phi'_{t,X}(0) = \sum_{j=1}^J \left(\sum_{i \neq j}^J \theta_{ij} \mathbb{1}_{I_j=0} I_i - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j=1} \right) \quad (4.9)$$

which does not depend on $A(t)$

- Minimise $\phi''_{t,X}(0)$

From equation 4.8, we have:

$$\begin{aligned} \phi''_{t,X}(u) &= \frac{d}{du} \phi'_{t,X}(u) = \sum_j \frac{d}{du} \mathbb{E}[\mathbb{1}_{I_j(u)>0}] \\ &= \sum_{ji} \theta_{ij} \frac{d}{du} \mathbb{E}[\mathbb{1}_{I_j(u)=0} I_i(u)] - \sum_j (\gamma + \tau_j + \vec{\theta}_j) \frac{d}{du} \mathbb{E}[\mathbb{1}_{I_j(u)=1}] \end{aligned}$$

Again, using the generator of the Markov process, we find:

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t)=0} I_i(t)] \Big|_{t=0} &= -(\mathbb{1}_{I_j=0}(I_i + 1) - \mathbb{1}_{I_j=0} I_i) \beta_i e_v \frac{I_i}{N_i} S_i A_i(t) + \Xi \\ &= -\beta_i e_v \frac{I_i}{N_i} S_i \mathbb{1}_{I_j=0} A_i(t) + \Xi \\ &= -\beta_i e_v \frac{I_i}{N_i} S_i \mathbb{1}_{I_j=0} A_i(t) + \Xi \end{aligned}$$

where Ξ is independent of $A(t)$ ¹. And:

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t)=1}] \Big|_{t=0} &= -\beta_j e_v (\mathbb{1}_{I_j+1=1} - \mathbb{1}_{I_j=1}) A_j(t) \frac{I_j}{N_j} S_j + \Xi \\ &= -\beta_j e_v (\mathbb{1}_{I_j=0} - \mathbb{1}_{I_j=1}) \frac{I_j}{N_j} S_j A_j(t) + \Xi \\ &= \beta_j e_v \mathbb{1}_{I_j=1} \frac{S_j}{N_j} A_j(t) + \Xi \end{aligned}$$

where Ξ is independent of $A(t)$. Thus

$$\begin{aligned} \phi''_{t,X}(0) &= -e_v \sum_{j=1}^J \sum_{i \neq j}^J \theta_{ij} \left[\beta_i \frac{I_i}{N_i} S_i \mathbb{1}_{I_j=0} A_i(t) \right] - e_v \sum_{j=1}^J \beta_j (\gamma + \tau_j + \vec{\theta}_j) \left[\mathbb{1}_{I_j=1} \frac{S_j}{N_j} A_j(t) \right] + \Xi \\ &= -e_v \sum_{j=1}^J \sum_{i \neq j}^J \theta_{ji} \left[\beta_j \mathbb{1}_{I_i=0} \frac{I_j}{N_j} S_j A_j(t) \right] - e_v \sum_{j=1}^J \beta_j (\gamma + \tau_j + \vec{\theta}_j) \left[\mathbb{1}_{I_j=1} \frac{S_j}{N_j} A_j(t) \right] + \Xi \end{aligned} \quad (4.10)$$

¹later we keep noting as Ξ all terms that are not of interest for the minimisation problem, since they do not depend on $A(t)$, but we remark that their actual value varies through-out the document.

where Ξ is independent of $A(t)$. Then:

$$\begin{aligned} \min_{A(t)} \phi''_{t,X}(0) &\iff \min_{A(t)} -e_v \sum_{j=1}^J \left(\sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i=0} \beta_j \frac{I_j}{N_j} S_j A_j(t) + \beta_j (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j=1} \frac{S_j}{N_j} A_j(t) \right) \\ &\iff \max_{A(t)} \sum_{j=1}^J \left(\beta_j (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j=1} \frac{S_j}{N_j} + \beta_j \frac{I_j}{N_j} S_j \sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i=0} \right) A_j(t) \end{aligned} \quad (4.11)$$

Since expressions 4.7 and 4.9 do not depend on $A(t)$, minimising 4.5 is equivalent to minimising 4.10. From 4.11, this can be achieved if we assign the $A_j(t) = 1$ or 0 according to the scoring function:

$$\left[\beta_j \frac{I_j(t)}{N_j(t)} S_j(t) \left(\mathbb{1}_{I_j(t)=1} (\gamma + \tau_j + \vec{\theta}_j) + \sum_{i \neq j}^J \theta_{ji} \mathbb{1}_{I_i(t)=0} \right) \right]_{j=1, \dots, J} \quad (4.12)$$

4.7.1.2 Minimise a function of the total number of infected animals

In the optimisation problem, we define $J_I^A(t) = \sum_{j=1}^J I_j(t)$.

Here, we use a first-order approximation of $\phi_{t,X}(u)$ evaluated at 0. Indeed, it was not possible to arrive to an analytic expression of $\phi''_{t,X}(0)$ for this case. So we minimise 4.6 with respect to A .

- Minimise $\phi_{t,X}(0)$

$$\phi_{t,X}(0) = \mathbb{E} [J_I^A(t) | X(t) = X] = \mathbb{E} \left[\sum_{j=1}^J I_j(t) | X(t) = X \right] = \sum_{j=1}^J I_j \quad (4.13)$$

which does not depend on $A(t)$

- Minimise $\phi'_{t,X}(0)$

We note that $\phi_{t,X}(u) = \phi_{0,X}(u) = \mathbb{E}[J_I^A(u)]$ if $X(0) = X$:

$$\begin{aligned} \phi_{t,X}(u) &= \mathbb{E} [J_I^A(t+u) | X(t) = X] \\ &= \mathbb{E} [J_I^A(u) | X(0) = X] \quad (\text{Markov}) \\ &= \mathbb{E}[J_I^A(u)] \text{ if } X(0) = X \end{aligned}$$

In our model:

$$\left. \frac{d}{dt} \mathbb{E}[I_j(t)] \right|_{t=0} = \beta_j \frac{I_j}{N_j} S_j (1 - e_v A_j(t)) + \sum_{i \neq j}^J \theta_{ij} I_i - (\gamma + \tau_j + \vec{\theta}_j) I_j$$

Then

$$\begin{aligned}
\phi'_{t,X}(0) &= \sum_{j=1}^J \beta_j \frac{I_j}{N_j} S_j (1 - e_v A_j(t)) + \sum_{j=1}^J \sum_{i \neq j}^J \theta_{ij} I_i - \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) I_j \\
&= -e_v \sum_{j=1}^J \beta_j \frac{I_j}{N_j} S_j A_j(t) + \Xi
\end{aligned} \tag{4.14}$$

where Ξ is independent of $A(t)$. Then:

$$\begin{aligned}
\min_{A(t)} \phi'_{t,X}(0) &\iff \min_{A(t)} -e_v \sum_{j=1}^J \beta_j \frac{I_j}{N_j} S_j A_j(t) \\
&\iff \max_{A(t)} \sum_{j=1}^J \beta_j \frac{I_j}{N_j} S_j A_j(t)
\end{aligned} \tag{4.15}$$

Since 4.13 does not depend on $A(t)$, minimising 4.6 is equivalent to minimising 4.14. From 4.15, this can be achieved if we assign the $A_j(t) = 1$ or 0 according to the scoring function:

$$\left[\beta_j \frac{I_j(t)}{N_j(t)} S_j(t) \right]_{j=1, \dots, J} \tag{4.16}$$

4.7.2 Treatment

4.7.2.1 Minimise a function of the number of infected herds

For the treatment, we assumed that b_{fix} is not a number of doses, but is the number of herds that can be treated, i.e. of attainable herds. The constraint in the optimisation problem is therefore by $\sum_{j=1}^J A_j(t) \leq b_{fix}$. We assume the treatment increases the recovery rate of treated infected animals by an additional factor γ' . That is, the mean duration of the infectious period for these animals becomes $(\gamma + \gamma')^{-1}$.

In the optimisation problem, we define $J_I^A(t) = \sum_{j=1}^J \mathbb{1}_{I_j(t) > 0}$.

Here, we use a second-order approximation of $\phi_{t,X}(u)$ evaluated at 0. So we minimise 4.5 with respect to $A(t)$.

- Minimise $\phi_{t,X}(0)$.

$$\phi_{t,X}(0) = \mathbb{E} [J_I^A(t) | X(t) = X] = \mathbb{E} \left[\sum_{j=1}^J \mathbb{1}_{I_j(t) > 0} | X(t) = X \right] = \sum_{j=1}^J \mathbb{1}_{I_j > 0} \tag{4.17}$$

which does not depend on $A(t)$

- Minimise $\phi'_{t,X}(0)$

We note that $\phi_{t,X}(u) = \phi_{0,X}(u) = \mathbb{E}[J_I^A(u)]$ if $X(0) = X$:

$$\begin{aligned}
\phi_{t,X}(u) &= \mathbb{E} [J_I^A(t+u) | X(t) = X] \\
&= \mathbb{E} [J_I^A(u) | X(0) = X] \quad (\text{Markov}) \\
&= \mathbb{E}[J_I^A(u)] \text{ if } X(0) = X
\end{aligned}$$

Thus, we are looking for:

$$\phi'_{t,X}(u) = \frac{d}{du} \mathbb{E}[J_I^A(u)] = \frac{d}{du} \mathbb{E} \left[\sum_j \mathbb{1}_{I_j(u) > 0} \right] = \sum_j \frac{d}{du} \mathbb{E}[\mathbb{1}_{I_j(u) > 0}]. \quad (4.18)$$

From now on, we note $\vec{\theta}_j = \sum_{i \neq j} \theta_{ji}$. Now, let $Q = \{Q_{(x,x')}, x, x' \in \mathbb{N}^{3J}\}$ be the infinitesimal generator matrix of the Markov process $\{X(t)\}_t = \{(S(t), I(t), R(t))\}_t$. Let the canonical basis of \mathbb{N}^{3J} be written as $(e_j^S, \dots, e_j^S, e_1^I, \dots, e_J^I, e_1^R, \dots, e_J^R)$.

In the present case, where the resource to be allocated is a treatment that increases the recovery rate, the infinitesimal transition rates of the process are given by:

$$Q_{(x,x')} = \left\{ \begin{array}{lll} \beta_j \frac{I_j}{N_j} S_j & \text{if} & x' = x + e_j^I - e_j^S \quad (\text{infection in } j) \\ (\gamma + \gamma' A_j(t)) I_j & \text{if} & x' = x + e_j^R - e_j^I \quad (\text{recovery in } j) \\ \theta_{ji} S_j & \text{if} & x' = x + e_i^S - e_j^S \quad (\text{susceptible animal going from } j \text{ to } i) \\ \theta_{ji} I_j & \text{if} & x' = x + e_i^I - e_j^I \quad (\text{infected animal going from } j \text{ to } i) \\ \theta_{ji} R_j & \text{if} & x' = x + e_i^R - e_j^R \quad (\text{recovered animal going from } j \text{ to } i) \\ \tau_j S_j & \text{if} & x' = x - e_j^S \quad (\text{death of a susceptible animal in } j) \\ \tau_j I_j & \text{if} & x' = x - e_j^I \quad (\text{death of an infected animal in } j) \\ \tau_j R_j & \text{if} & x' = x - e_j^R \quad (\text{death of a recovered animal in } j) \\ \mu_j N_j & \text{if} & x' = x + e_j^S \quad (\text{birth of a susceptible animal in } j) \\ 0 & \text{otherwise} & \end{array} \right.$$

We use Q as the derivation operator of the law of the process, as described in Méléard 2016, to obtain $\frac{d}{dt} \mathbb{E}[f(X(t))] = \mathbb{E}[g(X(t))]$, where $g(x) = \sum_{x \neq x'} Q_{x,x'} [f(x') - f(x)]$, for the function $f(X(t)) = \mathbb{1}_{I_j(t) > 0}$:

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t) > 0}] &= \mathbb{E} \left[\mathbb{1}_{I_j(t)=0} \left(\beta_j \frac{I_j(t)}{N_j(t)} S_j(t) + \sum_{i \neq j} \theta_{ij} I_i(t) \right) \right] \\ &\quad - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{E}[\mathbb{1}_{I_j(t)=1} I_j(t)] - \gamma' \mathbb{E}[\mathbb{1}_{I_j(t)=1} I_j(t) A_j(t)] \\ &= \sum_{i \neq j} \theta_{ij} \mathbb{E}[\mathbb{1}_{I_j(t)=0} I_i(t)] - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{E}[\mathbb{1}_{I_j(t)=1}] - \gamma' \mathbb{E}[\mathbb{1}_{I_j(t)=1} A_j(t)] \end{aligned} \quad (4.19)$$

And so:

$$\left. \frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t) > 0}] \right|_{t=0} = \sum_{i \neq j} \theta_{ij} \mathbb{1}_{I_j=0} I_i - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j=1} - \gamma' \mathbb{1}_{I_j=1} A_j(t)$$

Then:

$$\begin{aligned} \phi'_{t,X}(0) &= \sum_{j=1}^J \left(\sum_{i \neq j} \theta_{ij} \mathbb{1}_{I_j=0} I_i - (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j=1} - \gamma' \mathbb{1}_{I_j=1} A_j(t) \right) \\ &= \sum_{j=1}^J -\gamma' \mathbb{1}_{I_j=1} A_j(t) + \Xi \end{aligned} \quad (4.20)$$

where Ξ is independent of $A(t)$. Thus

$$\begin{aligned} \min_{A(t)} \phi'_{t,X}(0) &\iff \min_{A(t)} -\gamma' \sum_{j=1}^J \mathbb{1}_{I_j=1} A_j(t) \\ &\iff \max_{A(t)} \gamma' \sum_{j=1}^J \mathbb{1}_{I_j=1} A_j(t). \end{aligned}$$

This sum is maximal if only the herds that have 1 infected are chosen, i.e. if $I_j(t) \neq 1 \implies A_j(t) = 0$. Thus, if we treat only the herds with one infected animal, the total number of herds we treat will be $\sum_{j=1}^J \mathbb{1}_{I_j=1} A_j(t) = \min(b_{fix}, \sum_{j=1}^J \mathbb{1}_{I_j=1})$.

- Minimise $\phi''_{t,X}(0)$

From 4.19, we have:

$$\begin{aligned} \phi''_{t,X}(u) &= \frac{d}{du} \phi'_{t,X}(u) = \sum_j \frac{d}{du} \mathbb{E}[\mathbb{1}_{I_j(u)>0}] \\ &= \sum_{ji} \theta_{ij} \frac{d}{du} \mathbb{E}[\mathbb{1}_{I_j(u)=0} I_i(u)] - \sum_j (\gamma + \tau_j + \vec{\theta}_j) \frac{d}{du} \mathbb{E}[\mathbb{1}_{I_j(u)=1}] - \gamma' \frac{d}{du} \mathbb{E} \left[\sum_j \mathbb{1}_{I_j(u)=1} A_j(u) \right] \end{aligned} \quad (4.21)$$

Again, using the infinitesimal generator of the Markov process, we find:

$$\begin{aligned} \left. \frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t)=0} I_i(t)] \right|_{t=0} &= \gamma' [-\mathbb{1}_{I_j=0} I_i A_i(t)] + \Xi \\ &= -\gamma' \mathbb{1}_{I_j=0} I_i A_i(t) + \Xi \\ &= -\gamma' \mathbb{1}_{I_j=0} A_i(t) + \Xi \end{aligned}$$

where Ξ is independent of $A(t)$, and where we have used that if $I_j(t) \neq 1 \implies A_j(t) = 0$. Also:

$$\begin{aligned} \left. \frac{d}{dt} \mathbb{E}[\mathbb{1}_{I_j(t)=1}] \right|_{t=0} &= \beta_j (\mathbb{1}_{I_j+1=1} - \mathbb{1}_{I_j=1}) \frac{I_j}{N_j} S_j + \sum_{i \neq j} \theta_{ij} (\mathbb{1}_{I_j+1=1} - \mathbb{1}_{I_j=1}) I_i \\ &\quad + (\gamma + \tau_j + \vec{\theta}_j) (\mathbb{1}_{I_j-1=1} - \mathbb{1}_{I_j=1}) I_j + \gamma' (\mathbb{1}_{I_j-1=1} - \mathbb{1}_{I_j=1}) I_j A_j(t) \\ &= \beta_j (\mathbb{1}_{I_j=0} - \mathbb{1}_{I_j=1}) \frac{I_j}{N_j} S_j + \sum_{i \neq j} \theta_{ij} (\mathbb{1}_{I_j=0} - \mathbb{1}_{I_j=1}) I_i \\ &\quad + (\gamma + \tau_j + \vec{\theta}_j) (\mathbb{1}_{I_j=2} - \mathbb{1}_{I_j=1}) I_j + \gamma' (\mathbb{1}_{I_j=2} - \mathbb{1}_{I_j=1}) I_j A_j(t) \\ &= -\beta_j \mathbb{1}_{I_j=1} \frac{I_j}{N_j} S_j + \sum_{i \neq j} \theta_{ij} (\mathbb{1}_{I_j=0} - \mathbb{1}_{I_j=1}) I_i + (\gamma + \tau_j + \vec{\theta}_j) (\mathbb{1}_{I_j=2} - \mathbb{1}_{I_j=1}) I_j \\ &\quad + \gamma' (\mathbb{1}_{I_j=2} - \mathbb{1}_{I_j=1}) I_j A_j(t) \\ &= \gamma' (\mathbb{1}_{I_j=2} - \mathbb{1}_{I_j=1}) I_j A_j(t) + \Xi \\ &= \gamma' \mathbb{1}_{I_j=2} I_j A_j(t) - \gamma' \mathbb{1}_{I_j=1} A_j(t) + \Xi \\ &= -\gamma' A_j(t) + \Xi \end{aligned}$$

where Ξ is independent of $A(t)$, and where we have used that if $I_j(t) \neq 1 \implies A_j(t) = 0$.

Now, following the procedure in Scaman et al. 2015, we show that $\frac{d}{dt} \mathbb{E} \left[\sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) \right]$ does not depend on $A(t)$.

We have that the constraint is $\sum_{j=1}^J A_j(t) \leq b_{fix}$. From the minimisation of $\phi'_{t,X}(0)$ we obtain that we only treat herds with one infected, i.e. if $I_j(t) \neq 1 \implies A_j(t) = 0$. The total number of herds treated will be $\sum_{j=1}^J A_j(t) = \sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) = \min(b_{fix}, \sum_{j=1}^J \mathbb{1}_{I_j(t)=1})$. Let $H(t) = \min(b_{fix}, \sum_{j=1}^J \mathbb{1}_{I_j(t)=1})$. We can calculate $\frac{d}{dt} \mathbb{E} \left[\sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) \right]$:

$$\frac{d}{dt} \mathbb{E} \left[\sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) \right] = \lim_{\Delta_t \rightarrow 0} \frac{\mathbb{E}[H(t + \Delta_t)] - \mathbb{E}[H(t)]}{\Delta_t}$$

Let $N_I(t) = \sum_{j=1}^J \mathbb{1}_{I_j(t)=1}$. Let Δ_t be a sufficiently small time interval. Three scenarios are possible:

- $N_I(t) > b_{fix}$. And for $t' \in [t, t + \Delta_t]$, $H(t')$ is stationary (since $N_I(t')$ can at most increase or decrease by 1),
- either $N_I(t) < b_{fix}$. And $H(t') = N_I(t')$ for $t' \in [t, t + \Delta_t]$,
- or the last possibility is $N_I(t) = b_{fix}$. And in this case it is sufficient to consider the case where $N_I(t + \Delta_t) = b_{fix} - 1$ (by construction $H(t')$ will not change if $N_I(t)$ increases).

We can then write:

$$\begin{aligned} & E[H(t + \Delta_t) | X(t)] \\ &= \mathbb{1}_{N_I(t) > b_{fix}} b_{fix} \\ &+ \mathbb{1}_{N_I(t) = b_{fix}} \left[b_{fix} - \sum_{j=1}^J \beta_j \mathbb{1}_{I_j(t)=1} \frac{S_j(t)}{N_j(t)} \Delta_t - \sum_{j=1}^J \sum_{i \neq j}^J \theta_{ij} \mathbb{1}_{I_j(t)=1} I_i(t) \Delta_t \right. \\ &\quad \left. - \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j(t)=1} \Delta_t - \gamma' \sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) \Delta_t \right] \\ &+ \mathbb{1}_{N_I(t) < b_{fix}} \left[N_I(t) - \sum_{j=1}^J \beta_j \mathbb{1}_{I_j(t)=1} \frac{S_j(t)}{N_j(t)} \Delta_t + \sum_{j=1}^J \sum_{i \neq j}^J \theta_{ij} (\mathbb{1}_{I_j(t)=0} - \mathbb{1}_{I_j(t)=1}) I_i(t) \Delta_t \right. \\ &\quad \left. + \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) (\mathbb{1}_{I_j(t)=2} - \mathbb{1}_{I_j(t)=1}) I_j(t) \Delta_t - \gamma' \sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) \Delta_t \right] + o(\Delta_t) \\ &= H(t) \\ &- \mathbb{1}_{N_I(t) \leq b_{fix}} \left[\sum_{j=1}^J \beta_j \mathbb{1}_{I_j(t)=1} \frac{S_j(t)}{N_j(t)} + \sum_{j=1}^J \sum_{i \neq j}^J \theta_{ij} \mathbb{1}_{I_j(t)=1} I_i(t) \right. \\ &\quad \left. + \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j(t)=1} + \gamma' \sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) \right] \Delta_t \\ &+ \mathbb{1}_{N_I(t) < b_{fix}} \left[\sum_{j=1}^J \sum_{i \neq j}^J \theta_{ij} \mathbb{1}_{I_j(t)=0} I_i(t) + \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j(t)=2} I_j(t) \right] \Delta_t + o(\Delta_t) \end{aligned}$$

We have therefore:

$$\begin{aligned} \frac{d}{dt} \mathbb{E} \left[\sum_{j=1}^J \mathbb{1}_{I_j(t)=1} A_j(t) \right] &= \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E} [\mathbb{E}[H(t + \Delta t) | X(t)]] - \mathbb{E}[H(t)]}{\Delta t} \\ &= -\mathbb{E} \left[\mathbb{1}_{N_I(t) \leq b_{fix}} \left(\sum_{j=1}^J \beta_j \mathbb{1}_{I_j(t)=1} \frac{S_j(t)}{N_j(t)} + \sum_{j=1}^J \sum_{i \neq j} \theta_{ij} \mathbb{1}_{I_j(t)=1} I_i(t) \right. \right. \\ &\quad \left. \left. + \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j(t)=1} + \gamma' H(t) \right) \right] \\ &\quad + \mathbb{E} \left[\mathbb{1}_{N_I(t) < b_{fix}} \left(\sum_{j=1}^J \sum_{i \neq j} \theta_{ij} \mathbb{1}_{I_j(t)=0} I_i(t) + \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) \mathbb{1}_{I_j(t)=2} I_j(t) \right) \right] \end{aligned}$$

which does not depend on $A(t)$

All in all, we have:

$$\begin{aligned} \phi''_{t,X}(0) &= -\gamma' \sum_{j=1}^J \sum_{i \neq j} \theta_{ij} \mathbb{1}_{I_j=0} A_i(t) + \gamma' \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) A_j(t) + \Xi \\ &= -\gamma' \sum_{j=1}^J \sum_{i \neq j} \theta_{ji} \mathbb{1}_{I_i=0} A_j(t) + \gamma' \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) A_j(t) + \Xi \end{aligned} \quad (4.22)$$

where Ξ is independent of $A(t)$. And so:

$$\begin{aligned} \min_{A(t)} \phi''_{t,X}(0) &\iff \min_{A(t)} -\gamma' \sum_{j=1}^J \left(\sum_i \theta_{ji} \mathbb{1}_{I_i=0} - (\gamma + \tau_j + \vec{\theta}_j) \right) A_j(t) \\ &\iff \max_{A(t)} \sum_{j=1}^J \left(\sum_i \theta_{ji} \mathbb{1}_{I_i=0} - (\gamma + \tau_j + \vec{\theta}_j) \right) A_j(t) \end{aligned}$$

Then, the minimisation of $\phi'_{t,X}(0)$ and $\phi''_{t,X}(0)$ in 4.5 can be achieved simultaneously if we assign the $A_j(t) = 1$ or 0 according to the scoring function:

$$\left[\sum_i \theta_{ji} \mathbb{1}_{I_i(t)=0} - (\gamma + \tau_j + \vec{\theta}_j) \right]_{j: I_j(t)=1}$$

Since γ , the recovery rate from the disease, is the same for all herds, the scoring function can be rewritten as:

$$\left[-\tau_j - \sum_i \theta_{ji} \mathbb{1}_{I_i(t)>0} \right]_{j: I_j(t)=1} \quad (4.23)$$

4.7.2.2 Minimise a function of the total number of infected animals

We assume that in the optimisation problem, $J_I^A(t) = \sum_{j=1}^J I_j^A(t)$.

Here, we use a first-order approximation of $\phi_{t,X}(u)$ evaluated at 0. Indeed, it was not possible to arrive to an analytic expression of $\phi''_{t,X}(0)$ for this case. So we minimise 4.6 with respect to A .

- Minimise $\phi_{t,X}(0)$

$$\phi_{t,X}(0) = \mathbb{E} [J_I^A(t) | X(t) = X] = \mathbb{E} \left[\sum_{j=1}^J I_j(t) | X(t) = X \right] = \sum_{j=1}^J I_j \quad (4.24)$$

which does not depend on $A(t)$

- Minimise $\phi'_{t,X}(0)$

We note that $\phi_{t,X}(u) = \phi_{0,X}(u) = \mathbb{E}[J_I^A(u)]$ if $X(0) = X$:

$$\begin{aligned} \phi_{t,X}(u) &= \mathbb{E} [J_I^A(t+u) | X(t) = X] \\ &= \mathbb{E} [J_I^A(u) | X(0) = X] \quad (\text{Markov}) \\ &= \mathbb{E}[J_I^A(u)] \text{ if } X(0) = X \end{aligned}$$

In our model with treatment:

$$\left. \frac{d}{dt} \mathbb{E}[I_j(t)] \right|_{t=0} = \beta_j \frac{I_j}{N_j} S_j + \sum_{i \neq j} \theta_{ij} I_i - (\gamma + \tau_j + \vec{\theta}_j) I_j - \gamma' I_j A_j(t)$$

Thus:

$$\begin{aligned} \phi'_{t,X}(0) &= \sum_{j=1}^J \beta_j \frac{I_j}{N_j} S_j + \sum_{j=1}^J \sum_{i \neq j} \theta_{ij} I_i - \sum_{j=1}^J (\gamma + \tau_j + \vec{\theta}_j) I_j - \gamma' \sum_{j=1}^J I_j A_j(t) \\ &= -\gamma' \sum_{j=1}^J I_j A_j(t) + \Xi \end{aligned} \quad (4.25)$$

where Ξ is independent of $A(t)$. Then:

$$\begin{aligned} \min_{A(t)} \phi'_{t,X}(0) &\iff \min_{A(t)} -\gamma' \sum_{j=1}^J I_j A_j(t) \\ &\iff \max_{A(t)} \sum_{j=1}^J I_j A_j(t) \end{aligned} \quad (4.26)$$

Since 4.24 does not depend on $A(t)$, minimising 4.6 is equivalent to minimising 4.25. From 4.26, this can be achieved if we assign the $A_j(t) = 1$ or 0 according to the scoring function:

$$[I_j(t)]_{j=1, \dots, J} \quad (4.27)$$

Chapter 5

Application study: Learning and strategic imitation in modelling farmers' dynamic decisions on BVD vaccination

Contents

5.1	Introduction	117
5.2	The BVD model	118
5.2.1	Data description	118
5.2.2	Within-herd epidemiological-demographic dynamics	122
5.3	Farmer's decision-making on vaccination	129
5.3.1	Decision-mechanism	129
5.3.2	Economic-epidemiological cost	131
5.4	Simulation setting	133
5.5	Results	134
5.6	Discussion	140

This chapter presents an extension and an adaptation of the model in chapter 3. The extension regards the decision-making component of the model: farmers can either take into account the decisions and observed costs of their trade neighbours, or of their geographic neighbours. The adaption is built upon a specific disease, bovine viral diarrhoea (BVD), for which an existing individual-based stochastic model is considered.

The modelling choices in this chapter were motivated by the state of the art in section 2.1.1.1.2, section 2.1.1.1.4 and section 2.1.3.4, regarding epidemiological modelling, and in section 2.2.2 and section 2.3, regarding decentralised decision-making modelling.

5.1 Introduction

Accounting for farmer's dynamic decision regarding control measures is a key to better understand livestock disease spread at a large scale, and to better predict it. In Cristancho Fajardo et al. 2021 (chapter 3) an integrative model was proposed for this purpose. The model accounted for the dynamic decision-making process of farmers regarding the adoption of a control measure. In particular, phenomena such as learning and strategic imitation were considered. Yet, this model lied on a theoretical SIR infection dynamics, and for the decision dynamics it was assumed that

farmers shared information relative to their decisions only through the trade network on which the disease spread. This chapter presents an extension of such model that lies on an application to a specific disease: bovine viral diarrhoea (BVD).

BVD is a viral disease of cattle that causes economic losses and reductions in animals well-being (abortions, calving delays, and mortality) worldwide (Gates et al. 2013a; Piniot et al. 2017). In particular, the infection of a pregnant female during mid-gestation, can lead to the birth of persistently infected (PI) calves (Fray et al. 2000). Although PI animals generally have a reduced lifespan (Houe 1993), they are highly infectious during their whole lives. Therefore, they can not only threaten the immunity of the herd they belong to, by causing new infections, but also spread the infection to neighbouring herds through pasture contacts and trade movements (Gates et al. 2013b; Graham et al. 2016).

It has previously been established that combining testing with compulsory vaccination could eradicate the virus (Gethmann et al. 2019). However, since such a combined strategy has been proven not to be economically cost-effective, control is often left to farmers' voluntary vaccination. Indeed, regular vaccination can constitute a viable alternative for farmers to reduce new infections (Gethmann et al. 2019; Moennig et al. 2018). In particular, European countries use vaccines that can prevent vertical transmission if administered before pregnancy (Moennig et al. 2018). For such vaccines, and for some breeding systems, it has been shown that vaccination can be an economically interesting measure to control BVD virus spread (Arnoux et al. 2021).

The objective of the present study was two-fold. On the one hand, we sought to extend the decision-mechanism of the generic model initially proposed, considering several scenarios with respect to how information is shared among farmers. In particular, we assumed that information is shared through a neighbourhood defined on the basis of the trade network or by the means of a proximity network, i.e. the geographic position of herds. On the other hand, the objective was to adapt the generic model for a specific disease (BVD).

The chapter is structured as follows. First, section 5.2 presents the BVD model and the data on which it was built on. Then, section 5.3 details farmers' decision-making regarding vaccination. In particular, this section presents the extensions made to the original generic model. Section 5.4 defines the scenarios for simulations, whose results are included in section 5.5. Finally, section 5.6 presents a discussion and conclusion of this work on the basis of such results.

5.2 The BVD model

We used a BVD discrete-time stochastic individual-based model conceived by the DYNAMO team of the BIOEPAR unit of INRAE-Oniris, in which the introduction of the BVD virus into a BVD-free herd occurs through the purchase of infected animals, or through close contact with infected neighbours during the grazing season. For animal movements, real observed data were plugged into the model, so this component of the model is deterministic. For the other components of the model, life-cycle dynamics and health-related dynamics, some transitions were assumed to be deterministic, while others are stochastic. The parameters that determine transitions were either estimated from data, or fixed according to existing literature or expert knowledge.

5.2.1 Data description

We considered data for a five year period (January 2010- December 2014) relative to holdings in Saône-et-Loire, a French department densely populated with Charolais, a breed raised of beef cattle. Indeed, Arnoux et al. 2021 have shown that the combined characteristics of the Charolaise breeding system can yield to a strong economic impact of BVD virus spread and of vaccination against it, when compared to other breeding systems.

The data was obtained from the French cattle identification database (FCID), which records the complete life history of each cattle animal from birth to death. In particular, it includes all move-

ments between holdings. Hence, for each animal the available data are: country code, national identification number, sex, breed, birth date, date of entry(exit) into(from) each holding it belonged to, and cause (entry: birth or purchase, exit: death or sale).

In order to partly reduce the computational complexity of the model, we restricted ourselves to holdings whose yearly mean size was at least 30 animals for at least one of the five years i.e. we excluded holdings that each of the five years had a relatively small yearly mean size, and hence were unlikely to play a major role in pathogen spread. We remark that all herds that satisfy this condition were included in the metapopulation, irrespective to whether they exchanged animals or not. Additionally, 155 holdings were excluded either because they presented a null size on January 1st 2010, and could therefore not be initialised for simulation, or because their herd-specific parameters of the model could not be computed, since there were no births nor gestating animals in these herds during the study period (see section 5.2.2.1 for the definition of such parameters). Under this selection, the proportion of animal trade carried by excluded holdings was only 7%. The selection corresponded to 3685 out of the 5155 herds that were active at least once during the five years of study. Table 5.1 contains a summary of the distribution of movements between groups of herd, given by their belonging to the metapopulation, to the department, and to France.

Origin	Destination			
	metapop	small	outside	undetermined
metapop	0.139	0.031	0.187	0.418
small	0.015	0.002	0.008	0.018
outside	0.133	0.039	0.000	0.000
undetermined	0.009	0.002	0.000	0.000

Table 5.1: Distribution of movements according to the origin/destination group (as a proportion of the total movements concerning herds in Saône-et-Loire from 2010-2014): the group of herds included in the metapopulation to be modelled ('metapop'), the group of excluded herds from Saône-et-Loire ('small'), and the group of herds in France but outside Saône-et-Loire ('outside'). Movement whose origin/destination was not known are grouped under the term 'undetermined'.

The herd size distribution for herds in Saône-et-Loire within and outside the selection can be found in figure 5.1. Figure 5.2 represents all holdings in Saône-et-Loire that were active at some point during the five years of study, highlighting a heterogeneity in the spatial density of herds. The exact location of each holding was randomly assigned according to the coordinates of the administrative area (565 in total) the holding belongs to. Figure 5.3 represents the geographic location of the selected holdings, and their exchanges over the study period, also indicating a high spatial heterogeneity in trade patterns. Three geographic groups of holdings can be identified as being the most active in cattle exchange (in particular having a higher share of buying): one group in the south-west zone of the department, one in the north, and one in the west. The first and second groups are very interconnected (many links, some with a high trading volume). Herds in the latter group are not very connected within the group nor with the rest of the population, except from some long distance connections.

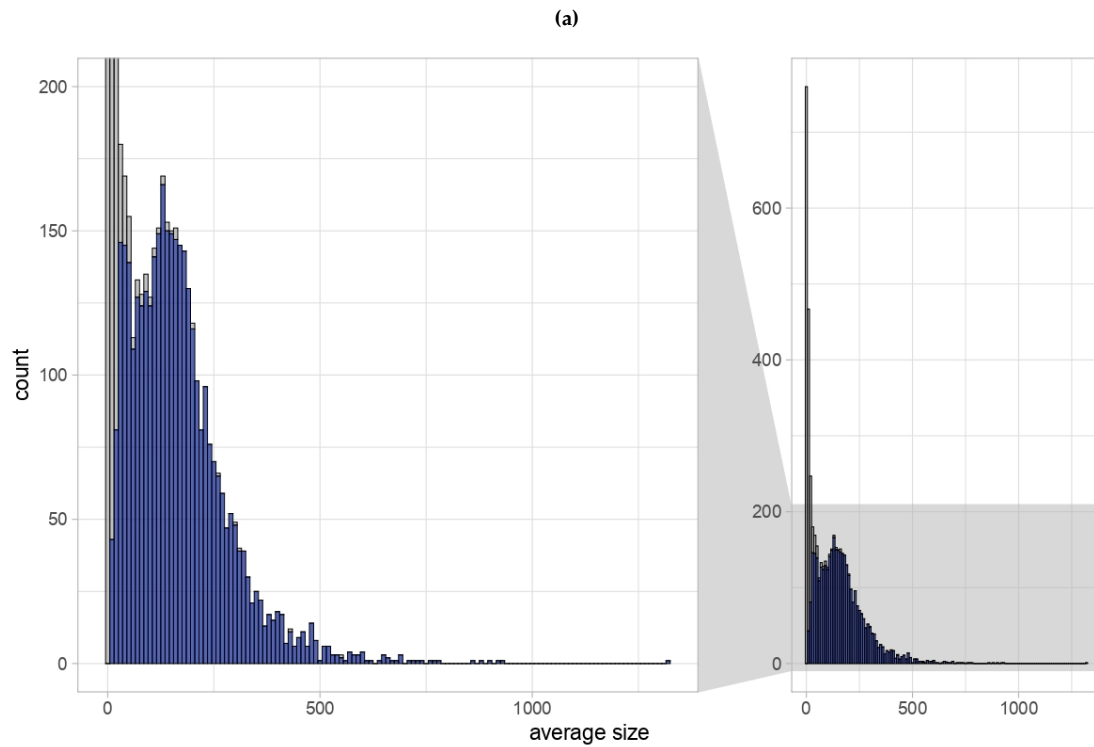


Figure 5.1: Herd size distributions in Saône-et-Loire French department from real data (FCID) for years 2010-2014. Distribution of the mean herd sizes over the five year period. Gray bars represent the distribution for all herds that were active (size > 0) at least once within the period (5155 herds). Blue bars (superimposed) represent the distribution for all herds whose yearly mean size was at least 30 animals for at least one of the five years (3685 herds). The picture is courtesy of Gaël Beaunée.



Figure 5.2: Geographic location of exchanging cattle holdings in Saône-et-Loire from 2010-2014 (5155 holdings). The location of each holding was randomly assigned according to the coordinates of the administrative area the holding belongs to. The picture is courtesy of Gaël Beaunée.

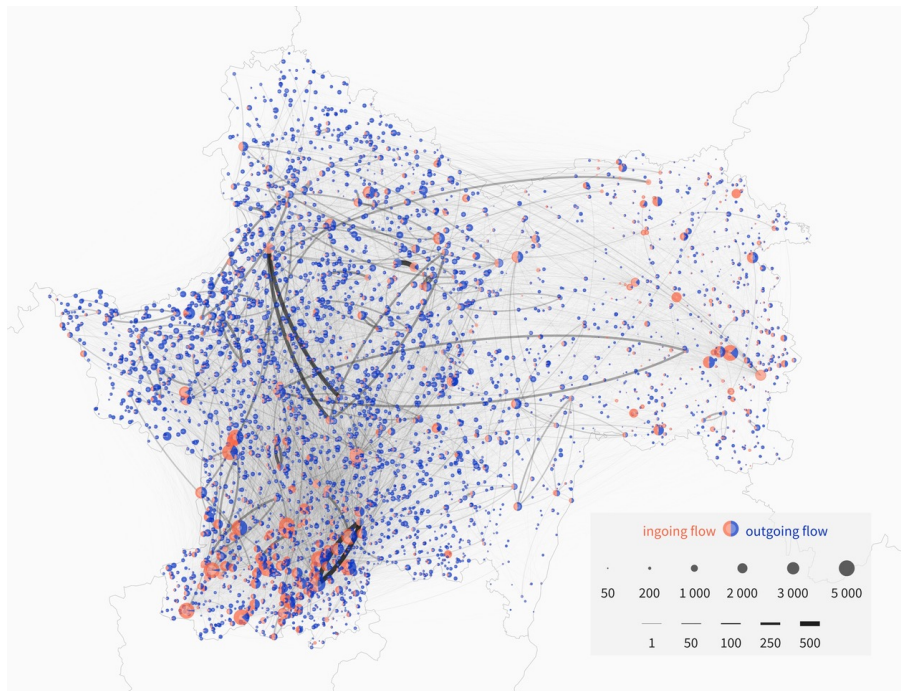


Figure 5.3: Geographic location of the selected exchanging cattle holdings in the metapopulation (3685 holdings). The location of each holding was randomly assigned according to the coordinates of the administrative area the holding belongs to. The size of a node represents its mean (over years) population. Colours of nodes represent the mean (over years) share of buying (orange) and selling (blue). The edges represent the existence (at least once within the period) of a movement between two holdings in the dataset, with their width indicative of the observed trading volume along this edge. The picture is courtesy of Gaël Beaunée.

5.2.2 Within-herd epidemiological-demographic dynamics

The BVD stochastic individual-based model represents each animal and its main characteristics : sex (Male, Female), race (Beef, Dairy), parity (P0, P1, P2, P3, P4, P5), age (in days), life-cycle and health-state. Parity P0 corresponds to nulliparous animals (i.e. never having given birth before). This includes females fattened or to be bred, and males. Parity P1 corresponds to females that have had a first calving, etc. Parity P5 corresponds to females that have had five births or more.

This model enables to represent most possible herd types: dairy, beef, breeding herds, fattening unit of veal calves and of young beef bulls.

5.2.2.1 Life-cycle and health-state dynamics

Regarding life-cycle there are seven possible categories. They are specified in table 5.2. Figure 5.4 presents the scheme of the life-cycle dynamics.

Name	Description
YJ	young juveniles. Female calves kept for breeding, under 6 months of age.
OJ	old juveniles. Female calves kept for breeding, between 6 months of age and the beginning of the first pregnancy.
G	gestating females.
NG	non-gestating females, i.e. period between calving and the start of the next pregnancy.
Fadult	fattened adult females, i.e. females in the period between last calving and culling.
YFbirth	young animals fattened from birth, i.e. animals under 6 months of age that will never calve (male and non-breeding female calves).
OFbirth	old animals fattened from birth, i.e. animals over 6 months of age that will never calve (male and non-breeding female animals).

Table 5.2: Life-cycle categories

Regarding health-states, there are five possible categories: P (persistently infected), M (protected by maternal antibodies), S (susceptible), T (transiently infected), R (recovered). Figure 5.5 presents the scheme for health-state dynamics.

The herd-specific parameters of the model were calculated from the dates of birth, death, calving and movements in the database, except for $(\phi_{Fbirth}^{B,Female})^{-1}$ and the proportions of culled animals. The distribution of $(\phi_{Fbirth}^{B,Female})^{-1}$ was set to be the same as $(\phi_{Fbirth}^{B,Male})^{-1}$. This is a good approximation and keeps the total population within reasonable limits. The proportion of culled dairy and beef animals were respectively set to 0.33 and 0.25 for all herds. Table 5.3 presents the definition and main descriptive statistics (mean, 10th percentile, median, 90th percentile) of these parameters. Details on their calibration can be found in the public repository specified in section 5.4.

Life-cycle and health-related parameters are defined in table 5.4.

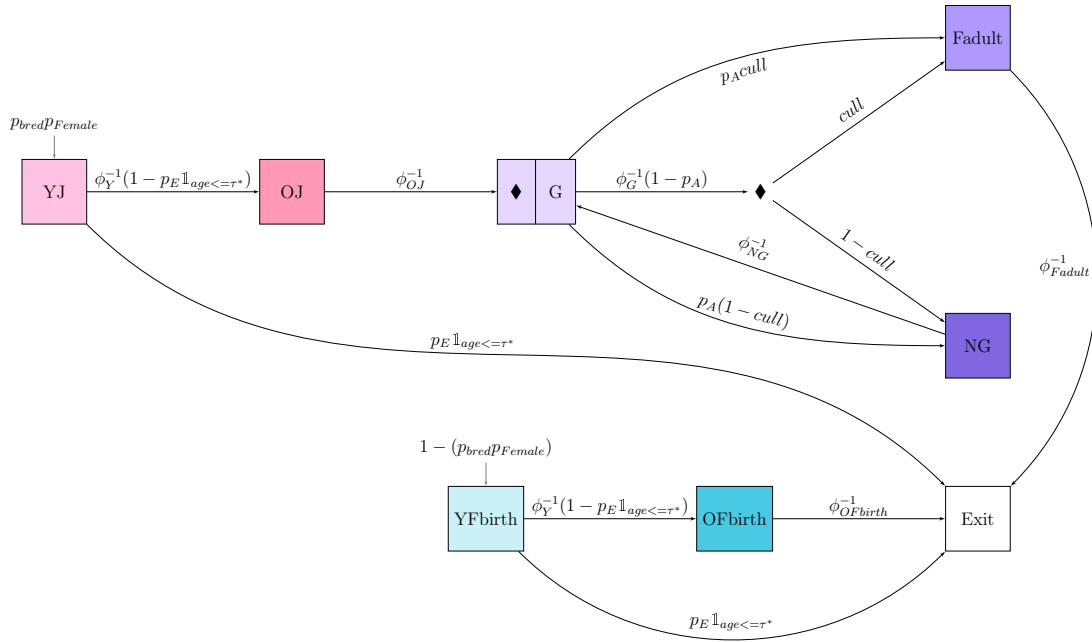


Figure 5.4: Life-cycle dynamics. Arrows represent transitions between life-cycle compartments. The coefficients on the arrows are either transition rates or transition probabilities. Compartment Exit represents culled breeding females or dead young cattle. Diamond indicates actions to perform: on entering G, variable cull and the health-state of unborn calf are defined; after the duration of gestation, i.e. if calf was not aborted, calf is born and cow's parity increases by one. See main text for parameter definition.

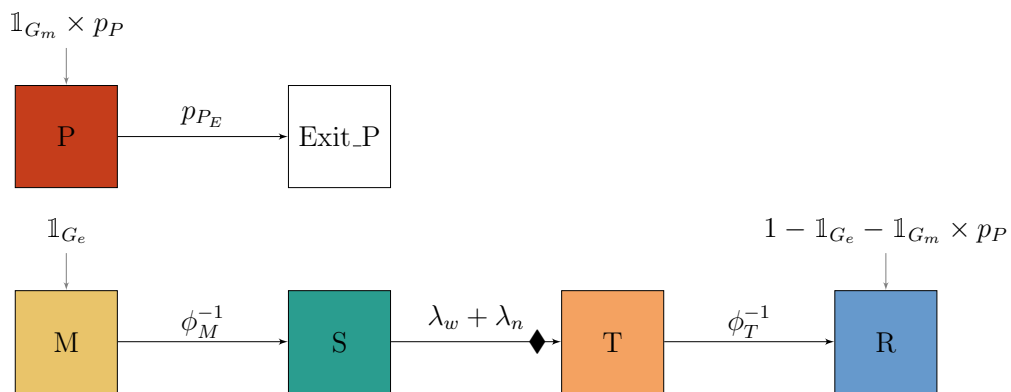


Figure 5.5: Health-state dynamics. Arrows represent transitions between health-state compartments. The coefficients on the arrows are either transition rates or transition probabilities. Compartment *Exit_P* represents dead persistently infected animals. Diamond indicates actions to perform: on transitioning from S to T, if animal is G, the health-state of the future calf can change or calf can be aborted. See main text for parameter definition.

Parameter	Definition	Mean	10th	Median	90th
p_{Female}	proportion of females born in the herd	0.49	0.25	0.49	0.75
p_{bred}^D	proportion of dairy female calves going to breeding	0.04	0.00	0.00	0.00
p_{bred}^B	proportion of beef female calves going to breeding	0.16	0.00	0.00	0.75
$(\phi_J^D)^{-1}$	duration for dairy animals in J	744.05	710.24	743.44	804.50
$(\phi_{NG}^D)^{-1}$	duration for dairy animals in NG	163.80	107.25	160.25	201.00
$(\phi_{Fadult}^D)^{-1}$	duration for dairy animals in Fadult	334.39	301.79	345.96	362.94
$(\phi_J^B)^{-1}$	duration for beef animals in J	789.86	715.20	790.02	859.15
$(\phi_{NG}^B)^{-1}$	duration for beef animals in NG	131.09	85.58	122.08	180.00
$(\phi_{Fadult}^B)^{-1}$	duration for beef animals in Fadult	321.49	292.78	327.11	346.86
$(\phi_{Fbirth}^{D, Male})^{-1}$	duration for dairy male animals in Fbirth	339.71	125.00	184.00	696.00
$(\phi_{Fbirth}^{D, Female})^{-1}$	duration for dairy female animals in Fbirth	558.11	106.00	367.71	1176.50
$(\phi_{Fbirth}^{B, Male})^{-1}$	duration for beef male animals in Fbirth	697.05	90.00	554.00	1530.00
$(\phi_{Fbirth}^{B, Female})^{-1}$	duration for beef female animals in Fbirth	697.05	90.00	554.00	1530.00
p_{cull}^D	proportion of dairy females culled	0.33	0.33	0.33	0.33
p_{cull}^B	proportion of beef females culled	0.25	0.25	0.25	0.25
p_E^D	proportion of dairy calves dying before 21 days of age	0.01	0.00	0.00	0.00
p_E^B	proportion of beef calves dying before 21 days of age	0.04	0.00	0.00	0.11

Table 5.3: Summary statistics of the herd-specific parameters of the model. All durations are in days.

Parameter	Description	Value	Source
ϕ_Y^{-1}	duration of the period when calves are with their mother in beef herds, determines duration in YJ and YFbirth	180	expert knowledge
ϕ_G^{-1}	duration of gestation	274	expert knowledge
τ^*	age before which natural mortality is applied	21	expert knowledge ¹
p_{A_e}	probability to abort when infected in early gestation (0 to 42 days)	0.8	Fray et al. 2000, Carlsson et al. 1989, McGowan et al. 1993
p_{A_m}	probability to abort when infected in mid gestation (43 to 150 days)	0.25	Fray et al. 2000, McClurkin et al. 1984
ϕ_M^{-1}	duration of the protection given by maternal immunity	150	Kendrick 1971
ϕ_T^{-1}	duration of the transient infection	7	McGowan et al. 1993
p_P	probability of vertical transmission in mid-gestation (43 to 150 days)	0.937	Fray et al. 2000, McClurkin et al. 1984, Kendrick 1971
p_{E_P}	probability of disease-related mortality, neglecting mortality at birth	0.00189	Houe 1993
β_{T_w}	transmission rate per transient animal within its group	0.03	Innocent et al. 1997 ² , Viet et al. 2004 ²
β_{P_w}	transmission rate per persistent animal within its group	0.5	Moerman et al. 1993, Viet et al. 2004
β_{P_b}	transmission rate per persistent animal to other groups	0.01	Ezanno et al. 2008 ²
β_{P_n}	transmission rate per persistent animal located in neighbours and that can access to pasture, with a density-dependent transmission	0.001	Qi et al. 2019 ²

¹ According to observed data (FCID) most of natural mortality occurs in the first weeks of age.

² No experiments nor field studies have allowed to estimate these parameters yet. Therefore, we choose similar values as in previous modelling studies.

Table 5.4: Life-cycle and health-state parameters of the model

Other parameters, that are based on the parameters in table 5.3 are:

$$\begin{aligned}
p_{bred} &= \mathbb{1}_D \times p_{bred}^D + \mathbb{1}_B \times p_{bred}^B \\
\phi_J^{-1} &= \mathbb{1}_D \times (\phi_J^D)^{-1} + \mathbb{1}_B \times (\phi_J^B)^{-1} \\
\phi_{OJ}^{-1} &= \mathbb{1}_D \times ((\phi_J^D)^{-1} - \phi_Y^{-1}) + \mathbb{1}_B \times ((\phi_J^B)^{-1} - \phi_Y^{-1}) \\
(\phi_{OFbirth}^{D, Male})^{-1} &= (\phi_{Fbirth}^{D, Male})^{-1} - \phi_Y^{-1} \\
(\phi_{OFbirth}^{D, Female})^{-1} &= (\phi_{Fbirth}^{D, Female})^{-1} - \phi_Y^{-1} \\
(\phi_{OFbirth}^{B, Male})^{-1} &= (\phi_{Fbirth}^{B, Male})^{-1} - \phi_Y^{-1} \\
(\phi_{OFbirth}^{B, Female})^{-1} &= (\phi_{Fbirth}^{B, Female})^{-1} - \phi_Y^{-1} \\
\phi_{OFbirth}^{-1} &= \mathbb{1}_D \times (\mathbb{1}_{Male} \times (\phi_{OFbirth}^{D, Male})^{-1} + \mathbb{1}_{Female} \times (\phi_{OFbirth}^{D, Female})^{-1}) \\
&\quad + \mathbb{1}_B \times (\mathbb{1}_{Male} \times (\phi_{OFbirth}^{B, Male})^{-1} + \mathbb{1}_{Female} \times (\phi_{OFbirth}^{B, Female})^{-1}) \\
\phi_{Fadult}^{-1} &= \mathbb{1}_D \times (\phi_{Fadult}^D)^{-1} + \mathbb{1}_B \times (\phi_{Fadult}^B)^{-1} \\
\phi_{NG}^{-1} &= \mathbb{1}_D \times (\phi_{NG}^D)^{-1} + \mathbb{1}_B \times (\phi_{NG}^B)^{-1} \\
p_{cull}^D &= \mathbb{1}_{P0} \times p_{cull}^{D, P0} + \mathbb{1}_{P1} \times p_{cull}^{D, P1} + \mathbb{1}_{P2} \times p_{cull}^{D, P2} + \mathbb{1}_{P3} \times p_{cull}^{D, P3} + \mathbb{1}_{P4} \times p_{cull}^{D, P4} \\
&\quad + \mathbb{1}_{P5} \times p_{cull}^{D, P5} \\
p_{cull}^B &= \mathbb{1}_{P0} \times p_{cull}^{B, P0} + \mathbb{1}_{P1} \times p_{cull}^{B, P1} + \mathbb{1}_{P2} \times p_{cull}^{B, P2} + \mathbb{1}_{P3} \times p_{cull}^{B, P3} + \mathbb{1}_{P4} \times p_{cull}^{B, P4} \\
&\quad + \mathbb{1}_{P5} \times p_{cull}^{B, P5} \\
p_{cull} &= \mathbb{1}_D \times p_{cull}^D + \mathbb{1}_B \times p_{cull}^B \\
cull &= \text{Bernoulli}(p_{cull}). \text{ Defined when a female enters state G.} \\
p_A &= \mathbb{1}_{G_e} \times p_{A_e} + \mathbb{1}_{G_m} \times p_{A_m} \\
p_E &= \mathbb{1}_D \times p_E^D + \mathbb{1}_B \times p_E^B
\end{aligned}$$

where $\mathbb{1}$ represents the indicator function. Then, for $x \in [D, B, Male, Female, P0, P1, P2, P3, P4, P5]$ (D for dairy, B for beef), $\mathbb{1}_x = 1$ if animal belongs to compartment x , 0 otherwise. Also, $\mathbb{1}_{G_e} = 1$ if the gestating female gets infected during early pregnancy (0 to 42 days), while $\mathbb{1}_{G_m} = 1$ if it gets infected during mid pregnancy (43 to 150 days).

The sex of a calf is determined according to p_{Female} at the moment of calving. Young male calves all enter the YFbirth compartment when born. Meanwhile, young female calves can either be kept for breeding (enter YJ) with a probability p_{bred} , or for fattening (enter YFbirth). Females in YJ can enter compartment OJ and animals in YFbirth can enter compartment OFbirth after spending ϕ_J^{-1} days in YJ and OJ, respectively. Before that time, they can die due to calf mortality (if their age is at most τ^* , i.e. if $\mathbb{1}_{age < \tau^*} = 1$) with a probability p_E .

Non breeding animals (males and non breeding females) are culled after spending $\phi_{OFbirth}^{-1}$ days in OJ. Meanwhile, females in OJ enter G after spending ϕ_{OJ}^{-1} days in OJ. At entering G, the health state of the calf is defined (which can later change during the pregnancy). Furthermore, the decision to cull the cow at calving or not is also attributed at the beginning of gestation. At the moment of calving (i.e. after ϕ_G^{-1} days if no abortion occurred) the parity of the breeding female increases by one, unless it is already P5. A female that will be culled enters compartment Fadult, and is culled after ϕ_{Fadult}^{-1} days. Females that will not be culled enter compartment NG, and reenter G after ϕ_{NG}^{-1} days. If a G female gets infected during early or mid pregnancy, the calf can be aborted with a probability p_A . Hence in such case, the female can either enter Fadult, if it will be culled, or NG, if it will not.

The force of infection within herd (λ_w), and the force of infection through the neighbours in the metapopulation (λ_n) are defined as:

$$\lambda_w = \lambda^{D,juv} + \lambda^{D,bred} + \lambda^{B,juv} + \lambda^{B,bred} + \lambda^{fat}, \quad (5.1)$$

$$\lambda_n = \mathbb{1}_{pasture} \times \mathbb{1}_{bred} \times \beta_{P_n} \times P_n, \quad (5.2)$$

where *bred* concerns the breeding group (all females in NG or G, and beef animals in YJ or YF-birth, as unweaned beef calves are located with their mother until weaning), *juv* concerns the juvenile group (all animals in OJ and dairy animals in YJ, as dairy unweaned calves are separated from their mothers in French farming systems), and *fat* concerns the fattening group (all females in Fadult, all animals in OFbirth, and dairy animals in YFbirth). Then, in λ_w (equation 5.1), each $\lambda^{x,y}$ is the force of infection of animals that are of type x and in group y , for $x \in \{D, B\}$ and $y \in \{juv, bred\}$. Similarly, λ^{fat} is the force of infection of animals in the fattening group. These are computed as:

$$\begin{aligned} \lambda^{D,juv} &= \frac{\mathbb{1}_D \times \mathbb{1}_{juv}}{N^{D,juv}} [\beta_{T_w} T^{D,juv} + \beta_{P_w} P^{D,juv} + \beta_{P_b} (P^{D,bred} + P^{fat} + P^{B,bred} + P^{B,juv})], \\ \lambda^{D,bred} &= \frac{\mathbb{1}_D \times \mathbb{1}_{bred}}{N^{D,bred}} [\beta_{T_w} T^{D,bred} + \beta_{P_w} P^{D,bred} + \beta_{P_b} (P^{D,juv} + P^{fat} + P^{B,bred} + P^{B,juv})], \\ \lambda^{B,juv} &= \frac{\mathbb{1}_B \times \mathbb{1}_{juv}}{N^{B,juv}} [\beta_{T_w} T^{B,juv} + \beta_{P_w} P^{B,juv} + \beta_{P_b} (P^{D,bred} + P^{fat} + P^{B,bred} + P^{D,juv})], \\ \lambda^{B,bred} &= \frac{\mathbb{1}_B \times \mathbb{1}_{bred}}{N^{B,bred}} [\beta_{T_w} T^{B,bred} + \beta_{P_w} P^{B,bred} + \beta_{P_b} (P^{D,bred} + P^{fat} + P^{B,juv} + P^{D,juv})], \\ \lambda^{fat} &= \frac{\mathbb{1}_B \times \mathbb{1}_{bred}}{N^{fat}} [\beta_{T_w} T^{fat} + \beta_{P_w} P^{fat} + \beta_{P_b} (P^{D,juv} + P^{B,juv} + P^{D,bred} + P^{B,bred})], \end{aligned}$$

where $N^{x,y}$ is the total number of animals that are of type x and that are in group y . $T^{x,y}$ and $P^{x,y}$ are defined in an analogous manner for transiently and persistently infected animals, respectively. N^{fat} , T^{fat} , P^{fat} are respectively the total number, the number of transiently infected, and the number of persistently infected animals in the fattening group.

In λ_n (equation 5.2), $\mathbb{1}_{pasture}$ equals 1 if current time is during pasture period, 0 otherwise. P_n is the number of P animals in neighbouring herds, among categories that go on pasture during pasture period (females in the breeding group and Fadult females), i.e.:

$$P_n = (P_n^{D,NG} + P_n^{D,G} + P_n^{B,NG} + P_n^{B,G} + P_n^{B,YFbirth} + P_n^{B,OFbirth}) + (P_n^{D,Fadult} + P_n^{B,Fadult}),$$

where $P_n^{x,y}$ is the number of persistent animals in neighbours, that are of type x and are in state y .

The health-state of a calf is initialised at the beginning of the mother's gestation: P for calves born to P mothers, S for calves born to S mothers, and in all other cases the calf will be M. Calves born with maternal antibodies are immune to infection during ϕ_M^{-1} days, when they enter the S compartment. Susceptible animals can become transiently infected either by contact with infected (T or P) animals within the herd, with a rate λ_w (equation 5.1), or by contact with P animals of neighbouring herds at a maximum distance of 2 km (expert knowledge) during the pasture period (which takes place every year from March 1st to November 1st (expert knowledge), with a rate λ_n (equation 5.2). See figure 5.6 for the distribution of the number of neighbours according to different values of the threshold distance.

If the new transiently infected animal is a G cow then the health-state of the future calf changes, if not aborted: M if infection occurs during early gestation, P with a probability p_P if it occurs during mid gestation, and R in all other cases.

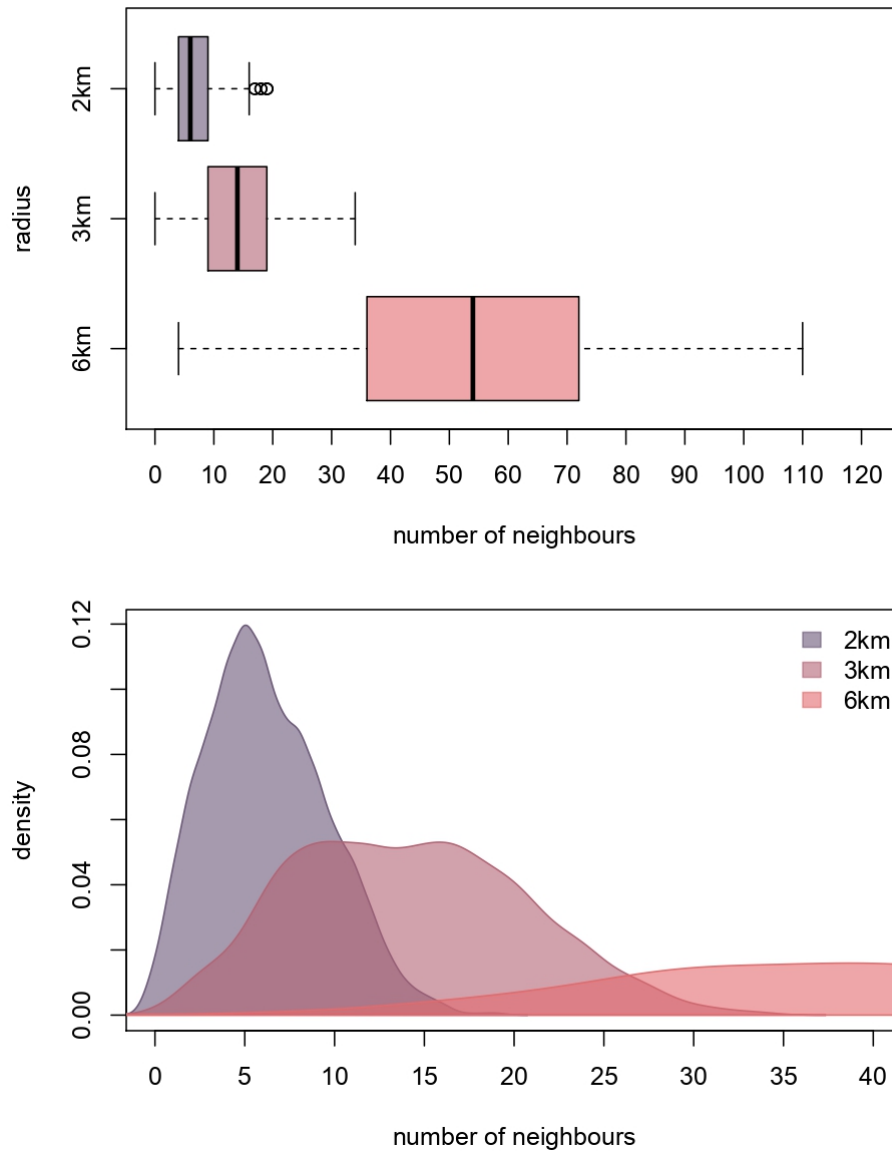


Figure 5.6: Distribution of the number of neighbours when the radius to define neighbourhood is fixed to 6km , 3km and 2km (for the selection of 3685 holdings). The bottom figure is a zoom of the distributions of the number of neighbours up to 40 neighbours.

5.2.2.2 Handling trade movements

Since the FCID does not contain information on life-cycle or health state, these elements must be initially assigned. Regarding health-state, unless specified differently in the initial conditions, the initial health-state of all animals is assigned to S, either at the beginning of a simulation or when entering the metapopulation from outside (i.e. there is no risk of virus entering the metapopulation from outside). To study infection dynamics, an initial number of infected animals (P or T) needs to be specified for some herds of the metapopulation (together with their life-cycle). For those herds, the health-state of the specified number of animals with the required life-cycle is updated from S to the specified state (P or T). If the specified herd does not have animals in the required life-cycle, there will be no initially infected animals in the herd.

Regarding life-cycle, it is assigned from available data of the animal (age, sex, breed, etc.) and

the parameters of the source herd (if it belongs to the metapopulation), or of the destination herd (if the source herd does not belong to the metapopulation). Males are sent to fattening: they are assigned to YFbirth if their age is less than ϕ_Y^{-1} , OFbirth otherwise. Females can either be sent to breeding or fattening. If information on a next calving date is found in the FCID, the female belongs to the breeding group. In that case, it can be a juvenile if its age is less than ϕ_J^{-1} (YJ if less than ϕ_Y^{-1} , OJ otherwise), or it can be G or NG (G if the time to the next calving is less than ϕ_G^{-1} , NG otherwise). If the animal is G, the gestation stage is obtained from the time of the next calving. If there is no information on a next calving, the female is sent to fattening. In that case, if the age of death is known and at most equal to τ^* , the female is assigned to YJ or YFbirth uniformly at random. If there is no information on age of death, or if such age is higher than τ^* , the female is assigned to a fattening category, depending on its parity and age. If there is no parity information or if parity is P0, the female can be YFbirth, if its age is less than ϕ_Y^{-1} , OFbirth otherwise. Finally, if its parity is at least P1, the female is assigned to Fadult. At the beginning of a simulation, the step to exit the life cycle state is assigned, considering the duration in the corresponding life-cycle, and that the animal entered the life cycle when entering the herd, except for G females (for which it corresponds to the time of next calving).

Selling is handled by identifying the animals that each herd is expected to sell at the end of that time step according to data. The right amount of animals is selected to be sold by choosing them uniformly at random among the animals in the origin holding that have the required characteristics. If the destination herd is within the metapopulation, animals to be sold are sent. Otherwise, they are simply removed from the origin herd. If the origin herd does not have enough animals to sell with the right specifications (breed, sex, parity, life cycle, age, step to exit the life cycle), such an animal is created in the destination herd. This also occurs if trade is coming from herds which are not within the metapopulation. In both cases, the incoming animal is assumed to be susceptible, so the associated infection risk is null. Movements whose origin and destination are both outside the metapopulation are ignored.

5.2.2.3 Epidemiological effect of vaccination

On the basis of the previously described model, we consider that a vaccine can have the effect of preventing vertical transmission of BVD virus if applied before the beginning of breeding. Then, the probability that a female vaccinated before being G produces a P calf if infected in mid-gestation is $p_P^v = p_P \times (1 - e_v)$ during mid-gestation. Other parameters are not modified, so animals can still get infected, be contagious, etc. In particular, if a G cow was not vaccinated right before breeding, the probability of producing a P calf in mid-gestation will remain to be p_P in mid-gestation.

In theory, e_v could be between 0 and 1, yet since available BVD vaccines are quite effective, so we set $e_v = 0.95$.

5.3 Farmer's decision-making on vaccination

If not applied before the breeding period, vaccination can have adverse effects, yet in the following we assume that this is never the case as farmers avoid these problems by applying the vaccine (if vaccination is adopted) at the right period. Furthermore, we assume vaccination only concerns breeding females.

5.3.1 Decision-mechanism

We adapt (Algorithm 4) the decision-mechanism presented in Cristancho Fajardo et al. 2021 (chapter 3).

Algorithm 4 Adapted exponential weighting stochastic mechanism with strategic imitation**Input:** $0 < p_v^{init} < 1; \kappa \geq 0; \rho \geq 0, B$: neighbourhood

- | | | |
|-----|--|--|
| 1: | $a_j(0) \leftarrow \text{Bernoulli}(p_v^{init}); j = 1, \dots, J$ | ▷ All farmers take an initial decision |
| 2: | for $t = \Delta_d, 2\Delta_d, \dots$ do | ▷ For other decision times t |
| 3: | for $j = 1, \dots, J$ do | ▷ Each farmer j |
| 4: | $C_{a_j(t-\Delta_d)}^j(t)$ | ▷ Observes the cost associated to his/her last decision (integrated over $]t - \Delta_d, t]$) |
| 5: | $B_j^*(t) \subseteq B_j(t)$ | ▷ Chooses a subset of neighbours among the set of eligible neighbours at time t |
| 6: | $\left\{ a_l(t - \Delta_d), C_{a_l(t-\Delta_d)}^l(t) \right\}_{l \in B_j^*(t)}$ | ▷ Observes neighbours' last decision and associated cost |
| 7: | $p_v^j(t) = \frac{p_v^j(t - \Delta_d) e^{-\kappa C_1^j(t) - \rho Z_1^{B_j^*(t)}}}{p_v^j(t - \Delta_d) e^{-\kappa C_1^j(t) - \rho Z_1^{B_j^*(t)}} + (1 - p_v^j(t - \Delta_d)) e^{-\kappa C_0^j(t) - \rho Z_0^{B_j^*(t)}}} \quad (5.3)$ | |
| | ▷ Updates the probability of vaccinating, where for $k \in \{0, 1\}$ | |
| | $Z_k^{B_j^*(t)} = \frac{n_k^{B_j^*(t)}}{n^{B_j^*(t)}} \sum_{l \in B_j^*(t)} C_k^l(t),$ | |
| | $n_k^{B_j^*(t)} = \text{Card}(\{l \in B_j^*(t) \text{ s.t. } a_l(t - \Delta_d) = k\}), n^{B_j^*(t)} = \text{Card}(B_j^*(t)), \text{ and the costs of the non taken decisions are equal to 0, i.e. } \forall i \in \{j\} \cup B_j^*(t): C_k^i(t) = C_{a_i(t-\Delta_d)}^i(t) \text{ if } k = a_i(t - \Delta_d), 0 \text{ otherwise}$ | |
| 8: | $a_j(t) \leftarrow \text{Bernoulli}(p_v^j(t))$ | ▷ Takes a decision |
| 9: | end for | |
| 10: | end for | |

At the first decision time, each farmer j decides to vaccinate or not, according to p_v^{init} , i.e. his/her initial probability of vaccinating, which is assumed equal for all farmers. From the second decision time, each farmer observes the costs associated to the previous decisions (his/her own and the one of his/her chosen neighbour(s) at that time), updates his/her probability of vaccinating (as a function of such decisions and costs) and takes a new decision based in this updated probability. Note that we consider the first decision to take place at time $t = 0$.

In the update of the probability to vaccinate, parameters κ and ρ , represent farmer's sensitivity to his/her own observed cost and to the cost observed by his/her selected neighbours. Then, the probability of vaccinating is updated proportional to the previous probability and an exponential function of a weighted sum. This weighted sum concerns the farmer's cost of vaccinating (equal to 0 if the farmer did not vaccinate at the previous decision time), whose weight is κ , and the mean cost observed by the farmer's neighbours that vaccinated at the previous decision time (weighted by the proportion of neighbours that vaccinated), whose weight is ρ . Then, this product is standardised to yield a probability. This exponential weighting system accounts for cognitive considerations regarding human decision-making, and the algorithm itself accounts for farmers learning. In particular, since farmers account for their neighbours decisions and observations, the algorithm accounts for social learning, which allows for farmers having specific behaviours such as strategic imitation: beginning to vaccinate/not vaccinate if neighbourhood vaccinates/does not vaccinate, or consciously free-riding, i.e. searching to benefit from neighbours' vaccination without having to bear the economic cost. See chapter 3 for a more in depth discussion on the form of this decision-mechanism.

Regarding the definition of B (the neighbourhood that could influence a farmer's decision on vaccination), we considered either the geographic neighbourhood B^g or the selling neighbourhood B^s . More specifically:

- *geographic* neighbourhood: based on geographic proximity and fixed in time. That is, the

neighbours are defined as a function of the maximum threshold distance used for defining pathogen transmission through pasture, i.e. 2km. With this value, the mean number of neighbours is 6.37, and there is a certain heterogeneity in this number across herds. Figure 5.6 shows the distribution of the number of neighbours when this radius is fixed to 6km, 3km, and 2km.

- *selling* neighbourhood: based on the trade network and therefore time-dependent. The selling neighbours of herd j at time t are the herds from which j has bought animals within the period $]t - \Delta_d, t]$.

At a given decision time, only herds that are active at that time (i.e. that have animals) can be eligible for being part of the neighbourhood of any other herd. Furthermore, only herds that have actually taken a decision at time $t - \Delta_d$ (i.e. that had gestating females within the period $]t - \Delta_d, t]$), could share this information with their neighbours. Hence, for a given definition of neighbourhood B (B^s or B^g), the set of eligible neighbours of herd j at time t , $B_j(t)$, is the set of neighbours of herd j at time t that have taken a decision at the previous decision step. That is, neighbours at time t that had a non-null quantity of gestating females within the period $]t - \Delta_d, t]$

Regarding the way each farmer j chooses the subset of neighbours $B_j^*(t) \subseteq B_j(t)$, we considered three options:

- *random*: choosing a neighbour uniformly at random at each decision time among the set of eligible neighbours at that time.
- *friend*: always choosing the same neighbour (initially chosen uniformly at random). If the chosen friend is not eligible at a given decision time, farmer decides only as a function of his/her own observed cost.
- *all*: using information of all eligible neighbours at each decision time.

Finally, we considered a *all-all* option ($B = B^s \cup B^g$), where farmer uses at the same time information of all his/her trade neighbours and of all his/her geographic neighbours.

Figure 5.7 presents the scheme of the full integrative model for vaccination decision-making in a context of BVD dynamic spread.

5.3.2 Economic-epidemiological cost

For the cost function we adapted the one in Cristancho Fajardo et al. 2021 (chapter 3), in the following way. The cost $C_{a_j(t-\Delta_d)}^j(t)$ that each farmer observes at time t , associated with decision $a_j(t - \Delta_d)$ is computed as:

$$C_{a_j(t-\Delta_d)}^j(t) = \frac{c_v(t - \Delta_d, t) \times a_j(t - \Delta_d) + c_i(t - \Delta_d, t)}{N_{\rightarrow G}(t - \Delta_d, t)}, \quad (5.4)$$

where

$$\begin{aligned} c_v(t - \Delta_d, t) &= cf_v + [cu_v \times N_{\rightarrow G}(t - \Delta_d, t)], \\ c_i(t - \Delta_d, t) &= r_1[N_{\rightarrow P}(t - \Delta_d, t) + N_{\rightarrow A}(t - \Delta_d, t)] \\ &\quad + r_2 N_{\rightarrow T}(t - \Delta_d, t), \end{aligned}$$

This standardised cost gathers the cost farmers pay to apply the vaccine to all the new gestating females within a given decision period, $c_v(t - \Delta_d, t)$, and the cost of the following new infections within the same period, $c_i(t - \Delta_d, t)$.

Indeed, the first term in the numerator of equation 5.4 is non-zero if the vaccine was applied ($a_j(t - \Delta_d) = 1$). First, $c_v(t - \Delta_d, t)$ considers a fixed cost of applying vaccination per herd cf_v ,

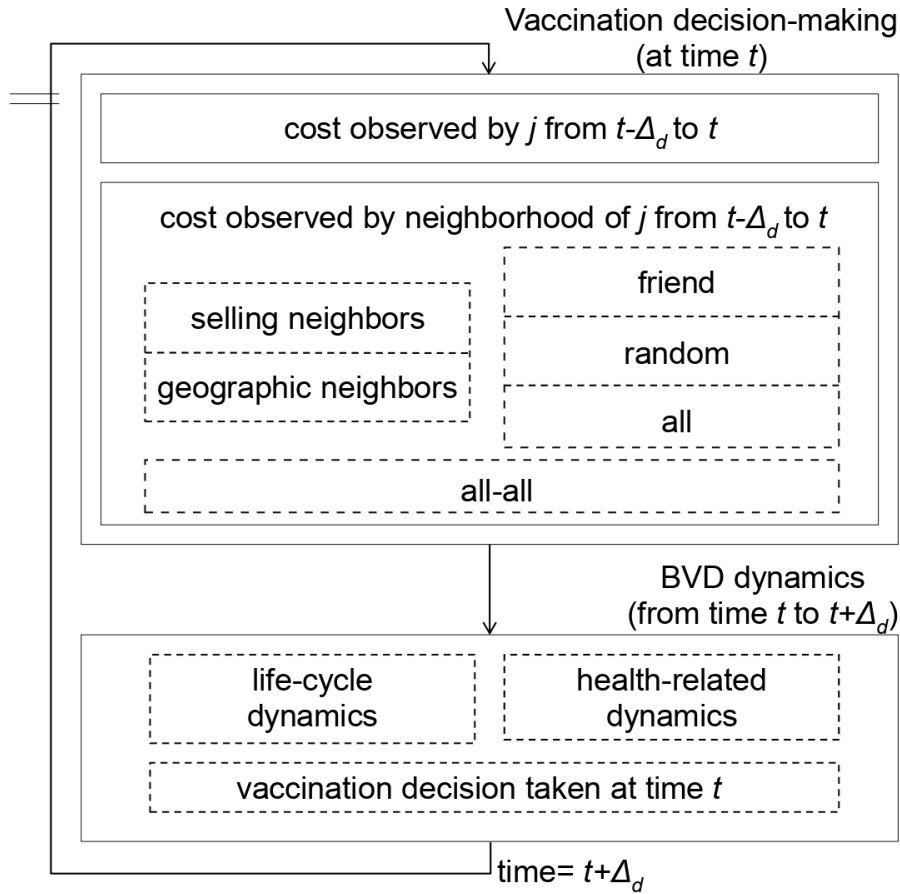


Figure 5.7: Scheme of the epidemiological-decision dynamical model for a herd j regarding BVD dynamics and farmers' vaccination decision-making.

which could correspond to the veterinary fees for vaccination. Second, cu_v , is the unitary cost of the vaccine per animal, which is multiplied by the number of vaccinated females for the period, if vaccination is decided. Neglecting breeding females that did not gestate within the period, this number is supposed to equal $N_{\rightarrow G}(t - \Delta_d, t)$, the number new gestating females over the period $[t - \Delta_d, t]$. Since we supposed that farmers apply their vaccination decision right before the breeding period, then all new gestating females over this period must have been vaccinated before becoming G , if vaccination was decided for that decision time.

Even if vaccines only have a direct effect on preventing vertical transmission, we can consider a broader indirect impact. Indeed, the presence of P animals in the herd can impact the number of new infections. On the one hand, these new infections can concern gestating females, which can cause more P newborns to be born, if infection occurs in mid-pregnancy, or can cause their abortion, if infection occurs in early or mid-pregnancy. On the other hand, there can be new infections of young animals (i.e. YJ or YF_{birth}), which can become transitory infected (T) by the contact with P animals. These animals are known to be at higher risk of enteric and respiratory disease (Stott et al. 2010). Hence, in $c_i(t - \Delta_d, t)$, r_1 is the monetary value of a healthy calf, which is completely lost if this calf is aborted or born but persistently infected (P). And r_2 is the loss associated with a new transient infection of a young animal.

Then $N_{\rightarrow A}(t - \Delta_d, t)$, $N_{\rightarrow P}(t - \Delta_d, t)$ and $N_{\rightarrow T}(t - \Delta_d, t)$ are respectively the cumulative number of aborted animals, the cumulative number of calves P , and the cumulative numbers of young T in the herd within the period $]t - \Delta_d, t]$.

Finally, to account for differences among the cost in different herds, related to the variation of the number of new G over the period, the cost is standardised by $N_{\rightarrow G}(t - \Delta_d, t)$.

The definition and values of economic and decision-related parameters are summarised in table 5.5. For the parameters in $c_v(t - \Delta_d, t)$, we arbitrarily supposed $cf_v = 50$ euros, as in Cristancho Fajardo et al. 2021, $cu_v = 5.11$ euros, the mean cost of one vaccine dose for four different brands of BVD vaccines (Arnoux et al. 2021). For the parameters in $c_i(t - \Delta_d, t)$, we set $r_1 = 800$ euros, i.e. roughly the price of a calf of the Charolais breed (Chambres d’agriculture de Bourgogne Franche-Comté 2022). Parameter r_2 is somewhat more difficult to fix as the precise loss associated to an animal becoming T is harder to estimate, although it is expected to be much lower than r_1 . Yet, the associated overall cost may be high, by the means of having many new T animals. We choose to set $r_2 = 0.01 \times r_1 = 8$ euros. We note that in Stott et al. 2010 a value of 3 £ is assumed for the cost of a immunocompromised calf, and 31 £ for a calf born with congenital defects, growth retardation, etc. Yet it is unclear how these expected costs were calculated. Regarding the values of decision-related parameters, we set p_v^{init} , κ and ρ arbitrarily. Finally, since farmers are assumed to decide whether to vaccinate females before gestation (roughly 9 months), the decision periodicity to be of 1 year, i.e. farmers take a decision for each year on January 1st.

Parameter	Description	Value	Source
cf_v	fixed cost of applying vaccination per herd	50 €	arbitrarily chosen
cu_v	unitary cost of the vaccine per animal	5.11 €	Arnoux et al. 2021
r_1	monetary value of a healthy calf	800 €	Chambres d’agriculture de Bourgogne Franche-Comté 2022
r_2	loss associated with a new transient infection of young animal (YJ or YFbirth)	8 €	arbitrarily chosen
p_v^{init}	initial probability of vaccinating	0.01	arbitrarily chosen
κ	farmer’s sensitivity to his/her own observed cost	1	arbitrarily chosen
ρ	farmer’s sensitivity to the cost observed by his/her selected neighbours	0.5	arbitrarily chosen
Δ_d	decision periodicity	1 year	expert knowledge

Table 5.5: Economic and decision-related parameters of the model

5.4 Simulation setting

The simulation step was set to 7 days, due to the short duration of transiently infections. Stochastic transitions determined by a probability were converted accordingly, when necessary.

We explored infection-dynamics, and the decision-dynamics during four years (from January 2010 to December 2013) for $(2 \times 3 + 1 + 1)$ scenarios. Simulations for the year 2014 were not included for the moment, as they exhibited problems regarding a strong decrease in the number of animals for some life-cycle categories. Such a problem is most probably related to the parameter calibration for that year, and will be looked into in the near future. The scenarios are given by the combination of the neighbourhood definition (geographic or selling), the way the neighbour is selected (random, friend, all) and the all-all scenario. An additional base-line scenario (*no_decision*) was considered for the dynamics without the vaccination decision component of the model.

Given the intrinsic stochasticity and the computational cost of the model, we considered 5 runs (as a first attempt, more runs will be further considered) for each scenario. For each run, we

considered that for a subset of 4% herds, 0.2% of their calves (YJ + YFbirth) were initially P. That is, $147 (=3685 \cdot 0.04)$ herds were initially infected with 1 P calf each. Indeed, the initial total number of calves (YJ + YFbirth) in the metapopulation for the year 2010 was 66549. Hence, the 0.2% P calves corresponded to 133 P calves in total, and each positive herd had 1 ($= 133/147$) P calf. We defined these initial conditions on the basis of the information from the animal health services (GDS: *groupement de défense sanitaire*) of Saône-et-Loire (*GDS de Bourgogne - Franche-Comté 2022*) for the year 2021, indicating 273 positive calves out of 184,930 tested (i.e. 0.2%), belonging to 109 of the 2900 herds tested. That is 3.8% ($=109/2900$) of the herds had at least one P among the tested. We note however, that there was no measurement of vaccination practices in such study, which can partly explain such a low prevalence for a first generalised screening.

The P calf for each initially infected herd was chosen among the YJ or YFbirth, beef or dairy animals of the selected herd, according to a multinomial distribution, where the probabilities of the P animal being a dairy/beef YJ/YFbirth animal were set proportional to the initial proportion of young animals in the chosen herd that were dairy/beef YJ/YFbirth. We remark that only herds that had initially at least 1 YJ or YFbirth animal were eligible to be initially infected, so as to ensure the same initially infected P animals, and the same initial number of infected herds.

The simulation code of the model and parameter calibration can be found in the public repository github.com/CristanchoLina/BVD_farmersdecisions.

5.5 Results

In the following, preliminary results over 5 runs are presented. As the number of runs is low, such results are not yet conclusive and additional explorations are needed.

Figure 5.8 presents infection dynamics in the metapopulation for each of the considered scenarios. Without the vaccination component (no_decision scenario) the pathogen spreads quite rapidly, and at the end of the fourth year 50% of herds have at last 1 P calf. The proportion of P animals among calves presents a seasonal behaviour, with waves whose peaks are attained roughly at the end of each year, from the second year. The highest peak (almost 0.12) occurs at the end of the third year. The final (at the end of the fourth year) proportion of P calves in the metapopulation is roughly 0.11.

Regarding the proportion of T animals in the metapopulation, figure 5.9 shows seasonal peaks, this proportion being almost 0.01 at the largest peak, attained between the second and third year, and concerning 35% herds. Yet, between the third and the fourth year, the proportion of T animals in the metapopulation attains levels in the previous year. Even more, the proportion of herds concerned by the presence of T animals increases (to 0.37) with respect to the highest level registered during the third year (0.35).

Regarding scenarios with decision-making, the first remark is that all of them achieve to decrease pathogen spread (P and T animals and herds) when compared to the no_decision scenario (figures 5.8 and 5.9). Unsurprisingly, the group of scenarios where the pathogen spread is better controlled is the one where farmers take into account all the information of their neighbourhood, through a weighted mean, constituted by the selling-all, geographic-all and all-all scenarios. For this group, the highest peak in the proportion of infected calves in the metapopulation (figure 5.8(b)) is decreased to less than 0.03, attained at the end of the second year. In the following years, this proportion almost consistently decreases, even if a small peak is still observed before the end of the third year, until reaching less than 0.005 at the fourth year. Of the three scenarios in this group, the worst control is obtained when decisions rely only on selling neighbours (selling-all scenario). In particular, figure 5.9 shows that it is the only scenario of this group for which a peak is clearly observed in the proportion of T animals (at almost 0.003), and in the proportion of herds with T animals (0.15), in the third year. Similarly, the proportions of P animals and of herds with P animals (figure 5.8) are higher than in the other two scenarios of the group, from the second year to the end of the fourth year. Indeed, the final values attained for these two proportions are respectively almost 0.002 and 0.05 higher for the selling-all scenario than in the other two

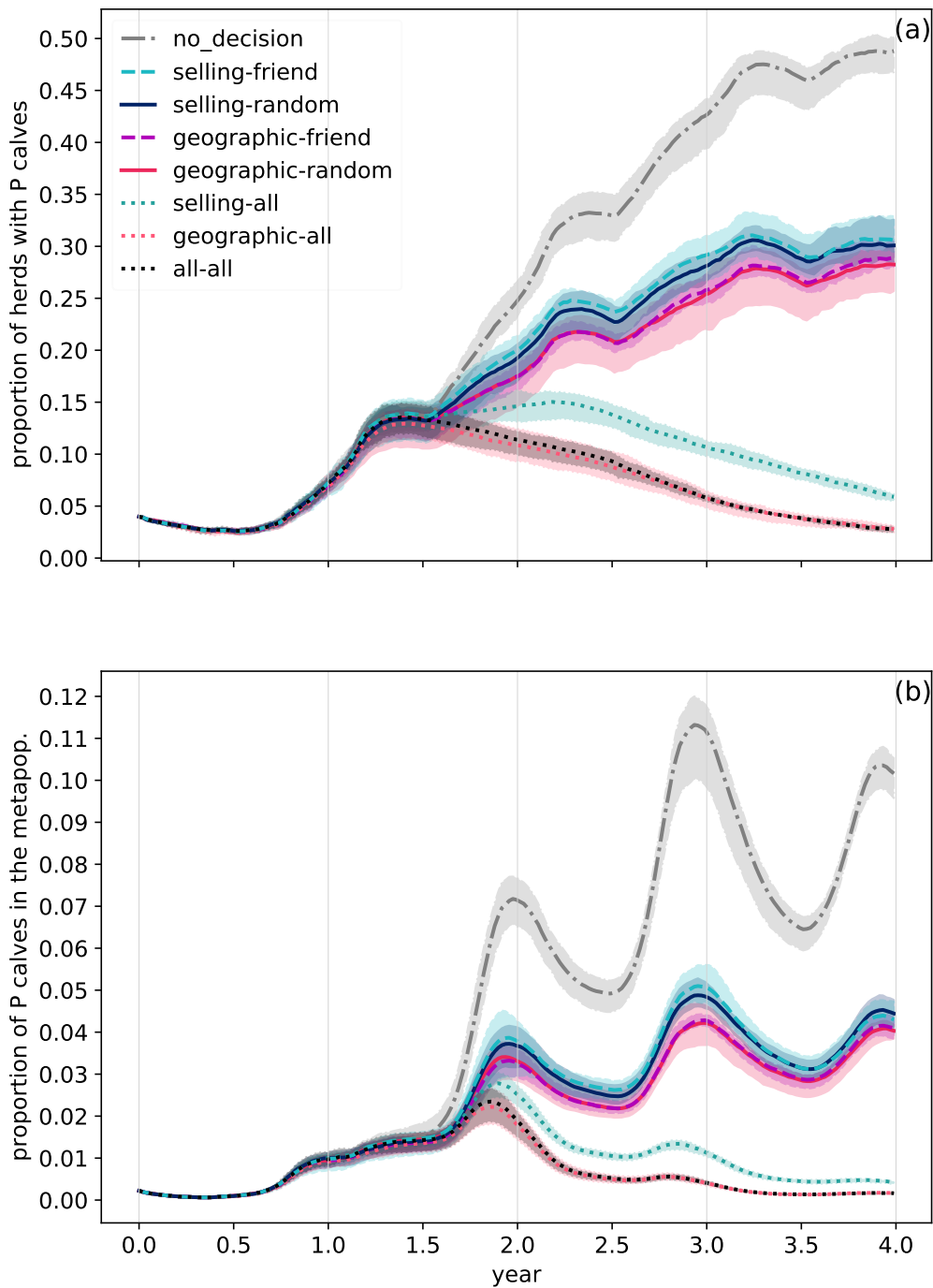


Figure 5.8: Dynamics of BVD virus spread in the metapopulation. Proportion of herds with P calves (a) and proportion of P calves (out of all calves) in the metapopulation (b). Each colour represents the neighbourhood (geographic or selling) and the way neighbours are selected for observation (random, friend or all), together with the all-all selection. Gray vertical lines represent decision times. Mean results and confidence bands (min to max) over 5 runs.

scenarios. Clearly, the performance of the all-all scenario is driven by that of the geographic-all scenario, as in the first both the geographic and the selling neighbours are taken into account. In particular, for both of these scenarios, the final levels of each of the fourth studied outputs are almost the same to the ones given by initial conditions. The group of scenarios where only the information of a single neighbour is taken into account at each decision time is found to be

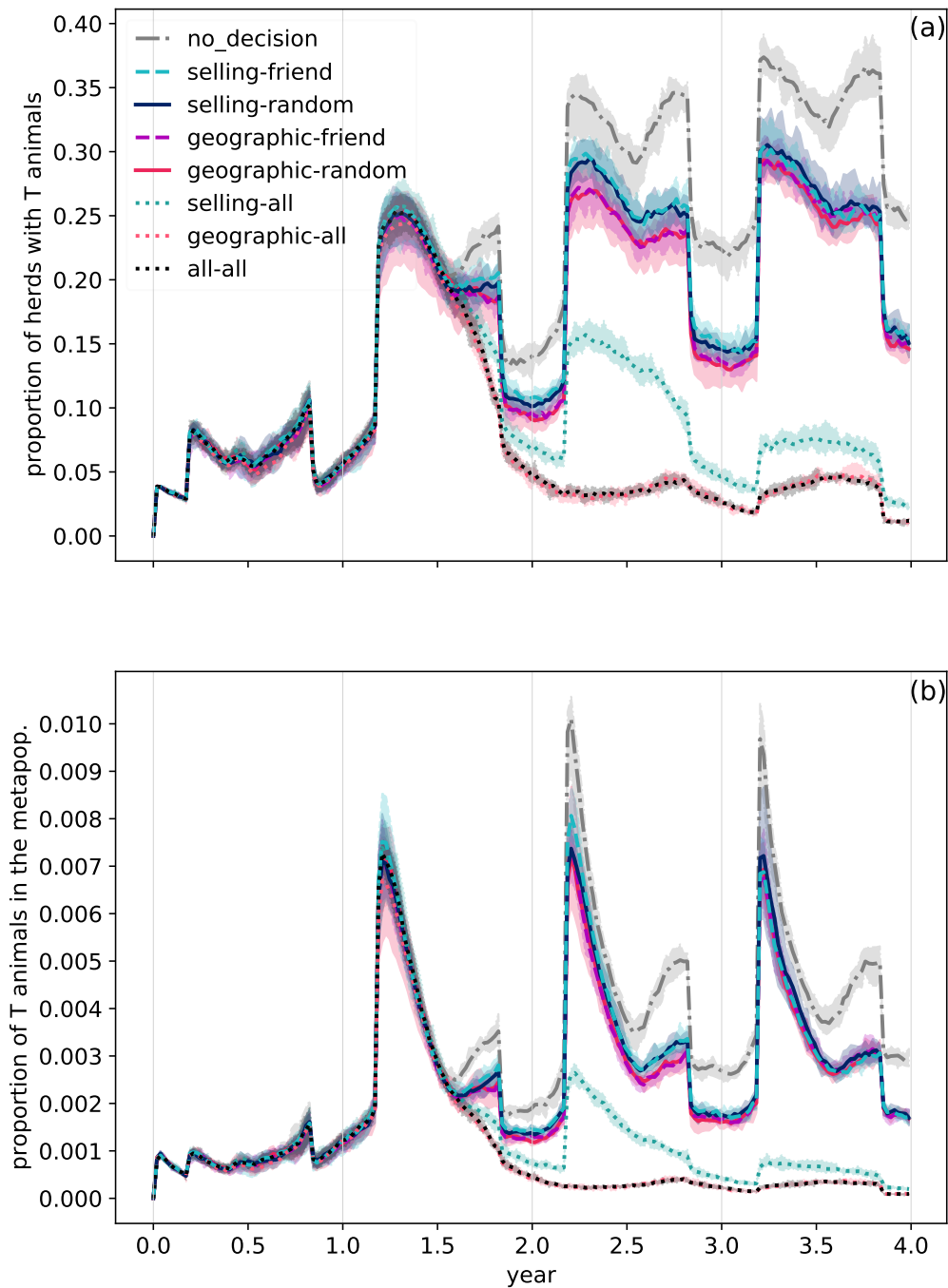


Figure 5.9: Dynamics of BVD virus spread in the metapopulation. Proportion of herds with T animals (a) and proportion of T animals in the metapopulation (b). Each colour represents the neighbourhood (geographic or selling) and the way neighbours are selected for observation (random, friend or all), together with the all-all selection. Gray vertical lines represent resource decision times. Mean results and confidence bands (min and max) over 5 runs.

in between the `no_decision` scenario and the previously mentioned group. With respect to the baseline scenario in the scenarios of this group the proportion of herds with P calves shows a lower increase, to attain roughly 0.3 at the end of the fourth year. For the proportion of P calves in the metapopulation, the highest peak is reduced to 0.05, attained at the end of the third year (figure 5.8). A similar behaviour is observed for the proportion of herds with T animals, and the proportion of T animals in the metapopulation, whose highest peaks are respectively reduced to

0.30 and 0.008 (figure 5.9). As in the scenarios where all information was taken into account, the geographic scenarios when using only one neighbour led to a better control of pathogen spread when compared to the respective selling scenarios, even if the difference is not as striking as when using all neighbours. For both the geographic and selling neighbours, randomly choosing the neighbour exhibited a better result in average, yet showed more variation.

Regarding vaccination dynamics, figure 5.10 shows that the proportion of herds that vaccinate increases from the initial value (roughly 0.01) to 0.25 in the geographic-random and the geographic-friend scenarios, to 0.27 in the selling-random and the selling-friend scenarios. The gap between these two sets of scenarios increases over time, and at the end of the simulation (which corresponds to the beginning of a fifth year), in the geographic scenarios (random and friend) the proportion that vaccinates is around 0.42, while in the respective selling scenarios this proportion attains 0.52. No difference can be appreciated regarding the way farmers choose their observed neighbour (random or friend), once the neighbourhood is fixed. Regarding the scenarios where all information from the chosen neighbourhood is used, in the cases where the geographic neighbourhood was considered (either alone or with the selling neighbourhood), the proportion that vaccinates increases to more than 0.98 from the first to the second decision time and stays at the same level throughout the years. Meanwhile, when considering solely the selling neighbourhood, this proportion increases to 0.7 at the beginning of the second year, and progressively increases to attain roughly 0.9 at the beginning of the fourth year. Such a proportion is slightly increased at the beginning of a fifth year.

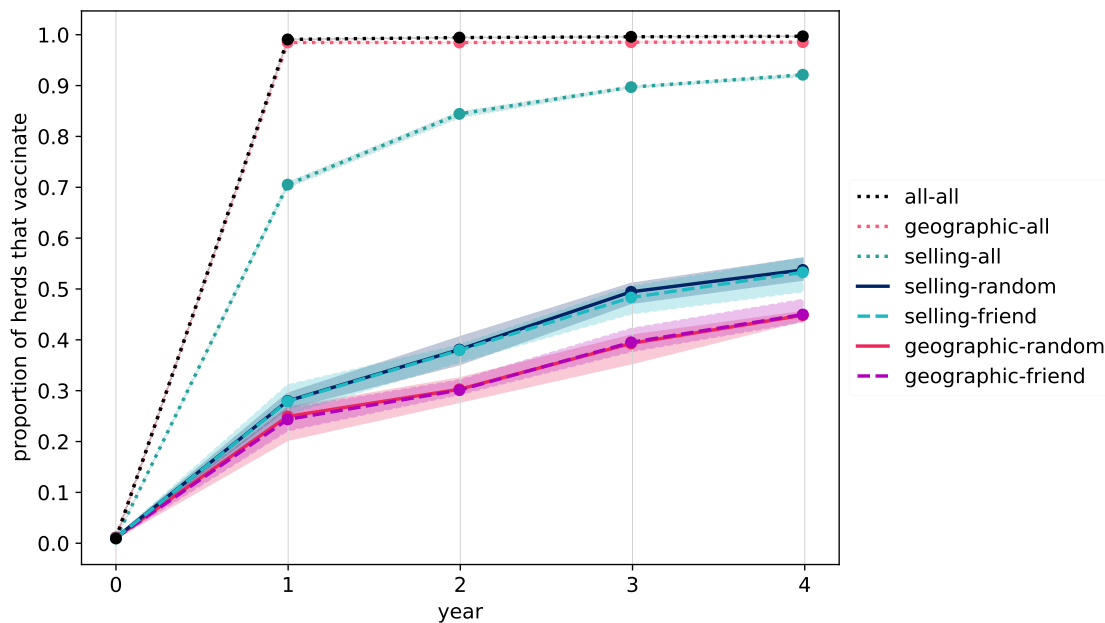


Figure 5.10: Dynamics of the proportion of herds that vaccinate. Each colour represents the neighbourhood (geographic or selling) and the way neighbours are selected for observation (random, friend or all), together with the all-all selection. Mean results and confidence bands (min and max) over 5 runs.

The exploration of the vaccination patterns (for one run only) shows that in the scenarios where farmers use their geographic neighbours to decide, the proportion that never vaccinates is higher by at least 0.05 with respect to the scenarios where farmers use the selling neighbourhood, and that the vaccination patterns where herds vaccinate most of the time are also more frequent in the selling scenarios (figure 5.11). Finally, figure 5.12 shows that, as expected, the proportion of herds that always vaccinate (excluding the first decision time) is higher (almost 1) in the geographic-all and the all-all scenarios, than in the selling-all scenario (0.7). Even if the pattern where herds vaccinate from the second decision time is the next more frequent one for the latter scenario (almost 0.15), the proportion of herds that never vaccinate is higher (0.06) when compared to the other two scenarios (less than 0.02).

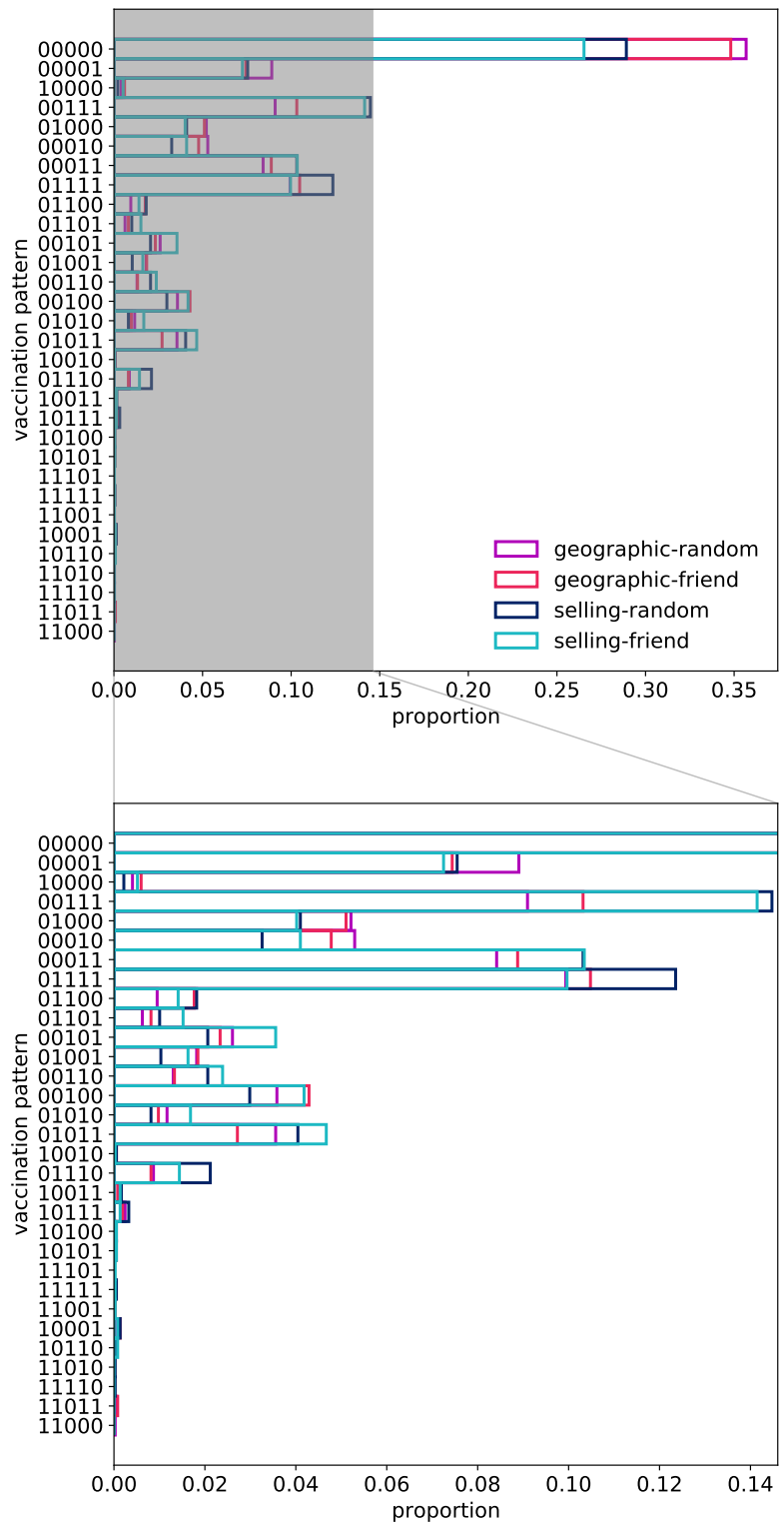


Figure 5.11: Vaccination patterns for the scenarios where farmers observe the information from only one neighbour (random or friend). 0 stands for not vaccinating, while 1 for vaccinating (e.g. the pattern 00010 concerns herds that only vaccinated at the beginning of the fourth year). The bottom figure is a zoom of the top figure that focuses on vaccination patterns that were observed at most in 15% of herds. Results for one run.

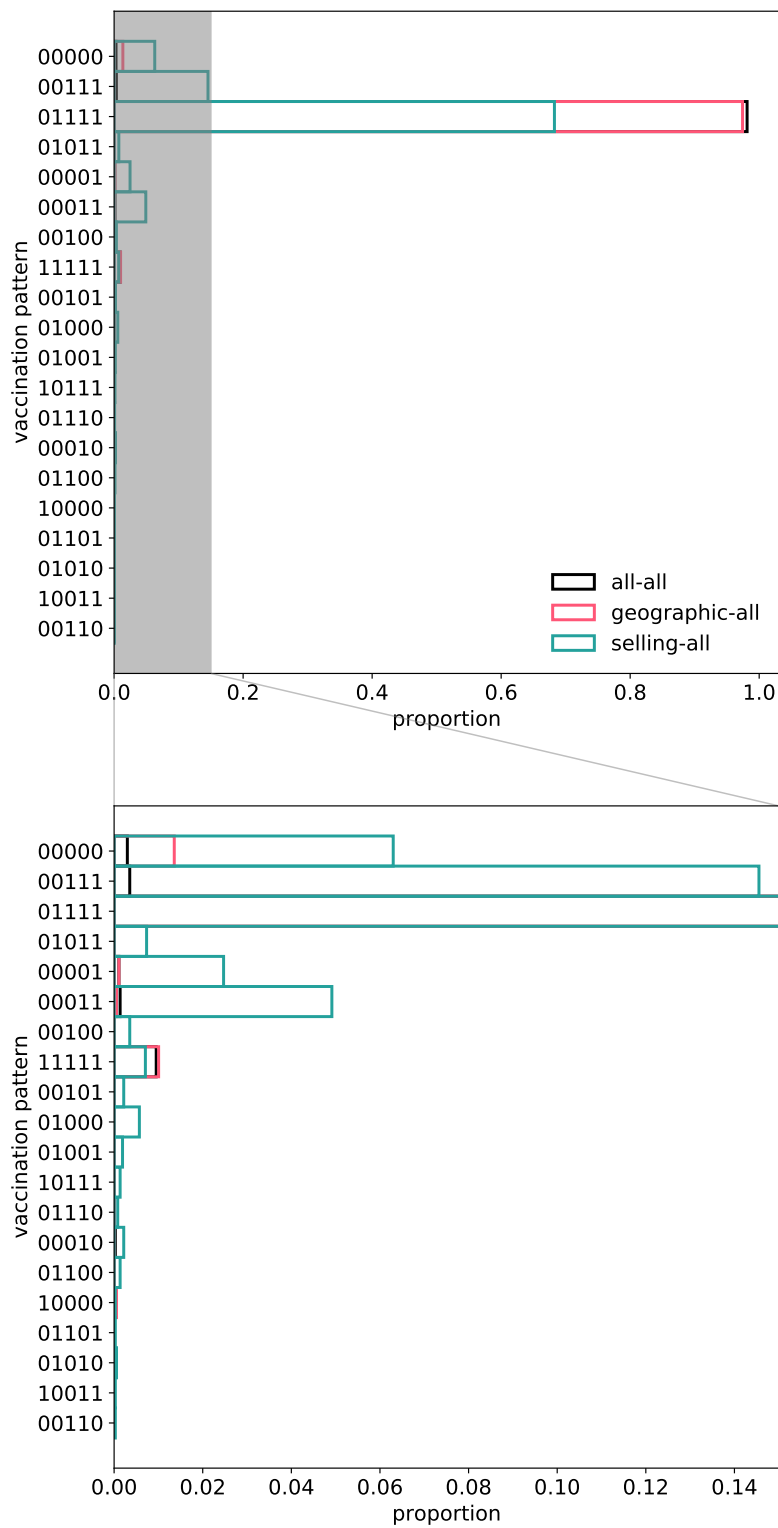


Figure 5.12: Vaccination patterns for the scenarios where farmers observe the information from all their neighbours in the chosen neighbourhood (geographic and/or selling). 0 stands for not vaccinating, while 1 for vaccinating (e.g. the pattern 00010 concerns herds that only vaccinated at the beginning of the fourth year). The bottom figure is a zoom of the top figure that focuses on vaccination patterns that were observed at most in 15% of herds. Results for one run.

5.6 Discussion

In this chapter, the generic integrative model proposed in Cristancho Fajardo et al. 2021 (chapter 3) is adapted for a real-life specific disease: the BVD. The adaptation concerned from one part, the economic-epidemiological cost on which farmers base their decisions. Indeed, the fact that BVD model used is structured according to host-heterogeneities (in particular age and life-cycle), and the vaccine has an effect on vertical transmission, allowed for a more detailed evaluation of the economic and epidemiological consequences of vaccination decisions. Furthermore, as BVD virus can spread through different ways (geographical proximity or the trade network), the initially proposed decision-mechanism was generalised to account for different neighbourhoods and ways of selecting neighbours that can influence farmers decisions regarding the adoption of a control measure (vaccination).

Preliminary simulations show that in all three scenarios where farmers observed at each decision time the previous actions and costs of all their neighbours (geographic and/or selling), BVD virus did not spread as much as in the scenarios where only one neighbour was observed at a time (figures 5.8 and 5.9). Even more, for the scenarios where the full information of the geographic neighbourhood was observed, the prevalence of the pathogen was the same at the one given at the beginning of simulations. Additionally, for the scenarios where farmers only choose a single neighbour from which they obtain information, the geographic neighbourhood appeared to be slightly better for pathogen control than the selling neighbourhood (figures 5.8 and 5.9). Quite surprisingly, this occurs although under the geographic neighbourhood the proportion of herds that vaccinate is almost always less than under the selling neighbourhood (figure 5.10), the proportion that never vaccinates is lower, and vaccination patterns where herds mostly vaccinate are less frequent.

Two main remarks can be made out of such observations. First, regardless of the chosen neighbourhood, using the full information of all neighbours at each decision time seems to have the most important impact for BVD control. Second, sharing information through the geographic neighbourhood seems to be consistently better for targeting herds that should get vaccinated to reduce BVD virus spread. Meanwhile with the selling neighbourhood, even in cases where more herds slightly get vaccinated, vaccination is not adopted by some herds that are key for the control of such spread.

It is to be stressed however, that such observations are not conclusive as many important aspects of the simulations have been neglected for the moment. To the very short term, the calibration of herd-specific parameters will be reevaluated, as some of these parameters presented unexpected ranges. In particular, the duration for animals in Fbirth presented some very high values that may unrealistically increase herd sizes in the simulations. For the proportion of culled and beef animals, although their values were set to realistic values, in the future they will be calibrated from data and will depend not only on the herd, but also on the parity of the female. For the moment, this is not the case as in the calibration procedure the value of this parameter could not be computed for several herds. This reevaluation of the herd-specific parameters may allow to stabilise the dynamics of the number of animals in each life-cycle category that are observed in simulations for the year 2014 (excluded in the present study). Once this issue is settled, the next point to address will be the low number of runs (only 5 runs were considered for each scenario). More simulations should be performed so as to consider the observations conclusive. And a numerical analysis should be done regarding the values of decision-parameters (p_v^{init} , κ and ρ) as they were arbitrarily chosen. Indeed, qualitative remarks may not hold under different numerical settings, and hence it will be of interest to explore the variation in results, as a function of the variation in the values of these parameters. Finally, we could further explore and describe results, for example through an association between geographic position and vaccination patterns.

This work gives insights for the control of BVD virus spread through vaccination, by focusing on a zone densely populated with Charolais cattle, a breeding system for which vaccination has proven to be a health measure of interest (Arnoux et al. 2021). The model represents how farmers voluntary vaccination practices can be determined by their own past experiences (vaccinating or not), and the information shared through their neighbourhood regarding other herds experi-

ences.

Regarding the limitations of this study (excluding those mentioned before, as they will soon be addressed), we may include the same one made for the original integrative model in (Cristancho Fajardo et al. 2021) (chapter 3) regarding the fact that in the model farmers perfectly observe costs associated to their own and their neighbours decisions. Additionally, farmers consider in the same way the observations of any of their chosen neighbours. In reality, farmers could give greater weight to the observations made by their 'closer' neighbours, i.e. account for their geographic neighbours proportionally to the distance to their herd, and account for their selling neighbours proportional to the amount of animals sold. Through this mechanism, the model could consider an additional strategic behaviour of farmers, as they could better account for how the vaccination practices of their neighbours can impact the health status of their own herd. Finally, regarding the BVD virus spread model itself, we assumed that the metapopulation was initially infected, but that there was no risk of introduction of the virus from outside the metapopulation. As the area modelled is not that large, this assumption could be revisited and exposure to trade and geographic neighbours from other areas could be accounted for.

Chapter 6

Conclusions and perspectives

Contents

6.1 Main contributions	143
6.1.1 Integrative model for pathogen spread over an animal trade network accounting for farmers' dynamic decisions regarding the adoption of a health measure	144
6.1.2 Strategies for the dynamic resource allocation of a limited resource for controlling disease spread on a metapopulation network	144
6.1.3 Study on farmers' vaccination decisions integrated in a model of BVD's spread at a large scale	145
6.2 Perspectives	145
6.2.1 On decentralised decision-making	145
6.2.2 On centralised decision-making	147
6.2.3 On decentralised and centralised decision-making	148
6.3 Conclusion	149

This thesis is concerned with the dynamics of infectious diseases spreading on a large metapopulation network, and the dynamic decision-making processes there can exist regarding the adoption of control measures to contain such spread, particularly in a context of livestock endemic diseases spreading through trade. In particular, in such a multi-agent framework, there can be dynamic decisions that are taken locally (i.e. by the agents of the system themselves), and/or dynamic decisions taken in a centralised manner (i.e. by a central authority).

Although in epidemiology, a large diversity of mechanistic mathematical models are used for describing and predicting the spread of infectious diseases, the precise study of decision-making in this context is most often neglected, particularly for the management of livestock diseases (Hidano et al. 2018). On the one hand, the behaviour of the individuals with a decisionary aptitude in the system where the pathogen spreads is usually not taken into account at all. When it is, the focus is often on how a certain decision or behaviour could impact disease spread, but not on how the behaviour itself can emerge as a function of the infection dynamics (Wang et al. 2016). On the other hand, regarding decisions made by a central entity for an optimal management of the disease, their study most often concerns unstructured and/or relatively small populations, as the solution to a most general problem is computationally intractable (Nowzari et al. 2015).

6.1 Main contributions

In the context described above, this thesis proposed two methodological contributions, and one contribution regarding an application study. In the following, each of these contributions is briefly summarised.

6.1.1 Integrative model for pathogen spread over an animal trade network accounting for farmers' dynamic decisions regarding the adoption of a health measure

The first contribution (Chapter 3) concerned an integrative model of pathogen spread on an animal trade metapopulation network, that accounts for farmers' voluntary decision-making regarding the adoption of a health control measure against infection in their own herd. The epidemiological component is built upon a stochastic SIR intra-herd model with demography, that accounts for inter-herd animal transfers. The decision-making component consists in setting a dynamic decision problem under uncertainty, and proposing a dynamic decision-mechanism that accounts for farmers' stochastic behaviour, learning, and strategic imitation dynamics. Thus, there is a dynamic feed-back loop between the two components, as the control decisions are function of the observed disease spread, and the disease spread is impacted at its turn by the decisions previously taken. The model was detailed and simulated for a specific measure (vaccination). The infection and vaccination decision dynamics for contrasted scenarios were studied, and the robustness regarding the values of the model's parameters was assessed. Experiments showed that decision related parameters highly impact vaccination decisions, and therefore the course of the disease spread.

Previous studies (Bauch 2005; Fu et al. 2010) were pioneers in accounting for real-life aspects of human behaviour regarding vaccination against human diseases, such as stochastic behaviour and imitation. Yet, such type of study has rarely considered an explicit learning process for representing human-decision making. Furthermore, these works had not yet been adapted for the specificities related to the management of livestock diseases. Hence, the first model proposed in this thesis is a significant step towards accounting for human behaviour in epidemiological models. Additionally, its generic formalisation allows it to be relevant for a variety of epidemiological models or measures. Potential extensions and uses can be further facilitated by the public availability of the simulation code.

6.1.2 Strategies for the dynamic resource allocation of a limited resource for controlling disease spread on a metapopulation network

The second contribution (Chapter 4) concerned the optimisation of a social planner's decisions regarding the allocation of a limited resource among the sub-populations of a metapopulation network. This problem was studied for a livestock disease spreading on an animal trade network. Again, for each sub-population (herd) of the network, a stochastic SIR intra-herd model with demography was used, accounting for inter-herd animal transfers and grounded on data for its non epidemiological components. The contribution is based on an extension to this metapopulation framework of an approach initially proposed for networks without epidemiological-demographic within-node dynamics (i.e. a node of the network corresponded to an individual, not a sub-population). The approach consists in deriving analytic scores minimising an approximated variant of an objective function and then dynamically allocating the resource to each sub-population according to their ranking given by the sub-populations scores, and the limited amount of resource. Such optimised scoring functions were obtained for two different objective functions, and for two different types of resources (vaccine and treatment). Their performances in containing disease dynamics and their relevance in practice for allocating the resource were assessed, along with the ones of several heuristic scores. It was shown that the analytic scoring functions can allow for a significant control of the pathogen spread, although they are not always truly optimal in practice. However, they perform in general much better than the scoring functions based only on information on demographic or trade data, as in most of the evaluated scenarios, information on epidemiological status appeared as necessary to achieve disease eradication.

This work represents an advance into optimising socially planned interventions in a complex system of disease propagation on a large metapopulation network. Indeed, related studies refer either to small metapopulation networks (e.g. Mbah et al. 2011), or to large non metapopulation

networks (e.g. Wijayanto et al. 2019). The same approach could also be used for different epidemiological compartmental models in the metapopulation framework. Moreover, other measures can be studied leading to other specific greedy scores, to be obtained. Finally, like for the work in Chapter 3, the simulation code was made available in a public repository.

6.1.3 Study on farmers' vaccination decisions integrated in a model of BVD's spread at a large scale

The final contribution (Chapter 5) consists of an application and extension of the integrative model in Chapter 3, to a specific real-life disease (BVD), supposing that farmers' decision-making can be influenced in other ways than the one initially considered. In this application, the methodological framework was extended to consider that farmers can make decisions as a function of the decisions and observations of their neighbours, not only in the animal trade network, as assumed in the original generative model, but also in their geographical neighbourhood (i.e. based on geographical proximity to other farms). Furthermore, several alternative ways on how the neighbours are chosen were evaluated. Finally, this ongoing work highlights how the components of the integrative model can be adapted to the study of a different type of epidemiological model (an individual-based model) for the control of a real-life disease, in this case BVD, through the adoption of a health measure (vaccination) that has been proven to be of interest for certain contexts (Arnoux et al. 2021).

6.2 Perspectives

The assumptions made throughout this thesis are accompanied by some advantages and limitations. Such limitations can be of inspiration for subsequent works as they can open up to perspectives of interest. In the following, certain perspectives are described. Some of these perspectives concern elements mentioned in the discussions of previous chapters, that are here summarised or deepened, while other perspectives are more detailed as they can have a wider opening.

6.2.1 On decentralised decision-making

First, regarding the works relative to farmers' decision-making (Chapters 3 and 5), four points could be addressed: observation of costs, variation of the decision algorithm, parameter robustness, and the consideration of more relevant health measures.

6.2.1.1 Noisy observation of costs

The epidemiological and farmer's decision-making components are linked through an economic cost that is function of the epidemiological dynamics following the previous decision, and on which the following decision is taken. For updating their decisions, it was supposed that farmers could perfectly observe the costs associated to their own previous action and the previous actions and associated costs of their neighbours, either by observing the number of new infected animals, or by directly observing the associated financial loss.

However, it can be a strong assumption that the observed cost precisely corresponds to the loss associated with the true number of new infections within the period. In reality, even if farmers could observe an increase/decrease in the number of infected animals (for example, through symptomatic animals, which would already be a noised signal of the number of infected animals), they would probably not be able to exactly know the number of new infections in the herd, i.e. distinguishing infected animals whose infection occurred within the herd, from imported infected animals (whose health status could reflect a decision of the origin herd), and also

because most of the time infection remains undetected. Furthermore, farmers most probably only have a rough idea of the observed costs of their neighbours.

Relaxing this hypothesis could impact the observed decision patterns and the local or overall disease dynamics. To this end, the cost could be reformulated to be dependent not on the number of true new infections, but on the increase of infected animals with respect to the previous decision period, i.e. assuming that farmers perfectly observe prevalence, not incidence. Alternatively, an appropriate additive noise could be accounted for in observed number of infections (or in the observed increase in the number of infected animals) within the period. Such an additive noise could be, for example, distributed as a binomial random variable whose parameters would need to be calibrated.

6.2.1.2 Variations of the decision-making algorithm

An algorithm intended to represent farmers' decision-making process was proposed and evaluated within the context of both a theoretical and a real-life infection relying on real-life aspects of human decision-making such as learning and strategic imitation. Although the algorithm considers heterogeneity in decision outcome, it has a structural homogeneity (i.e. all farmers have the same decision-making process). In particular, for vaccination, the same values for decision-related parameters were assumed: same initial probability of vaccinating, same sensitivity to their own observed costs and same sensitivity to the observed costs of neighbours. Aside from the fact that it is actually unknown to what extent this algorithm reflects the true decision process of a typical farmer, considering that such a typical farmer is a representative decision-maker for all the farmers in the trade network might be a strong assumption, as farmers could have different values of decision-related parameters, or even alternative decision-making processes. Yet, instead of considering more than one decision-making mechanism present in the metapopulation, a simple option to address this limitation would be for instance to assign, on the basis of the proposed decision-making mechanism, a low probability of applying the measure to half of the farmers, and a high probability to the other half, which could allow for additional explorations on the decision patterns. More generally, the hypothesis could be alleviated by considering different decision profiles, using a distribution for each of the three previously mentioned parameters.

Another assumption in the farmers' decision process was to consider in the same way any neighbour whose observation of actions and associated costs is explicitly included in the decision. Indeed, the neighbours are chosen uniformly at random. Furthermore, the sensitivity to the results of any neighbour is the same.

Although this seems an acceptable hypothesis as a starting point, more potentially realistic assumptions could be made. In particular, the probability of selecting a given neighbour in the trade network could be proportional to the amount of animals purchased from the neighbour. Such a weight could also be considered for the parameter regarding the sensitivity to neighbour's observed costs. Hence, farmer's decision-making would be not only based on the costs observed by his/her neighbours, but would also, in a way, take into account the impact that the neighbours' decisions may have on his/her own herd as a function of the amount of animals purchased. That is, by integrating the potential risk of importing infected animals, farmers might focus on the health practices of the herds from which they purchase most animals. A similar approach could be taken for the geographical neighbour, weighting by the geographical distance between herds. This could, in principle, strengthen the impact that a different neighbourhood, or a different way of choosing the neighbour, may have.

6.2.1.3 Robustness to parameter values

The robustness of the initial integrative model presented in chapter 3 to the parameters was assessed through several sensitivity analysis experiments. However, the adaptation of the cost function for vaccination decisions on BVD involved new parameters with respect to the original generic model. The values of some of the parameters (the decision-related ones) were arbitrarily

chosen. Therefore, results and conclusions may be different for other values of such parameters. Additionally, the impact of the variation of the tested parameters of the model might have changed in the presence of new parameters (Saltelli et al. 2008).

Hence, a new sensitivity analysis should eventually be performed for this new model, in particular on the new parameters, in order to evaluate the qualitative impact of their variation on the conclusions of this work.

6.2.1.4 More relevant field measures

The generic formulation of the integrative model allows it to be adapted to different epidemiological state structures, including other compartmental models, as evidenced by the BVD application study. Furthermore, the adaptation could also concern the use of other types of control measures.

However, only vaccination was explicitly modelled and simulated. Although this seemed the most important measure to look at for endemic unregulated diseases, other measures more relevant in the field could be evaluated. In particular, for the application study on BVD, a classical measure is to test animals at birth in order to detect persistently infected calves on ear tag sampling. Such a policy could be modelled (like vaccination in chapter 5) by making a decision at the beginning of each period of time, and applying the measure at each birth occurring within this period. Another measure consisting in changing pasture practices could be considered: delaying the pasture period according to the economic losses related to infected animals. Considering such a measure could widen the use of this model to other administrative areas, where vaccination might not to be an appropriate measure for controlling BVD, e.g. the French region of Brittany, mainly populated with dairy cattle, as BVD has been controlled in dairy herds without using vaccination. Additionally, as opposed to countries that use vaccination as the BVD eradication plan (e.g. Germany, Belgium, Ireland, Scotland), some countries (e.g. Switzerland, Austria) have chosen an approach without vaccination (Moennig et al. 2018), since their policies are based on the detection of BVD through serology, and vaccination interferes with serological results.

Depending on the measure, the adaptation of the decision-making model, and in particular of the economic-epidemiological cost, may not be so straightforward. For example, for test-and-cull, bulk tank milk could be followed to highlight new virus circulation in a herd, and then a decision could be taken as a function of such measurement. This could also imply that decisions are not taken at regular time steps, but following a change in herd status based on such screening (see section 6.2.3.2 for a discussion on this point).

6.2.2 On centralised decision-making

Second, concerning the work on centralised dynamic resource allocation on a large metapopulation network (Chapter 4), three aspects could be considered: the availability of health information, the observation of a sub-network, and the dynamic nature of the network.

6.2.2.1 Noisy or partial health information

For the resource allocation problem, it was assumed that the social planner could exactly compute the scores (heuristic or optimised) based on real-time health information. However, in the field, there is rarely such detailed information regarding the health-status of the farm, particularly for endemic animal diseases. Hence, such scores may not be possible to be observed, in particular those based on real-time intra-herd incidence or prevalence.

To overcome this limitation, an option on the basis of the work presented in this thesis would be to build a composite score as a linear combination of uncorrelated scoring functions with no necessarily highest performance but that are likely to be observed.

6.2.2.2 Sub-network

Instead of having partial information, like in the previous point, it may be possible to have no information at all for some sub-populations (health-related or not), and the resource may be hard to allocate for some of them (deployment constraints). A less restrictive hypothesis would be to assume that the social planner can only observe and access a sub-network for the resource allocation. This perspective could be addressed in the light of the work in Fekom et al. 2020 regarding a *restricted dynamic resource allocation (RDRA)* problem, in which at each time the social planner can allocate the resource only among random sample of reachable nodes of a (non metapopulation) network.

6.2.2.3 Dynamic network

In chapter 4, new optimisation based scores were found for allocating a resource (vaccine or treatment) among the sub-population of an arbitrary metapopulation network.

However, even if these new scores were evaluated through simulation on a dynamic metapopulation network, they were built on a static aggregated view of such network. Although this may be seen as an advantage, as their performance does not depend on real-time information on the topology of the network, scores that would actually take into account the time-varying structure of the network could potentially achieve better performances than the ones observed for the derived scores. The possibility of finding such scores could be explored with a new formal analysis. If obtained, such scores could be compared by simulation to the scores initially derived and give insights regarding the importance of the availability of real-time topological information for resource allocation in a metapopulation network.

6.2.3 On decentralised and centralised decision-making

6.2.3.1 Coupled centralised-decentralised decision-making

It would be of interest to revisit the assumption of separating the social planner's and the farmers' decision-making. This could be done by considering in the resource allocation problem, that farmers' behaviour is not limited to doing what the social planner finances. The objective would be to couple the farmers' decision model, and the social planner's resource allocation problem and proposed strategies. For a complete coupling, the planner should take into account farmers' behaviour and vice-versa.

Regarding the social planner, a possible approach is to consider that a score-based strategy is applied only on a group of sub-populations. This is not necessarily the same framework as the restricted dynamic resource allocation (RDRA) problem, as here the social planner can observe and have access to the whole metapopulation network, yet decides not to allocate the resource to some sub-populations so that farmers control the pathogen spread themselves, to a certain extent. There are several choices for the eligible group for resource allocation, yet the implications of one or another should be carefully studied (Zhang et al. 2013). A first simple approach, recently taken for a human theoretical disease in Meng et al. 2022, would be to allocate the resource on some sub-populations based on a predefined score, and leave the rest of the sub-populations decide whether to apply the measure or not, i.e. not actually coupling the decision-making processes but only assuming they occur in a certain order. As an alternative, it could be assumed that the social planner would only allocate the resource to the sub-populations that did not apply a control measure for themselves. However, in both options, counter-productive effects should be taken into account, as such options can represent an incentive for farmers not to apply the measure for themselves. Indeed, farmers could anticipate that they can be indirectly or directly protected from infection through the resource allocation performed by the social planner, even if they decided not paying the price of applying the measure. Another option could be to allocate the resource for sub-populations that did apply the measure (at the previous decision period),

hence rewarding them. This could, in principle, be an incentive for farmers to further apply the measure, but could cause overspending resources (e.g. the cost of increasing vaccination could be greater than the cost of infection) (Kuga et al. 2019). Finally, a more developed strategy for resource allocation could be based (apart from the allocation score) on the full history of information regarding control related practices of the farmers, similarly to what is proposed for a theoretical model of a human seasonal disease in (Ding et al. 2018).

A final important aspect to take into account on this matter is how to define decision times. For example, if farmers and the social planner make simultaneous decisions, the social planner might allocate resources to herds that would have applied a measure for themselves. Yet, if decisions were interspersed, farmers may change their decision in anticipation of the social planner's resource allocation policy.

Regarding farmers, independently of the strategy chosen by the social planner, a term could be added in the cost function to decrease the probability of a farmer applying the measure if, at the last decision step, the planner allocated the resource to the herd. Alternatively, farmers' decision-making could be left unmodified, and it could be explored to what extent the social planner's decisions influence farmers' decisions only through the bias of variations in pathogen spread.

6.2.3.2 Decision-step optimisation: when to allocate?

Either for farmers' or social planner's decision-making, a fixed heuristic value was considered for the decision-step Δ_d , which for vaccination, also defines the duration of the vaccine efficacy. Even if this assumption is not unrealistic when compared to field practice for vaccination, other manners to determine the decision step are more appropriate depending on the control measure in mind (see section 6.2.1.4). These alternative ways of determining decision times rely on considering that decisions can be made asynchronously across herds. A first option would be to consider a decision step for each herd that is determined adaptively as a function of an external input. A second, more elaborate option, would be to build an optimisation procedure to determine when to take decisions. Such methodological variations of the model are not straightforward, especially when considering the second option (i.e. determine the decision-step by optimisation) together with the problem considered in this thesis (i.e. determine the resource allocation by optimisation).

6.3 Conclusion

The work presented in this thesis provides insights on modelling and optimisation of the control of pathogens spreading on animal trade networks, when such control is carried out voluntarily by farmers or by a central authority. Through identifying and exploiting the bridges between epidemiology and fields related to control decisions in a multi-agent setting, important questions and elements for such control at a large scale have been raised and addressed, either to model the voluntary behaviour of system agents, when they are humans, or to optimise the control of the system as a whole.

Research works on this topic are useful to better understand and predict real life phenomena that are not only dependent on biological processes, but also on other aspects such as human behaviour or governmental decision-making. Such better understanding and predictions have the potential of improving coordination among the different actors of the system. In particular, regarding the management of livestock diseases, a better coordination could ultimately account for the practices, interests and motivations of farmers, and at the same time allow for a better global control of the pathogen spread on complex networks through animal trade.

To this end, an integrative point of view seems to be particularly appropriate for such a cross-sectional problem. This should account not only for approaches from epidemiological field, but also from all the fields that study, in some way or another, decision-making. Indeed, in epidemiology, the control of pathogens could be improved by the expertise of the other fields regarding

decision-making, while such fields could be enriched by the complexity of the problems dealt with in epidemiology.

Chapter 7

Résumé en français: contexte et contributions de la thèse

Contents

7.1	Contexte	151
7.1.1	Gestion dynamique de la propagation d'une maladie sur un réseau complexe	151
7.1.2	Maladies endémiques du bétail se propageant par le biais du commerce des animaux	153
7.2	Contributions	155
7.2.1	Modèle intégratif pour la propagation d'agents pathogènes sur un réseau de commerce d'animaux rendant compte des décisions dynamiques des agriculteurs concernant l'adoption d'une mesure sanitaire	155
7.2.2	Stratégies pour l'allocation dynamique d'une ressource limitée pour la gestion de la propagation des maladies sur un réseau de métapopulation	156
7.2.3	Etude sur les décisions de vaccination des agriculteurs intégrées dans un modèle de propagation de la diarrhée virale bovine à grande échelle	157
7.3	Conclusion	157

7.1 Contexte

L'objectif principal de cette thèse était d'identifier, d'adapter et de construire les approches appropriées pour modéliser et optimiser efficacement l'adoption de mesures de gestion pour limiter la propagation d'une maladie sur un grand réseau de métapopulation. Le réseau a été spécifiquement considéré comme un réseau commercial de bétail à travers lequel la maladie peut se propager entre fermes dans une certaine région.

7.1.1 Gestion dynamique de la propagation d'une maladie sur un réseau complexe

La structure des contacts entre individus est un élément clé à prendre en compte pour mieux comprendre la propagation de maladies infectieuses et finalement pour la contrôler (Keeling et al. 2005). Ceci est particulièrement vrai pour la transmission de maladies entre plusieurs sous-populations sur une grande zone géographique. Bien que les agents pathogènes puissent être transmis entre les sous-populations de plusieurs façons, en fonction de l'agent pathogène et du contexte (par exemple, environnement partagé, vecteurs animaux sauvages), l'une des voies de transmission les plus courantes est le déplacement d'individus infectés (Danon et al. 2011).

En effet, les mouvements d'individus entre différentes sous-populations forment une structure appelée réseau de métapopulation (Keeling et al. 2005), dans laquelle les noeuds sont les sous-populations et les liens sont représentés par les mouvements. Un tel réseau est dirigé, car les individus se déplacent d'une sous-population à l'autre; pondéré, car il y a un poids (flux d'individus) associé à chaque lien; et dynamique, car les flux changent dans le temps (les liens ne sont pas pérennes et/ou le nombre d'individus qui se déplacent entre deux sous-populations peut varier dans le temps.). De plus, lorsque les coordonnées spatiales sont explicitement prises en compte, les liens peuvent connecter des sous-populations même s'il n'y a pas de proximité géographique entre elles (réseau spatial).

En particulier, ce contexte permet de décrire deux phénomènes spécifiques. D'abord, l'introduction d'un agent pathogène dans une sous-population où il n'était pas présent, appelé événement de colonisation (Donahue et al. 2008). Le second type de phénomène est la persistance de la propagation d'un pathogène à grande échelle malgré son extinction à petite échelle, ce que l'on appelle un *rescue effect* (effet de sauvetage) (Brown et al. 1977). Il s'agit d'un concept bien connu en écologie: des épidémies de courte durée ont lieu au niveau des sous-populations, mais la maladie se maintient à grande échelle en raison des interactions entre les sous-populations.

La question de la gestion d'une maladie se propageant sur un réseau aussi complexe se pose rapidement. Le système est intrinsèquement dynamique (en raison des mouvements des individus, des changements démographiques des sous-populations au fil du temps et de la propagation de la maladie), tout comme peut l'être la prise de décision. En effet, la gestion de maladies infectieuses dans la vie réelle est généralement effectuée par des décisions répétées dans le temps (campagnes de vaccination, restrictions temporelles des mouvements, etc.)

De plus, étant donné que le réseau est vaste et englobe de nombreuses sous-populations, les décisions concernant l'adoption de mesures de gestion pour réduire la propagation de la maladie peuvent être prises à deux niveaux différents. D'une part, s'il existe un agent décisionnel humain associé à chaque sous-population du système, nous sommes en présence d'une *prise de décision humaine décentralisée*. D'autre part, la gestion peut être assurée par un planificateur central, généralement un planificateur social. On parle alors de *prise de décision centralisée du planificateur social*. La distinction entre ces deux types de prise de décision est importante car, selon le décideur, les décisions répondront à des motivations différentes. En outre, la prise de décision humaine décentralisée peut naturellement introduire un comportement stratégique et une mise en œuvre hétérogène des mesures de gestion parmi les nombreuses sous-populations du système (Kreps 1997).

Dans ce contexte, l'utilisation d'approches mathématiques appropriées pour décrire et contrôler un tel système s'avère très utile. En particulier, les modèles épidémiologiques mécanistes peuvent être des instruments pour représenter et comprendre le système complexe impliqué dans la propagation des agents pathogènes (Keeling et al. 2011). Comme le comportement humain peut avoir un rôle majeur dans la propagation des maladies, en particulier pour celles dont la gestion est volontaire, la prise en compte des décisions liées à la gestion de la propagation peut augmenter la précision de ces modèles, et donc aider à mieux comprendre et prédire cette propagation. De plus, sur la base de ces modèles mécanistes, il est possible de concevoir des stratégies pour la gestion de la propagation de maladies infectieuses (Manfredi et al. 2013). Cependant, dans le processus de modélisation et de la gestion d'un tel système, des défis méthodologiques se posent lorsqu'il s'agit de prendre en compte les décisions dynamiques, qu'elles soient prises par un planificateur social ou par les agents humains du système.

Bien que la modélisation mathématique mécaniste soit de plus en plus utilisée dans l'étude des maladies infectieuses, la plupart des modèles existants dans la littérature ne font pas référence aux décisions volontaires d'agents en interaction (Wang et al. 2016), ou ils ne considèrent pas que le processus de prise de décision est dynamique (Rat-Aspert et al. 2010). En outre, les approches classiques considèrent les humains comme des particules (Manfredi et al. 2013), ou lorsque les aspects économiques sont pris en compte les agents sont considérés comme parfaitement rationnels (par exemple Bauch et al. 2004; Zhang et al. 2012), ce qui est une hypothèse forte et controversée pour la prise de décision humaine (Wang et al. 2016). Enfin, les influences entre les décisions des agents sont souvent négligées (Shi et al. 2019), malgré le fait qu'il s'agisse d'une caractéristique

importante du comportement humain, notamment en ce qui concerne la gestion des maladies du bétail (Hidano et al. 2018).

Par conséquent, du point de vue de la modélisation, un premier défi consistait à construire un cadre qui formalise de manière appropriée la relation entre la dynamique de la propagation d'une maladie sur un grand réseau de métapopulation et la dynamique de l'adoption volontaire de mesures de gestion dans chaque sous-population, tout en tenant compte de considérations psychologiques, cognitives ou économiques pertinentes.

Un deuxième défi consistait à développer ou adapter des méthodes qui puissent être effectivement appliquées pour gérer de manière optimale la propagation de la maladie sur le réseau. Parmi les considérations à prendre en compte, l'une des plus importantes concernait la complexité computationnelle du problème. En effet, il s'agissait du principal défi pour la résolution d'un problème d'optimisation dans le cadre actuel, en raison de la grande dimension du réseau (Pellis et al. 2015).

Pour cette raison, les recherches sur ce sujet se sont principalement concentrées sur deux types de situations, soit le réseau est petit (par exemple Chernov et al. 2020; Viet et al. 2018), soit il s'agit d'un grand réseau d'individus, c'est-à-dire qu'il ne s'agit pas d'une métapopulation où chaque noeud du système a une dynamique interne propre (par exemple Lorch et al. 2018; Zhang et al. 2015). Dans le premier cas, les études se fondent le plus souvent sur les processus de décision de Markov (MDP) (Puterman 2014) ou sur la théorie des jeux (Myerson 1997). Dans le second, les auteurs ont généralement recours à des approximations de type champ moyen inspirées de la physique (Lasry et al. 2007), qui peuvent être inappropriées si l'on souhaite prendre en compte une rationalité limitée ou des hétérogénéités entre les agents. Cela peut concerner par exemple l'influence qu'un agent peut avoir sur les décisions des autres agents, qui peut être déterminée par sa place dans le réseau.

Les avancées dans ces deux objectifs, modélisation et optimisation, contribuent à mieux comprendre les différentes caractéristiques du système complexe impliqué dans la propagation des agents pathogènes sur un tel réseau, et finalement à le contrôler efficacement.

7.1.2 Maladies endémiques du bétail se propageant par le biais du commerce des animaux

Les modalités de transmission des maladies entre individus sont proches que l'on considère des humains ou des animaux (faune sauvage ou bétail), mais quelques différences existent néanmoins. Tout d'abord, les humains ont tendance à se déplacer librement entre les foyers, les lieux de travail, les lieux d'étude, les villes, les pays, etc., tandis que le bétail reste généralement longtemps dans une exploitation avant d'en rejoindre une autre par le biais du commerce d'animaux (Brooks-Pollock et al. 2015). Deuxièmement, les considérations économiques peuvent être plus centrales pour la gestion des maladies du bétail que pour les maladies humaines, où l'aspect santé publique prime généralement sur le reste. Ainsi, les avancées méthodologiques pour la gestion des premières ont en principe plus de chances de trouver l'adhésion des agents de terrain.

Suite à la crise de l'encéphalopathie spongiforme bovine (communément appelée maladie de la vache folle) qui a eu lieu notamment au Royaume-Uni entre 1986 et 2000, les pays européens maintiennent des bases de données nationales concernant les mouvements de bovins entre exploitations. En effet, les mouvements d'animaux occasionnés par les échanges constituent une voie majeure de transmission des agents pathogènes du bétail entre les exploitations (Fèvre et al. 2006), par exemple la paratuberculose (Beunée et al. 2015), la fièvre aphteuse (FMD) (Ferguson et al. 2001) et la tuberculose bovine (bTB) (Donnelly et al. 2003).

Ces échanges se font directement entre deux exploitations, ou peuvent passer par des structures intermédiaires. Les échanges directs concernent généralement des exploitations géographiquement proches, de sorte que la propagation de la maladie est surtout concentrée dans une petite zone géographique. Par ailleurs, les mouvements d'animaux peuvent se faire via des marchés

ou des centres de rassemblement. Ces structures facilitent le commerce d'animaux à une grande échelle géographique, en augmentant les échanges d'animaux provenant d'une exploitation géographiquement éloignée de l'exploitation de destination. Elles augmentent donc encore le risque de transmission d'agents pathogènes sur de vastes zones (Robinson et al. 2007). La figure 7.1 illustre l'exemple d'un réseau commercial de bovins formé par les mouvements d'animaux laitiers entre les troupeaux de bovins, qui s'est produit en 2009 en France selon la base de données nationale d'identification (BDNI), qui enregistre l'histoire de vie de chaque bovin de la naissance à la mort. Cette figure met en évidence les échanges d'animaux géographiquement proches et éloignés qui peuvent avoir lieu dans un réseau commercial d'animaux, et par lesquels une propagation de la maladie est susceptible d'atteindre l'ensemble du territoire français.



Figure 7.1: Flux commerciaux de bovins (animaux laitiers uniquement) pour l'année 2009 en France, à partir de la base de données française d'identification des bovins. Chaque nœud est une commune (la plus petite unité administrative française). Source: Gaël Beaunée.

Une maladie qui se propage par le biais des échanges entre animaux a de fortes chances de devenir *endémique* dans une zone donnée, c'est-à-dire présente à une prévalence donnée (généralement modérée) dans l'ensemble de la population pendant de longues périodes (Carslake et al. 2011). Comme mentionné pour le cadre général des métapopulations, dans les métapopulations animales, la persistance de la maladie peut être observée à grande échelle par le biais de deux processus. Premièrement, une longue durée d'infection dans une sous-population donne lieu à une longue infection de la métapopulation. Ensuite, la persistance peut être observée s'il existe un *rescue effect* dû aux interactions entre les sous-populations (voir par exemple Jesse et al. 2008).

Par ailleurs, une maladie peut présenter une dynamique *épidémique*, ce qui signifie par définition que l'on observe une propagation importante et rapide sur une courte période de temps. Des exemples de ce type sont la fièvre aphteuse (Ferguson et al. 2001), la peste porcine africaine (PPA) (Nigsch et al. 2013), et la grippe aviaire (Benincà et al. 2020). Pour une maladie donnée, la distinction entre dynamique endémique et dynamique épidémique se fait à partir des données de surveillance. Or, les données de surveillance détaillées sont principalement disponibles pour des maladies qui ont déjà été qualifiées d'épidémiques (Carslake et al. 2011).

Une fois qu'une maladie est associée à une dynamique épidémique, des efforts politiques importants sont déployés afin de l'éradiquer. En particulier, les politiques publiques ciblent généralement les maladies épidémiques du bétail en réglementant leur gestion de manière obligatoire. Dans ce cas, on dit que la maladie est *réglementée*. En revanche, les maladies endémiques intéressent généralement moins l'opinion publique et les décideurs politiques. Par conséquent, leur gestion est souvent laissée aux initiatives individuelles ou locales, et n'est donc pas obligatoire, ce qui fait que les maladies endémiques sont le plus souvent *non réglementées*.

Néanmoins, comme elles persistent sur de longues périodes, les maladies endémiques peuvent avoir une incidence cumulée importante, entraînant une réduction de la rentabilité économique des exploitations et du bien-être des animaux (Tomley et al. 2009). En outre, les zoonoses (les maladies infectieuses qui peuvent se propager entre animaux et humains) jouent un rôle majeur dans le nombre croissant de maladies humaines émergentes (Lefrançois et al. 2014). En effet, au moins 60% des maladies infectieuses humaines émergentes sont des zoonoses et, en particulier, plus de 30% des zoonoses infectieuses émergentes sont associées aux animaux destinés à l'alimentation (Otte et al. 2021). Par conséquent, la gestion de maladies endémiques du bétail représente un défi majeur pour la santé animale et pour des systèmes agroalimentaires durables, en particulier dans un contexte de sous-populations qui échangent des animaux, et de maladies endémiques pour lesquelles la gestion n'est pas obligatoire.

La prise de décision pour l'adoption d'une ou plusieurs mesures sanitaires est donc intéressante pour évaluer la gestion des maladies endémiques animales. Les objectifs de production des éleveurs n'étant pas seulement motivés par la santé animale mais aussi par des critères de temps de travail, de productivité, de rentabilité, etc., l'hétérogénéité de la mise en œuvre des mesures est accrue (Ezanno et al. 2020). Par exemple, un décideur central peut décider d'isoler un troupeau si cela implique une réduction globale de la maladie, à condition que ce soit l'objectif visé. Cependant, si les éleveurs décident eux-mêmes d'isoler ou non leur troupeau, ils prendront cette décision uniquement sur la base de leurs propres critères, ce qui n'est généralement pas optimal au niveau collectif (Krebs et al. 2018).

Pour les maladies non réglementées, il est naturel de penser que les décisions concernant la gestion du troupeau sont principalement prises par les agriculteurs eux-mêmes. Pourtant, les décisions relatives aux mesures de lutte peuvent également être prises à un niveau centralisé, ne serait-ce que par des organisations locales composées de plusieurs agriculteurs. Dans certaines régions administratives françaises, par exemple, les GDS (groupements de défense sanitaire) peuvent s'engager dans divers types d'actions de gestion pour réduire la prévalence locale d'une maladie. Ces actions peuvent impliquer, par exemple, l'allocation de ressources sanitaires pour intervenir dans le système sans exiger de tous les agriculteurs qu'ils appliquent la mesure, ou le recours à des campagnes pour encourager l'adoption volontaire de la mesure par les agriculteurs. L'intervention peut également se faire en facilitant l'accès aux informations sur le statut infectieux des troupeaux, afin que les éleveurs puissent prendre des décisions plus éclairées.

En outre, de nombreuses mesures de gestion ou de prévention sont possibles (application de vaccins ou de traitements, tests et abattage des animaux positifs, isolement des animaux infectés, etc.) Il peut être difficile de choisir, tant du point de vue de l'éleveur que du planificateur social, entre les différentes options dont l'efficacité n'est pas toujours connue ou comparable. De plus, anticiper l'impact de ces choix individuels ou collectifs sur la dynamique à grande échelle des maladies infectieuses reste un défi.

7.2 Contributions

Dans le contexte décrit ci-dessus, cette thèse a proposé deux contributions méthodologiques principales et une contribution concernant une étude d'application. Dans ce qui suit, chacune de ces contributions est brièvement résumée.

7.2.1 Modèle intégratif pour la propagation d'agents pathogènes sur un réseau de commerce d'animaux rendant compte des décisions dynamiques des agriculteurs concernant l'adoption d'une mesure sanitaire

La première contribution (chapitre 3) aborde le défi de la modélisation, à travers le développement d'un nouveau modèle intégratif de propagation d'agents pathogènes sur un réseau de métapopulation de commerce animal, qui tient compte des décisions volontaires des agriculteurs concernant l'adoption d'une mesure de gestion sanitaire contre l'infection dans leur propre trou-

peau. La composante épidémiologique est construite sur la base d'un modèle stochastique SIR intra-troupeau avec démographie, qui tient compte des transferts d'animaux entre troupeaux. La composante décisionnelle consiste à poser un problème de décision dynamique sous incertitude et à proposer un mécanisme de décision dynamique qui tient compte du comportement stochastique des éleveurs, de l'apprentissage et de la dynamique d'imitation stratégique. Ainsi, il existe une boucle de rétroaction dynamique entre les deux composantes, car les décisions de gestion sont fonction de la propagation observée de la maladie, et la propagation de la maladie est à son tour impactée par les décisions prises précédemment. Le modèle a été détaillé et simulé pour une mesure spécifique (la vaccination). Les dynamiques d'infection et des décisions de vaccination pour des scénarios contrastés ont été étudiées, et la robustesse concernant les valeurs des paramètres du modèle a été évaluée. Les expériences ont montré que les paramètres liés aux décisions ont un impact important sur les décisions de vaccination, et donc sur l'évolution de la propagation de la maladie.

Des études précédentes (Bauch 2005; Fu et al. 2010) ont été pionnières dans la prise en compte des aspects réels du comportement humain concernant la vaccination contre les maladies humaines, tels que le comportement stochastique et l'imitation. Pourtant, ce type d'étude a rarement pris en compte un processus d'apprentissage explicite pour représenter la prise de décision humaine. De plus, ces travaux n'avaient pas encore été adaptés aux spécificités liées à la gestion des maladies du bétail. Ainsi, le premier modèle proposé dans cette thèse est un pas important vers la prise en compte du comportement humain dans les modèles épidémiologiques. De plus, sa formalisation générique lui permet d'être pertinent pour une variété de modèles ou de mesures épidémiologiques. Les extensions et utilisations potentielles peuvent être facilitées par la disponibilité publique du code de simulation.

Les travaux de ce chapitre ont été publiés dans Scientific Reports (Cristancho Fajardo et al. 2021).

7.2.2 Stratégies pour l'allocation dynamique d'une ressource limitée pour la gestion de la propagation des maladies sur un réseau de métapopulation

La deuxième contribution (chapitre 4) concernait l'optimisation des décisions d'un planificateur social concernant l'allocation d'une ressource limitée entre les sous-populations d'un réseau de métapopulation. Ce problème a été étudié pour une maladie du bétail se propageant sur un réseau de commerce d'animaux. Là encore, pour chaque sous-population (troupeau) du réseau, un modèle stochastique SIR intra-troupeau avec démographie a été utilisé, tenant compte des transferts d'animaux entre troupeaux et s'appuyant sur des données pour ses composantes non épidémiologiques. La contribution est basée sur l'extension à ce cadre de métapopulation d'une approche initialement proposée pour les réseaux sans dynamique épidémiologique-démographique intra-nœud (i.e. un nœud du réseau correspondait à un individu et non à une sous-population). L'approche que nous avons proposée consiste à dériver des scores analytiques minimisant une approximation d'une fonction objectif, puis à allouer dynamiquement la ressource à chaque sous-population en fonction de leur classement donné par les scores des sous-populations, et de la quantité limitée de la ressource. De telles fonctions de score optimisées ont été obtenues pour deux fonctions objectives différentes (le nombre d'animaux infectés dans la métapopulation et le nombre de troupeaux infectés) et pour deux types de ressources différents (vaccin et traitement). Leurs performances pour contenir la propagation de la maladie et leur pertinence en pratique pour l'allocation des ressources ont été évaluées, ainsi que celles de plusieurs scores heuristiques. Il a été démontré que les fonctions de scores que nous avons obtenues analytiquement peuvent permettre une meilleure gestion de la propagation de l'agent pathogène, bien qu'elles ne soient pas toujours vraiment optimales en pratique. Cependant, elles sont en général beaucoup plus performantes que les fonctions de score basées uniquement sur des informations démographiques ou topologiques, car dans la plupart des scénarios évalués, les informations sur le statut épidémiologique sont apparues comme nécessaires pour parvenir à éradiquer la maladie.

Ce travail représente une avancée dans l'optimisation des interventions socialement planifiées dans un système complexe de propagation de maladies sur un grand réseau de métapopulation. En effet, les études similaires se réfèrent soit à de petits réseaux de métapopulation (par exemple, Mbah et al. 2011), soit à de grands réseaux d'individus (par exemple, Wijayanto et al. 2019). La même approche pourrait également être utilisée pour différents modèles épidémiologiques compartimentaux dans le cadre des métapopulations. De plus, d'autres mesures peuvent être étudiées conduisant à d'autres scores analytiques spécifiques, à obtenir. Enfin, comme pour le travail du chapitre 3, le code de simulation a été mis à disposition dans un dépôt public.

Ce chapitre a été publié dans le Journal of the Royal Society Interface (Cristancho-Fajardo et al. 2022).

7.2.3 Etude sur les décisions de vaccination des agriculteurs intégrées dans un modèle de propagation de la diarrhée virale bovine à grande échelle

La dernière contribution (chapitre 5) consiste en une application et une extension du modèle intégratif du chapitre 3, à une maladie spécifique de la vie réelle, la diarrhée virale bovine (BVD), en supposant que la prise de décision des agriculteurs peut être influencée de manière différente de celle initialement considérée. Dans cette application, le cadre méthodologique a été étendu pour considérer que les agriculteurs peuvent prendre des décisions en fonction des décisions et des observations de leurs voisins, non seulement dans le réseau de commerce des animaux, comme le suppose le modèle générique original, mais aussi dans leur voisinage géographique (en fonction de la proximité géographique avec d'autres exploitations). En outre, plusieurs méthodes alternatives sur la façon dont les voisins sont choisis ont été évaluées. Enfin, ce travail en cours met en évidence la façon dont les composantes du modèle intégratif peuvent être adaptées à l'étude d'un type différent de modèle épidémiologique (un modèle basé sur l'individu) pour la gestion d'une maladie réelle, dans ce cas la BVD, par l'adoption d'une mesure sanitaire (la vaccination) qui s'est avérée être intéressante pour certains contextes (Arnoux et al. 2021).

Le travail de chapitre a été consolidé et soumis pour publication dans Veterinary Research.

7.3 Conclusion

Le travail réalisé dans cette thèse fournit des informations sur la modélisation et l'optimisation de la gestion des agents pathogènes se propageant sur les réseaux de commerce animal, lorsque cette gestion est effectuée volontairement par les éleveurs ou par une autorité centrale. En identifiant et en exploitant les ponts entre l'épidémiologie et les domaines liés aux décisions de gestion dans un cadre multi-agents, des questions et des éléments importants pour une telle gestion à grande échelle ont été soulevés et traités, soit pour modéliser le comportement volontaire des agents du système, lorsqu'il s'agit d'humains, soit pour optimiser la gestion du système dans son ensemble.

Les travaux de recherche sur ce sujet sont utiles pour mieux comprendre et prédire les phénomènes de la vie réelle qui ne dépendent pas seulement des processus biologiques, mais aussi d'autres aspects tels que le comportement humain ou la prise de décision par les pouvoirs publics. Cette meilleure compréhension et ces prévisions ont le potentiel d'améliorer la coordination entre les différents acteurs du système. En particulier, en ce qui concerne la gestion des maladies du bétail, une meilleure coordination pourrait tenir compte des pratiques, des intérêts et des motivations des agriculteurs, tout en permettant une meilleure gestion globale de l'agent pathogène qui se propage sur des réseaux complexes par le biais du commerce des animaux.

À cette fin, un point de vue intégratif semble particulièrement approprié pour un problème aussi transversal. Il doit prendre en compte non seulement les approches issues du domaine de l'épidémiologie, mais aussi de tous les domaines qui étudient, d'une manière ou d'une autre, la prise de décision. En effet, en épidémiologie, la gestion des agents pathogènes pourrait être améliorée par l'expertise des autres domaines concernant la prise de décision, tandis que ces

domaines pourraient s'enrichir de la complexité des problèmes traités en épidémiologie.

Bibliography

- Alamo, T., Reina, D. G., Gata, P. M., Preciado, V. M., and Giordano, G. (2021). “Data-Driven Methods for Present and Future Pandemics: Monitoring, Modelling and Managing”. en. In: *Annual Reviews in Control*. arXiv: 2102.13130, S1367578821000419. ISSN: 13675788. DOI: [10.1016/j.arcontrol.2021.05.003](https://doi.org/10.1016/j.arcontrol.2021.05.003).
- Albert, R. and Barabási, A.-L. (2002). “Statistical mechanics of complex networks”. en. In: *Reviews of Modern Physics* 74.1, pp. 47–97. ISSN: 0034-6861, 1539-0756. DOI: [10.1103/RevModPhys.74.47](https://doi.org/10.1103/RevModPhys.74.47).
- Allen, L. J. S. (2008). “An Introduction to Stochastic Epidemic Models”. en. In: *Mathematical Epidemiology*. Ed. by Brauer, F., Driessche, P. van den, and Wu, J. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, pp. 81–130. ISBN: 978-3-540-78911-6. DOI: [10.1007/978-3-540-78911-6_3](https://doi.org/10.1007/978-3-540-78911-6_3).
- Alshamsi, A., Pianesi, F., Lepri, B., Pentland, A., and Rahwan, I. (2015). “Beyond Contagion: Reality Mining Reveals Complex Patterns of Social Influence”. en. In: *PLOS ONE* 10.8. Ed. by Bauch, C. T., e0135740. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0135740](https://doi.org/10.1371/journal.pone.0135740).
- Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., and Rohani, P. (2006). “Seasonality and the dynamics of infectious diseases”. en. In: *Ecology Letters* 9.4, pp. 467–484. ISSN: 1461-0248. DOI: [10.1111/j.1461-0248.2005.00879.x](https://doi.org/10.1111/j.1461-0248.2005.00879.x).
- Anderson, D. F. (2008). “Incorporating postleap checks in tau-leaping”. en. In: *The Journal of Chemical Physics* 128.5, p. 054103. ISSN: 0021-9606, 1089-7690. DOI: [10.1063/1.2819665](https://doi.org/10.1063/1.2819665).
- Anderson, D. and Watson, R. (1980). “On the spread of a disease with gamma distributed latent and infectious periods”. In: *Biometrika* 67.1, pp. 191–198. ISSN: 0006-3444. DOI: [10.1093/biomet/67.1.191](https://doi.org/10.1093/biomet/67.1.191).
- Andersson, H. and Britton, T. (2000). “Stochastic epidemics in dynamic populations: quasi-stationarity and extinction”. en. In: *Journal of Mathematical Biology* 41.6, pp. 559–580. ISSN: 0303-6812, 1432-1416. DOI: [10.1007/s002850000060](https://doi.org/10.1007/s002850000060).
- Andersson, H. and Britton, T. (2012). *Stochastic Epidemic Models and Their Statistical Analysis*. en. Springer Science & Business Media. ISBN: 978-1-4612-1158-7.
- Andriotis, C. and Papakonstantinou, K. (2019). “Managing engineering systems with large state and action spaces through deep reinforcement learning”. In: *Reliability Engineering & System Safety* 191, p. 106483.
- Arnoux, S., Bidan, F., Damman, A., Petit, E., Assié, S., and Ezanno, P. (2021). “To Vaccinate or Not: Impact of Bovine Viral Diarrhoea in French Cow-Calf Herds”. en. In: *Vaccines* 9.10. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute, p. 1137. DOI: [10.3390/vaccines9101137](https://doi.org/10.3390/vaccines9101137).
- Auer, P., Cesa-Bianchi, N., and Fischer, P. (2002a). “Finite-time analysis of the multiarmed bandit problem”. In: *Machine Learning* 47.2-3, pp. 235–256.
- Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. (2002b). “The Nonstochastic Multiarmed Bandit Problem”. en. In: *SIAM Journal on Computing* 32.1, pp. 48–77. ISSN: 0097-5397, 1095-7111. DOI: [10.1137/S0097539701398375](https://doi.org/10.1137/S0097539701398375).
- Bajardi, P., Barrat, A., Natale, F., Savini, L., and Colizza, V. (2011). “Dynamical patterns of cattle trade movements”. In: *PloS one* 6.5, e19869.
- Bajardi, P., Barrat, A., Savini, L., and Colizza, V. (2012). “Optimizing surveillance for livestock disease spreading through animal movements”. In: *Journal of the Royal Society Interface* 9.76, pp. 2814–2825.

- Bandura, A. and Walters, R. H. (1977). *Social learning theory*. Vol. 1. Englewood cliffs Prentice Hall.
- Barabási, A.-L. and Albert, R. (1999). "Emergence of Scaling in Random Networks". In: *Science* 286.5439. Publisher: American Association for the Advancement of Science, pp. 509–512. DOI: [10.1126/science.286.5439.509](https://doi.org/10.1126/science.286.5439.509).
- Bartlett, M. S. (1956). "Deterministic and Stochastic Models for Recurrent Epidemics". In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Health* 3.4. Publisher: University of California Press, pp. 81–110.
- Bartlett, M. S. (1957). "Measles Periodicity and Community Size". In: *Journal of the Royal Statistical Society. Series A (General)* 120.1. Publisher: [Royal Statistical Society, Wiley], pp. 48–70. ISSN: 0035-9238. DOI: [10.2307/2342553](https://doi.org/10.2307/2342553).
- Bauch, C. T. (2005). "Imitation dynamics predict vaccinating behaviour". In: *Proceedings of the Royal Society B: Biological Sciences* 272.1573. Publisher: Royal Society, pp. 1669–1675. DOI: [10.1098/rspb.2005.3153](https://doi.org/10.1098/rspb.2005.3153).
- Bauch, C. T. and Earn, D. J. D. (2004). "Vaccination and the theory of games". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 101.36, pp. 13391–13394. ISSN: 0027-8424. DOI: [10.1073/pnas.0403823101](https://doi.org/10.1073/pnas.0403823101).
- Bauch, C. T., Galvani, A. P., and Earn, D. J. D. (2003). "Group interest versus self-interest in smallpox vaccination policy". en. In: *Proceedings of the National Academy of Sciences* 100.18. Publisher: National Academy of Sciences Section: Biological Sciences, pp. 10564–10567. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.1731324100](https://doi.org/10.1073/pnas.1731324100).
- Beaunée, G., Vergu, E., and Ezanno, P. (2015). "Modelling of paratuberculosis spread between dairy cattle farms at a regional scale". In: *Veterinary Research* 46.1, p. 111.
- Beaunée, G., Vergu, E., Joly, A., and Ezanno, P. (2017). "Controlling bovine paratuberculosis at a regional scale: towards a decision modelling tool". In: *Journal of theoretical biology* 435, pp. 157–183.
- Behncke, H. (2000). "Optimal control of deterministic epidemics". en. In: *Optimal Control Applications and Methods* 21.6, pp. 269–285. ISSN: 1099-1514. DOI: [10.1002/oca.678](https://doi.org/10.1002/oca.678).
- Benincà, E., Hagenaars, T., Boender, G. J., Kassteele, J. v. d., and Boven, M. v. (2020). "Trade-off between local transmission and long-range dispersal drives infectious disease outbreak size in spatially structured populations". en. In: *PLOS Computational Biology* 16.7. Publisher: Public Library of Science, e1008009. ISSN: 1553-7358. DOI: [10.1371/journal.pcbi.1008009](https://doi.org/10.1371/journal.pcbi.1008009).
- Berry, D. A. and Fristedt, B. (1985). *Bandit problems*. en. Dordrecht: Springer Netherlands. DOI: [10.1007/978-94-015-3711-7](https://doi.org/10.1007/978-94-015-3711-7).
- Bidot, C., Lamboni, M., and Monod, H. (2018). *multisensi: Multivariate Sensitivity Analysis*. R package version 2.1-1. URL: <https://CRAN.R-project.org/package=multisensi>.
- Bigras-Poulin, M., Thompson, R., Chriél, M., Mortensen, S., and Greiner, M. (2006). "Network analysis of Danish cattle industry trade patterns as an evaluation of risk potential for disease spread". In: *Preventive veterinary medicine* 76.1-2, pp. 11–39.
- Bloembergen, D., Tuyls, K., Hennes, D., and Kaisers, M. (2015). "Evolutionary Dynamics of Multi-Agent Learning: A Survey". en. In: *Journal of Artificial Intelligence Research* 53, pp. 659–697. ISSN: 1076-9757. DOI: [10.1613/jair.4818](https://doi.org/10.1613/jair.4818).
- Börgers, T. and Sarin, R. (1997). "Learning Through Reinforcement and Replicator Dynamics". en. In: *Journal of Economic Theory* 77.1, pp. 1–14. ISSN: 0022-0531. DOI: [10.1006/jeth.1997.2319](https://doi.org/10.1006/jeth.1997.2319).
- Boyd, S., Kim, S.-J., Vandenberghe, L., and Hassibi, A. (2007). "A tutorial on geometric programming". en. In: *Optimization and Engineering* 8.1, p. 67. ISSN: 1573-2924. DOI: [10.1007/s11081-007-9001-7](https://doi.org/10.1007/s11081-007-9001-7).
- Brandeau, M., Zaric, G., and Richter, A. (2003). "Resource allocation for control of infectious diseases in multiple independent populations: Beyond cost-effectiveness analysis". In: *Journal of Health Economics* 22.4, pp. 575–598. DOI: [10.1016/S0167-6296\(03\)00043-2](https://doi.org/10.1016/S0167-6296(03)00043-2).
- Braunstein, A., Dall'Asta, L., Semerjian, G., and Zdeborová, L. (2016). "Network dismantling". In: *Proceedings of the National Academy of Sciences* 113.44, pp. 12368–12373.
- Bretó, C., He, D., Ionides, E. L., and King, A. A. (2009). "Time series analysis via mechanistic models". In: *The Annals of Applied Statistics* 3.1, pp. 319–348.

- Britton, T. (2010). "Stochastic epidemic models: A survey". en. In: *Mathematical Biosciences* 225.1, pp. 24–35. ISSN: 0025-5564. DOI: [10.1016/j.mbs.2010.01.006](https://doi.org/10.1016/j.mbs.2010.01.006).
- Britton, T. and Pardoux, E. (2019). *Stochastic Epidemic Models with Inference*. eng. Springer.
- Brooks-Pollock, E., De Jong, M., Keeling, M., Klinkenberg, D., and Wood, J. (2015). "Eight challenges in modelling infectious livestock diseases". en. In: *Epidemics. Challenges in Modelling Infectious Disease Dynamics* 10, pp. 1–5. ISSN: 1755-4365. DOI: [10.1016/j.epidem.2014.08.005](https://doi.org/10.1016/j.epidem.2014.08.005).
- Brown, J. H. and Kodric-Brown, A. (1977). "Turnover rates in insular biogeography: effect of immigration on extinction". In: *Ecology* 58.2, pp. 445–449.
- Busemeyer, J. (2001). "Dynamic Decision Making". In: *International Encyclopedia of the Social Behavioral Sciences*. Ed. by Smelser, N. J. and Baltes, P. B. Oxford: Pergamon, pp. 3903–3908. ISBN: 978-0-08-043076-8. DOI: <https://doi.org/10.1016/B0-08-043076-7/00641-0>.
- Bussell, E. H., Dangerfield, C. E., Gilligan, C. A., and Cunniffe, N. J. (2019). "Applying optimal control theory to complex epidemiological models to inform real-world disease management". In: *Philosophical Transactions of the Royal Society B* 374.1776, p. 20180284.
- Capasso, V. and Serio, G. (1978). "A generalization of the Kermack-McKendrick deterministic epidemic model". en. In: *Mathematical Biosciences* 42.1, pp. 43–61. ISSN: 0025-5564. DOI: [10.1016/0025-5564\(78\)90006-8](https://doi.org/10.1016/0025-5564(78)90006-8).
- Cardillo, A., Reyes-Suárez, C., Naranjo, F., and Gómez-Gardenes, J. (2013). "Evolutionary vaccination dilemma in complex networks". In: *Physical Review E* 88.3, p. 032803.
- Carli, R., Cavone, G., Epicoco, N., Scarabaggio, P., and Dotoli, M. (2020). "Model predictive control to mitigate the COVID-19 outbreak in a multi-region scenario". In: *Annual Reviews in Control*.
- Carlsson, U., Fredriksson, G., Alenius, S., and Kindahl, H. (1989). "Bovine Virus Diarrhoea Virus, a Cause of Early Pregnancy Failure in the Cow". en. In: *Journal of Veterinary Medicine Series A* 36.1-10, pp. 15–23. ISSN: 1439-0442. DOI: [10.1111/j.1439-0442.1989.tb00697.x](https://doi.org/10.1111/j.1439-0442.1989.tb00697.x).
- Carlslake, D., Grant, W., Green, L. E., Cave, J., Greaves, J., Keeling, M. J., McEldowney, J., Weldegebriel, H., and Medley, G. F. (2011). "Endemic cattle diseases: comparative epidemiology and governance". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 366.1573, pp. 1975–1986.
- Chambres d'agriculture de Bourgogne Franche-Comté (2022). *Poids et prix de vente des animaux Charolais en 2019*. fr. URL: <https://bourgognefranche-comte.chambres-agriculture.fr/publications/la-publication-en-detail/actualites/poids-et-prix-de-vente-des-animaux-charolais-en-2019-1/> (visited on 01/08/2022).
- Chang, S. L., Piraveenan, M., Pattison, P., and Prokopenko, M. (2019). "Game theoretic modelling of infectious disease dynamics and intervention methods: a mini-review". en. In: *arXiv:1901.04143 [physics, q-bio]*.
- Chernov, A. A., Kelbert, M. Y., and Shemendyuk, A. A. (2020). "Optimal vaccine allocation during the mumps outbreak in two SIR centres". In: *Mathematical medicine and biology: a journal of the IMA* 37.3, pp. 303–312.
- Choisy, M., Guégan, J.-F., and Rohani, P. (2007). "Mathematical modeling of infectious diseases dynamics". In: *Encyclopedia of infectious diseases: modern methodologies* 379.
- Chow, Y. S., Robbins 1915-2001 (viaf)81628103, H. E., and Siegmund, D. (1971). *Great expectations : the theory of optimal stopping*. eng. Boston (Mass.) : Houghton Mifflin. ISBN: 0-395-05314-5.
- Cohen, M., Jaffray, J. Y., and Said, T. (1985). "Individual behavior under risk and under uncertainty: An experimental study". In: *Theory and Decision* 18.2. Place: Germany Publisher: Springer, pp. 203–228. ISSN: 1573-7187. DOI: [10.1007/BF00134074](https://doi.org/10.1007/BF00134074).
- Cohen, R., Havlin, S., and ben-Avraham, D. (2003). "Efficient Immunization Strategies for Computer Networks and Populations". en. In: *Physical Review Letters* 91.24, p. 247901. ISSN: 0031-9007, 1079-7114. DOI: [10.1103/PhysRevLett.91.247901](https://doi.org/10.1103/PhysRevLett.91.247901).
- Colizza, V., Barrat, A., Barthélemy, M., and Vespignani, A. (2006). "The role of the airline transportation network in the prediction and predictability of global epidemics". en. In: *Proceedings of the National Academy of Sciences* 103.7, pp. 2015–2020. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.0510525103](https://doi.org/10.1073/pnas.0510525103).

- Colizza, V. and Vespignani, A. (2008a). "Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations". In: *Journal of theoretical biology* 251.3, pp. 450–467.
- Colizza, V. and Vespignani, A. (2008b). "Invasion threshold in heterogeneous metapopulation networks". en. In: *arXiv:0802.3636 [cond-mat, physics:physics, q-bio]*. arXiv: 0802.3636. DOI: [10.1103/PhysRevLett.99.148701](https://doi.org/10.1103/PhysRevLett.99.148701).
- Cousins, D. (2001). "Mycobacterium bovis infection and control in domestic livestock". In: *Revue Scientifique et Technique-Office International des Epizooties* 20.1, pp. 71–85.
- Cristancho Fajardo, L., Ezanno, P., and Vergu, E. (2021). "Accounting for farmers' control decisions in a model of pathogen spread through animal trade". In: *Scientific Reports* 11. DOI: [10.1038/s41598-021-88471-6](https://doi.org/10.1038/s41598-021-88471-6).
- Cristancho-Fajardo, L., Ezanno, P., and Vergu, E. (2022). "Dynamic resource allocation for controlling pathogen spread on a large metapopulation network". In: *Journal of the Royal Society Interface* 19.188, p. 20210744.
- Dangerfield, C. E., Vyska, M., and Gilligan, C. A. (2019). "Resource allocation for epidemic control across multiple sub-populations". In: *Bulletin of mathematical biology* 81.6, pp. 1731–1759.
- Danon, L., Ford, A. P., House, T., Jewell, C. P., Keeling, M. J., Roberts, G. O., Ross, J. V., and Vernon, M. C. (2011). "Networks and the epidemiology of infectious disease". In: *Interdisciplinary perspectives on infectious diseases* 2011.
- De Jong, M., Diekmann, O., and Heesterbeek, J. (1994). "How does transmission of infection depend on population size?" In: *Epidemic models: their structure and relation to data*. -Cambridge : Cambridge University Press, 1995. pp. 84-94.
- DeAngelis, D. L. (2018). *Individual-based models and approaches in ecology: populations, communities and ecosystems*. CRC Press.
- Delmas, J.-F., Dronnier, D., and Zitt, P.-A. (2021). "Targeted Vaccination Strategies for an Infinite-dimensional SIS Model". In: *arXiv preprint arXiv:2103.10330*.
- Diederich, A. (2001). "Sequential Decision Making". en. In: *International Encyclopedia of the Social & Behavioral Sciences*. Ed. by Smelser, N. J. and Baltes, P. B. Oxford: Pergamon, pp. 13917–13922. ISBN: 978-0-08-043076-8. DOI: [10.1016/B0-08-043076-7/00636-7](https://doi.org/10.1016/B0-08-043076-7/00636-7).
- Diekmann, O., Heesterbeek, J. a. P., and Roberts, M. G. (2010). "The construction of next-generation matrices for compartmental epidemic models". In: *Journal of The Royal Society Interface* 7.47. Publisher: Royal Society, pp. 873–885. DOI: [10.1098/rsif.2009.0386](https://doi.org/10.1098/rsif.2009.0386).
- Dijkstra, E. W. (1959). "A note on two problems in connexion with graphs". In: *Numerische mathematik* 1.1, pp. 269–271.
- Ding, H., Xu, J.-H., Wang, Z., Ren, Y.-Z., and Cui, G.-H. (2018). "Subsidy strategy based on history information can stimulate voluntary vaccination behaviors on seasonal diseases". en. In: *Physica A: Statistical Mechanics and its Applications* 503, pp. 390–399. ISSN: 0378-4371. DOI: [10.1016/j.physa.2018.03.003](https://doi.org/10.1016/j.physa.2018.03.003).
- Donahue, M. J. and Lee, C. T. (2008). "Colonization". en. In: *Encyclopedia of Ecology*. Ed. by Jørgensen, S. E. and Fath, B. D. Oxford: Academic Press, pp. 672–678. ISBN: 978-0-08-045405-4. DOI: [10.1016/B978-008045405-4.00472-9](https://doi.org/10.1016/B978-008045405-4.00472-9).
- Donnelly, C. A., Woodroffe, R., Cox, D. R., Bourne, J., Gettinby, G., Le Fevre, A. M., McNerney, J. P., and Morrison, W. I. (2003). "Impact of localized badger culling on tuberculosis incidence in British cattle". en. In: *Nature* 426.6968. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 6968 Primary_atype: Research Publisher: Nature Publishing Group, pp. 834–837. ISSN: 1476-4687. DOI: [10.1038/nature02192](https://doi.org/10.1038/nature02192).
- Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. (2001). "Giant strongly connected component of directed networks". In: *Physical Review E* 64.2, p. 025101.
- Drakopoulos, K., Ozdaglar, A., and Tsitsiklis, J. N. (2014). "An Efficient Curing Policy for Epidemics on Graphs". en. In: *IEEE Transactions on Network Science and Engineering* 1.2, pp. 67–75. ISSN: 2327-4697. DOI: [10.1109/TNSE.2015.2393291](https://doi.org/10.1109/TNSE.2015.2393291).
- Dubé, C., Ribble, C., Kelton, D., and McNab, B. (2008). "Comparing network analysis measures to determine potential epidemic size of highly contagious exotic diseases in fragmented monthly networks of dairy cattle movements in Ontario, Canada". In: *Transboundary and emerging diseases* 55.9-10, pp. 382–392.

- Duijzer, L. E., Jaarsveld, W. L. van, Wallinga, J., and Dekker, R. (2018). "Dose-optimal vaccine allocation over multiple populations". In: *Production and operations management* 27.1, pp. 143–159.
- Dutta, B. L., Ezanno, P., and Vergu, E. (2014). "Characteristics of the spatio-temporal network of cattle movements in France over a 5-year period". In: *Preventive Veterinary Medicine* 117.1, pp. 79–94.
- Dvijotham, K. and Todorov, E. (2013). "Linearly solvable optimal control". In: *Reinforcement learning and approximate dynamic programming for feedback control*. Vol. 17. Wiley Online Library, pp. 119–141.
- Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1952). "The Inventory Problem: I. Case of Known Distributions of Demand". In: *Econometrica* 20.2. Publisher: [Wiley, Econometric Society], pp. 187–222. ISSN: 0012-9682. DOI: [10.2307/1907847](https://doi.org/10.2307/1907847).
- Ellsberg, D. (1961). "Risk, Ambiguity, and the Savage Axioms". In: *The Quarterly Journal of Economics* 75.4, p. 643. ISSN: 00335533. DOI: [10.2307/1884324](https://doi.org/10.2307/1884324).
- Enayati, S. and Özaltn, O. Y. (2020). "Optimal influenza vaccine distribution with equity". In: *European Journal of Operational Research* 283.2, pp. 714–725.
- Enright, J. and Kao, R. R. (2018). "Epidemics on dynamic networks". In: *Epidemics* 24, pp. 88–97. ISSN: 1755-4365. DOI: <https://doi.org/10.1016/j.epidem.2018.04.003>.
- Erdős, P. and Rényi, A. (2011). "On the evolution of random graphs". In: *The Structure and Dynamics of Networks*. Princeton University Press, pp. 38–82. DOI: [doi:10.1515/9781400841356.38](https://doi.org/10.1515/9781400841356.38).
- Ethier, S. N. and Kurtz, T. G. (2009). *Markov Processes: Characterization and Convergence*. en. John Wiley & Sons. ISBN: 978-0-470-31732-7.
- Etner, J., Jeleva, M., and Tallon, J.-M. (2009). "Decision theory under uncertainty". en. In: p. 50.
- Euler, L. (1794). *Institutiones calculi integralis*. la. Academia Imperialis Scientiarum.
- Ezanno, P., Andraud, M., Beaunée, G., Hoch, T., Krebs, S., Rault, A., Touzeau, S., Vergu, E., and Widgren, S. (2020). "How mechanistic modelling supports decision making for the control of enzootic infectious diseases". In: *Epidemics* 32, p. 100398.
- Ezanno, P., Fourichon, C., Beaudeau, F., and Seegers, H. (2006). "Between-herd movements of cattle as a tool for evaluating the risk of introducing infected animals". In: *Animal Research* 55.3, pp. 189–208.
- Ezanno, P., Fourichon, C., and Seegers, H. (2008). "Influence of herd structure and type of virus introduction on the spread of bovine viral diarrhoea virus (BVDV) on the spread of bovine viral diarrhoea virus (BVDV) within a dairy herd". en. In: *Veterinary Research* 39.5. Publisher: EDP Sciences, p. 1. ISSN: 0928-4249, 1297-9716. DOI: [10.1051/vetres:2008016](https://doi.org/10.1051/vetres:2008016).
- Fefferman, N. H. and Ng, K. L. (2007). "How disease models in static networks can fail to approximate disease in dynamic networks". en. In: *Physical Review E* 76.3, p. 031919. ISSN: 1539-3755, 1550-2376. DOI: [10.1103/PhysRevE.76.031919](https://doi.org/10.1103/PhysRevE.76.031919).
- Fekom, M., Vayatis, N., and Kalogeratos, A. (2020). "Dynamic Epidemic Control via Sequential Resource Allocation". en. In: *arXiv:2006.07199 [cs, math, stat]*. arXiv: 2006.07199.
- Feng, Z., Hill, A. N., Curns, A. T., and Glasser, J. W. (2017). "Evaluating targeted interventions via meta-population models with multi-level mixing". eng. In: *Mathematical Biosciences* 287, pp. 93–104. ISSN: 1879-3134. DOI: [10.1016/j.mbs.2016.09.013](https://doi.org/10.1016/j.mbs.2016.09.013).
- Ferguson, N. M., Donnelly, C. A., and Anderson, R. M. (2001). "The Foot-and-Mouth Epidemic in Great Britain: Pattern of Spread and Impact of Interventions". EN. In: *Science*. Publisher: American Association for the Advancement of Science. DOI: [10.1126/science.1061020](https://doi.org/10.1126/science.1061020).
- Ferguson, T. S. (1989). "Who Solved the Secretary Problem?" In: *Statistical Science* 4.3. Publisher: Institute of Mathematical Statistics, pp. 282–289. ISSN: 0883-4237, 2168-8745. DOI: [10.1214/ss/1177012493](https://doi.org/10.1214/ss/1177012493).
- Fèvre, E. M., C. Bronsvoort, B. M. de, Hamilton, K. A., and Cleaveland, S. (2006). "Animal movements and the spread of infectious diseases". In: *Trends in Microbiology* 14.3, pp. 125–131. ISSN: 0966-842X. DOI: <https://doi.org/10.1016/j.tim.2006.01.004>.
- Forien, R., Pang, G., and Pardoux, E. (2021). "Recent Advances in Epidemic Modeling: Non-Markov Stochastic Models and their Scaling Limits". en. In: *arXiv:2106.08466 [math]*. arXiv: 2106.08466.

- Forster, G. A. and Gilligan, C. A. (2007). "Optimizing the control of disease infestations at the landscape scale". In: *Proceedings of the National Academy of Sciences* 104.12, pp. 4984–4989.
- Fray, M. D., Paton, D. J., and Alenius, S. (2000). "The effects of bovine viral diarrhoea virus on cattle reproduction in relation to disease control". en. In: *Animal Reproduction Science* 60-61, pp. 615–627. ISSN: 0378-4320. DOI: [10.1016/S0378-4320\(00\)00082-8](https://doi.org/10.1016/S0378-4320(00)00082-8).
- Fu, F., Rosenbloom, D. I., Wang, L., and Nowak, M. A. (2010). "Imitation dynamics of vaccination behaviour on social networks". In: *Proceedings of the Royal Society B: Biological Sciences* 278.1702, pp. 42–49.
- Fukuda, E., Kokubo, S., Tanimoto, J., Wang, Z., Hagishima, A., and Ikegaya, N. (2014). "Risk assessment for infectious disease and its impact on voluntary vaccination behavior in social networks". en. In: *Chaos, Solitons & Fractals* 68, pp. 1–9. ISSN: 0960-0779. DOI: [10.1016/j.chaos.2014.07.004](https://doi.org/10.1016/j.chaos.2014.07.004).
- Funk, S., Gilad, E., Watkins, C., and Jansen, V. A. (2009). "The spread of awareness and its impact on epidemic outbreaks". In: *Proceedings of the National Academy of Sciences* 106.16, pp. 6872–6877.
- Gates, M. C., Humphry, R. W., and Gunn, G. J. (2013a). "Associations between bovine viral diarrhoea virus (BVDV) seropositivity and performance indicators in beef suckler and dairy herds". en. In: *The Veterinary Journal* 198.3, pp. 631–637. ISSN: 1090-0233. DOI: [10.1016/j.tvjl.2013.09.017](https://doi.org/10.1016/j.tvjl.2013.09.017).
- Gates, M. C., Woolhouse, M. E. J., Gunn, G. J., and Humphry, R. W. (2013b). "Relative associations of cattle movements, local spread, and biosecurity with bovine viral diarrhoea virus (BVDV) seropositivity in beef and dairy herds". en. In: *Preventive Veterinary Medicine* 112.3, pp. 285–295. ISSN: 0167-5877. DOI: [10.1016/j.prevetmed.2013.07.017](https://doi.org/10.1016/j.prevetmed.2013.07.017).
- GDS de Bourgogne - Franche-Comté (2022). URL: <https://www.gdsbfc.org/> (visited on 01/08/2022).
- Gethmann, J., Probst, C., Bassett, J., Blunk, P., Hövel, P., and Conraths, F. J. (2019). "An Epidemiological and Economic Simulation Model to Evaluate Strategies for the Control of Bovine Virus Diarrhea in Germany". In: *Frontiers in Veterinary Science* 6, p. 406. ISSN: 2297-1769. DOI: [10.3389/fvets.2019.00406](https://doi.org/10.3389/fvets.2019.00406).
- Gilbert, M., Mitchell, A., Bourn, D., Mawdsley, J., Clifton-Hadley, R., and Wint, W. (2005). "Cattle movements and bovine tuberculosis in Great Britain". In: *Nature* 435.7041, pp. 491–496.
- Gillespie, D. T. (1976). "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions". In: *Journal of computational physics* 22.4, pp. 403–434.
- Gillespie, D. T. (2001). "Approximate accelerated stochastic simulation of chemically reacting systems". In: *The Journal of chemical physics* 115.4, pp. 1716–1733.
- Graham, D. A., Clegg, T. A., Thulke, H. .-, O'Sullivan, P., McGrath, G., and More, S. J. (2016). "Quantifying the risk of spread of bovine viral diarrhoea virus (BVDV) between contiguous herds in Ireland". en. In: *Preventive Veterinary Medicine* 126, pp. 30–38. ISSN: 0167-5877. DOI: [10.1016/j.prevetmed.2016.01.017](https://doi.org/10.1016/j.prevetmed.2016.01.017).
- Gunst, R. F. and Mason, R. L. (2009). "Fractional factorial design". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 1.2, pp. 234–244.
- Guy, R., Larédo, C., and Vergu, E. (2015). "Approximation of epidemic models by diffusion processes and their statistical inference". en. In: *Journal of Mathematical Biology* 70.3, pp. 621–646. ISSN: 1432-1416. DOI: [10.1007/s00285-014-0777-8](https://doi.org/10.1007/s00285-014-0777-8).
- Hagenaars, T. J., Donnelly, C. A., and Ferguson, N. M. (2004). "Spatial heterogeneity and the persistence of infectious diseases". en. In: *Journal of Theoretical Biology* 229.3, pp. 349–359. ISSN: 0022-5193. DOI: [10.1016/j.jtbi.2004.04.002](https://doi.org/10.1016/j.jtbi.2004.04.002).
- Hamer, W. H. (1906). *Epidemic disease in England: the evidence of variability and of persistency of type*. Bedford Press.
- Hansson, H. and Lagerkvist, C. J. (2015). "Identifying use and non-use values of animal welfare: Evidence from Swedish dairy agriculture". In: *Food Policy* 50, pp. 35–42.
- Heesterbeek, H., Anderson, R. M., Andreasen, V., Bansal, S., De Angelis, D., Dye, C., Eames, K. T., Edmunds, W. J., Frost, S. D., Funk, S., et al. (2015). "Modeling infectious disease dynamics in the complex landscape of global health". In: *Science* 347.6227.

- Heino, M., Kaitala, V., Ranta, E., and Lindström, J. (1997). "Synchronous dynamics and rates of extinction in spatially structured populations". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 264.1381, pp. 481–486.
- Hershey, J. C., Asch, D. A., Thumasathit, T., Meszaros, J., and Waters, V. V. (1994). "The roles of altruism, free riding, and bandwagoning in vaccination decisions". In: *Organizational Behavior and Human Decision Processes* 59.2, pp. 177–187.
- Hethcote, H. W. (2000). "The Mathematics of Infectious Diseases". In: *SIAM Review* 42.4. Publisher: Society for Industrial and Applied Mathematics, pp. 599–653. ISSN: 0036-1445.
- Hidano, A., Enticott, G., Christley, R. M., and Gates, M. C. (2018). "Modeling dynamic human behavioral changes in animal disease models: challenges and opportunities for addressing bias". In: *Frontiers in Veterinary Science* 5, p. 137.
- Hill, A. N. and Longini, I. M. (2003). "The critical vaccination fraction for heterogeneous epidemic models". en. In: *Mathematical Biosciences* 181.1, pp. 85–106. ISSN: 00255564. DOI: [10.1016/S0025-5564\(02\)00129-3](https://doi.org/10.1016/S0025-5564(02)00129-3).
- Hill, E. M., Prosser, N. S., Ferguson, E., Kaler, J., Green, M. J., Keeling, M. J., and Tildesley, M. J. (2021). "Modelling livestock infectious disease control policy under differing social perspectives on vaccination behaviour." en-US. In.
- Holme, P. and Litvak, N. (2017). "Cost-efficient vaccination protocols for network epidemiology". In: *PLoS computational biology* 13.9, e1005696.
- Holme, P. and Saramäki, J. (2012). "Temporal networks". en. In: *Physics Reports* 519.3, pp. 97–125. ISSN: 03701573. DOI: [10.1016/j.physrep.2012.03.001](https://doi.org/10.1016/j.physrep.2012.03.001).
- Horan, R. D., Fenichel, E. P., Wolf, C. A., and Gramig, B. M. (2010). "Managing infectious animal disease systems". In: *Annual Review of Resource Economics* 2.1, pp. 101–124.
- Hoscheit, P., Anthony, É., and Vergu, E. (2021). "Dynamic centrality measures for cattle trade networks". In: *Applied Network Science* 6.1, pp. 1–17.
- Hotaling, J. M., Fakhari, P., and Busemeyer, J. R. (2015). "Dynamic Decision Making". en. In: *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier, pp. 708–713. ISBN: 978-0-08-097087-5. DOI: [10.1016/B978-0-08-097086-8.43040-0](https://doi.org/10.1016/B978-0-08-097086-8.43040-0).
- Houe, H. (1993). "Survivorship of animals persistently infected with bovine virus diarrhoea virus (BVDV)". en. In: *Preventive Veterinary Medicine* 15.4, pp. 275–283. ISSN: 0167-5877. DOI: [10.1016/0167-5877\(93\)90099-F](https://doi.org/10.1016/0167-5877(93)90099-F).
- Huang, F., Cao, M., and Wang, L. (2020). "Learning enables adaptation in cooperation for multi-player stochastic games". en. In: *Journal of The Royal Society Interface* 17.172, p. 20200639. ISSN: 1742-5689, 1742-5662. DOI: [10.1098/rsif.2020.0639](https://doi.org/10.1098/rsif.2020.0639).
- Ibuka, Y., Li, M., Vietri, J., Chapman, G. B., and Galvani, A. P. (2014). "Free-riding behavior in vaccination decisions: an experimental study". In: *PLOS One* 9.1, e87164.
- Innocent, G., Morrison, I., Brownlie, J., and Gettinby, G. (1997). "A computer simulation of the transmission dynamics and the effects of duration of immunity and survival of persistently infected animals on the spread of bovine viral diarrhoea virus in dairy cattle". eng. In: *Epidemiology and Infection* 119.1, pp. 91–100. ISSN: 0950-2688. DOI: [10.1017/s0950268897007723](https://doi.org/10.1017/s0950268897007723).
- Jesse, M., Ezanno, P., Davis, S., and Heesterbeek, J. A. P. (2008). "A fully coupled, mechanistic model for infectious disease dynamics in a metapopulation: Movement and epidemic duration". en. In: *Journal of Theoretical Biology* 254.2, pp. 331–338. ISSN: 0022-5193. DOI: [10.1016/j.jtbi.2008.05.038](https://doi.org/10.1016/j.jtbi.2008.05.038).
- Jesse, M. and Heesterbeek, H. (2011). "Divide and conquer? Persistence of infectious agents in spatial metapopulations of hosts". In: *Journal of theoretical biology* 275.1, pp. 12–20.
- Jijón, S., Supervie, V., and Breban, R. (2017). "Prevention of treatable infectious diseases: A game-theoretic approach". en. In: *Vaccine* 35.40, pp. 5339–5345. ISSN: 0264-410X. DOI: [10.1016/j.vaccine.2017.08.040](https://doi.org/10.1016/j.vaccine.2017.08.040).
- Kahneman, D. and Tversky, A. (1979). "Prospect Theory: An Analysis of Decision under Risk". en. In: *Econometrica* 47.2, p. 263. ISSN: 00129682. DOI: [10.2307/1914185](https://doi.org/10.2307/1914185).
- Kao, R. R. (2003). "The impact of local heterogeneity on alternative control strategies for foot-and-mouth disease". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270.1533, pp. 2557–2564.

- Kao, R. R., Green, D. M., Johnson, J., and Kiss, I. Z. (2007). "Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the UK". In: *Journal of the Royal Society Interface* 4.16, pp. 907–916.
- Karlsson, C.-J. and Rowlett, J. (2020). "Decisions and disease: a mechanism for the evolution of cooperation". In: *Scientific Reports* 10.1, pp. 1–9.
- Karni, E. (2008). "Savage's Subjective Expected Utility Model". en. In: *The New Palgrave Dictionary of Economics*. Ed. by Palgrave Macmillan. London: Palgrave Macmillan UK, pp. 1–5. ISBN: 978-1-349-95121-5. DOI: [10.1057/978-1-349-95121-5_2467-1](https://doi.org/10.1057/978-1-349-95121-5_2467-1).
- Kaufmann, E. (2014). "Analyse de stratégies bayésiennes et fréquentistes pour l'allocation séquentielle de ressources". PhD thesis. Paris, ENST.
- Keeling, M. J. (2000). "Metapopulation moments: coupling, stochasticity and persistence". In: *Journal of Animal Ecology* 69.5, pp. 725–736.
- Keeling, M. J. and Eames, K. T. (2005). "Networks and epidemic models". In: *Journal of the Royal Society Interface* 2.4, pp. 295–307.
- Keeling, M. J. and Gilligan, C. (2000). "Bubonic plague: a metapopulation model of a zoonosis". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 267.1458, pp. 2219–2230.
- Keeling, M. J. and Rohani, P. (2011). *Modeling infectious diseases in humans and animals*. Princeton University Press.
- Kendrick, J. W. (1971). "Bovine viral diarrhoea-mucosal disease virus infection in pregnant cows". eng. In: *American Journal of Veterinary Research* 32.4, pp. 533–544. ISSN: 0002-9645.
- Kermack, W. O., McKendrick, A. G., and Walker, G. T. (1932). "Contributions to the mathematical theory of epidemics. II. —The problem of endemicity". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 138.834. Publisher: Royal Society, pp. 55–83. DOI: [10.1098/rspa.1932.0171](https://doi.org/10.1098/rspa.1932.0171).
- Kermack, W. O. and McKendrick, A. G. (1927). "A contribution to the mathematical theory of epidemics". In: *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character* 115.772, pp. 700–721.
- Kim, H. and Anderson, R. (2012). "Temporal node centrality in complex networks". In: *Physical Review E* 85.2, p. 026107.
- Kiss, I. Z., Cassell, J., Recker, M., and Simon, P. L. (2010). "The impact of information transmission on epidemic outbreaks". In: *Mathematical Biosciences* 225.1, pp. 1–10.
- Kiss, I. Z., Miller, J. C., Simon, P. L., et al. (2017). "Mathematics of epidemics on networks". In: *Cham: Springer* 598.
- Klovdahl, A. S. (1985). "Social networks and the spread of infectious diseases: The AIDS example". en. In: *Social Science & Medicine* 21.11, pp. 1203–1216. ISSN: 0277-9536. DOI: [10.1016/0277-9536\(85\)90269-2](https://doi.org/10.1016/0277-9536(85)90269-2).
- Kobilinsky, A., Bouvier, A., and Monod, H. (2020). *PLANOR: an R package for the automatic generation of regular fractional factorial designs*. R package version 1.5-3. INRA, MIA, Jouy en Josas, France.
- Kochenderfer, M. J. (2015). *Decision Making Under Uncertainty: Theory and Application*. en. Ed. by Ward, J. MIT Lincoln Laboratory Series. Cambridge, MA, USA: MIT Press. ISBN: 978-0-262-02925-4.
- Krebs, S., Picault, S., and Ezanno, P. (2018). *Modélisation multi-agents pour la gestion individuelle et collective d'une maladie infectieuse*. Ed. by Picard, G. Métabief, France: Christophe Lang, Nicolas Marilleau.
- Kreps, D. M. (1997). "Intrinsic motivation and extrinsic incentives". In: *The American economic review* 87.2, pp. 359–364.
- Kuga, K., Tanimoto, J., and Jusup, M. (2019). "To vaccinate or not to vaccinate: A comprehensive study of vaccination-subsidizing policies with multi-agent simulations and mean-field modeling". In: *Journal of Theoretical Biology* 469, pp. 107–126.
- Lagerkvist, C. J., Hansson, H., Hess, S., and Hoffman, R. (2011). "Provision of farm animal welfare: Integrating productivity and non-use values". In: *Applied Economic Perspectives and Policy* 33.4, pp. 484–509.
- Lamboni, M., Monod, H., and Makowski, D. (2011). "Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models". In: *Reliability Engineering & System Safety* 96.4, pp. 450–459.

- Lasry, J.-M. and Lions, P.-L. (2007). "Mean field games". en. In: *Japanese Journal of Mathematics* 2.1, pp. 229–260. ISSN: 1861-3624. DOI: [10.1007/s11537-007-0657-8](https://doi.org/10.1007/s11537-007-0657-8).
- Lefrançois, T. and Pineau, T. (2014). "Public health and livestock: Emerging diseases in food animals". In: *Animal Frontiers* 4.1, pp. 4–6. ISSN: 2160-6056. DOI: [10.2527/af.2014-0001](https://doi.org/10.2527/af.2014-0001).
- Lemaitre, J. C., Pasetto, D., Zanon, M., Bertuzzo, E., Mari, L., Miccoli, S., Casagrandi, R., Gatto, M., and Rinaldo, A. (2021). *Optimizing the spatio-temporal allocation of COVID-19 vaccines: Italy as a case study*. en. preprint. *Epidemiology*. DOI: [10.1101/2021.05.06.21256732](https://doi.org/10.1101/2021.05.06.21256732).
- Lenhart, S. and Workman, J. T. (2007). *Optimal Control Applied to Biological Models*. New York: Chapman and Hall/CRC. ISBN: 978-0-429-13805-8. DOI: [10.1201/9781420011418](https://doi.org/10.1201/9781420011418).
- Levandowsky, M. and Winter, D. (1971). "Distance between sets". In: *Nature* 234.5323, pp. 34–35.
- Levins, R. (1969). "Some demographic and genetic consequences of environmental heterogeneity for biological control". In: *American Entomologist* 15.3, pp. 237–240.
- Li, M., Liu, R.-R., Lü, L., Hu, M.-B., Xu, S., and Zhang, Y.-C. (2021). "Percolation on complex networks: Theory and application". In: *Physics Reports* 907, pp. 1–68.
- Li, Q., Li, M., Lv, L., Guo, C., and Lu, K. (2017). "A new prediction model of infectious diseases with vaccination strategies based on evolutionary game theory". en. In: *Chaos, Solitons & Fractals* 104, pp. 51–60. ISSN: 0960-0779. DOI: [10.1016/j.chaos.2017.07.022](https://doi.org/10.1016/j.chaos.2017.07.022).
- Liu, X.-T., Wu, Z.-X., and Zhang, L. (2012). "Impact of committed individuals on vaccination behavior". In: *Physical Review E* 86.5. Publisher: American Physical Society, p. 051132. DOI: [10.1103/PhysRevE.86.051132](https://doi.org/10.1103/PhysRevE.86.051132).
- Lloyd, A. L. (2001). "Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268.1470, pp. 985–993.
- Lloyd, A. L. and Jansen, V. A. A. (2004). "Spatiotemporal dynamics of epidemics: synchrony in metapopulation models". en. In: *Mathematical Biosciences. Topics in Biomathematics and Related Computational Problems: selected papers* 188.1, pp. 1–16. ISSN: 0025-5564. DOI: [10.1016/j.mbs.2003.09.003](https://doi.org/10.1016/j.mbs.2003.09.003).
- Long, E. F., Nohdurft, E., and Spinler, S. (2018). "Spatial resource allocation for emerging epidemics: A comparison of greedy, myopic, and dynamic policies". In: *Manufacturing & Service Operations Management* 20.2, pp. 181–198.
- Lorch, L., De, A., Bhatt, S., Trouleau, W., Upadhyay, U., and Gomez-Rodriguez, M. (2018). "Stochastic optimal control of epidemic processes in networks". In: *arXiv preprint arXiv:1810.13043*.
- Machina, M. J. (1987). "Choice Under Uncertainty: Problems Solved and Unsolved". en. In: p. 63.
- Manfredi, P. and D'Onofrio, A. (2013). *Modeling the interplay between human behavior and the spread of infectious diseases*. Springer Science & Business Media.
- Martinez-Lopez, B., Perez, A., and Sanchez-Vizcano, J. (2009). "Combined application of social network and cluster detection analyses for temporal-spatial characterization of animal movements in Salamanca, Spain". In: *Preventive veterinary medicine* 91.1, pp. 29–38.
- Masuda, N. and Lambiotte, R. (2016). *A guide to temporal networks*. World Scientific.
- Matos Baptista, F. and Nunes, T. (2007). "Spatial analysis of cattle movement patterns in Portugal". In: *Veterinaria Italiana* 43.3, pp. 611–619.
- Matrajt, L., Halloran, M. E., and Longini Jr, I. M. (2013). "Optimal vaccine allocation for the early mitigation of pandemic influenza". In: *PLoS Comput Biol* 9.3, e1002964.
- May, R. M. and Anderson, R. M. (1987). "Commentary transmission dynamics of HIV infection". In: *Nature* 326.137, pp. 10–1038.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. en. Cambridge University Press. ISBN: 978-0-521-28884-2.
- Mbah, M. L. N. and Gilligan, C. A. (2011). "Resource allocation for epidemic control in metapopulations". In: *PLoS one* 6.9, e24577.
- McCallum, H., Barlow, N., and Hone, J. (2001). "How should pathogen transmission be modelled?" In: *Trends in Ecology & Evolution* 16.6, pp. 295–300. ISSN: 0169-5347. DOI: [https://doi.org/10.1016/S0169-5347\(01\)02144-9](https://doi.org/10.1016/S0169-5347(01)02144-9).
- McClurkin, A. W., Littledike, E. T., Cutlip, R. C., Frank, G. H., Coria, M. F., and Bolin, S. R. (1984). "Production of cattle immunotolerant to bovine viral diarrhoea virus." In: *Canadian Journal of Comparative Medicine* 48.2, pp. 156–161. ISSN: 0008-4050.

- McGowan, M. R., Kirkland, P. D., Richards, S. G., and Littlejohns, I. R. (1993). "Increased reproductive losses in cattle infected with bovine pestivirus around the time of insemination". eng. In: *The Veterinary Record* 133.2, pp. 39–43. ISSN: 0042-4900. DOI: [10.1136/vr.133.2.39](https://doi.org/10.1136/vr.133.2.39).
- Méléard, S. (2016). *Modèles aléatoires en Ecologie et Evolution*. Springer.
- Mendes, Â. J., Haydon, D. T., McIntosh, E., Hanley, N., and Halliday, J. E. B. (2020). "Socially vs. Privately Optimal Control of Livestock Diseases: A Case for Integration of Epidemiology and Economics". English. In: *Frontiers in Veterinary Science* 7. Publisher: Frontiers. ISSN: 2297-1769. DOI: [10.3389/fvets.2020.558409](https://doi.org/10.3389/fvets.2020.558409).
- Meng, X., Han, S., Wu, L., Si, S., and Cai, Z. (2022). "Analysis of epidemic vaccination strategies by node importance and evolutionary game on complex networks". en. In: *Reliability Engineering & System Safety* 219, p. 108256. ISSN: 0951-8320. DOI: [10.1016/j.ress.2021.108256](https://doi.org/10.1016/j.ress.2021.108256).
- Mieghem, P., Stevanović, D., Kuipers, F., Li, C., Bovenkamp, R., Liu, D., and Wang, H. (2011). "Decreasing the spectral radius of a graph by link removals". In: *Physical review. E, Statistical, nonlinear, and soft matter physics* 84, p. 016101. DOI: [10.1103/PhysRevE.84.016101](https://doi.org/10.1103/PhysRevE.84.016101).
- Moennig, V. and Becher, P. (2018). "Control of Bovine Viral Diarrhea". eng. In: *Pathogens (Basel, Switzerland)* 7.1, E29. ISSN: 2076-0817. DOI: [10.3390/pathogens7010029](https://doi.org/10.3390/pathogens7010029).
- Moerman, A., Straver, P. J., Jong, M. C. de, Quak, J., Baanvinger, T., and Oirschot, J. T. van (1993). "A long term epidemiological study of bovine viral diarrhoea infections in a large herd of dairy cattle". eng. In: *The Veterinary Record* 132.25, pp. 622–626. ISSN: 0042-4900. DOI: [10.1136/vr.132.25.622](https://doi.org/10.1136/vr.132.25.622).
- Mohr, S., Beard, R., Nisbet, A. J., Burgess, S. T. G., Reeve, R., Denwood, M., Porphyre, T., Zadoks, R. N., and Matthews, L. (2020). "Uptake of Diagnostic Tests by Livestock Farmers: A Stochastic Game Theory Approach". In: *Frontiers in Veterinary Science* 7, p. 36. ISSN: 2297-1769. DOI: [10.3389/fvets.2020.00036](https://doi.org/10.3389/fvets.2020.00036).
- Montgomery, D. C. (2017). *Design and analysis of experiments*. John Wiley & Sons.
- Morin, B. R., Fenichel, E. P., and Castillo-Chavez, C. (2013). "SIR dynamics with economically driven contact rates". In: *Natural Resource Modeling* 26.4, pp. 505–525.
- Mortensen, S., Stryhn, H., Søgaard, R., Boklund, A., Stärk, K. D., Christensen, J., and Willeberg, P. (2002). "Risk factors for infection of sow herds with porcine reproductive and respiratory syndrome (PRRS) virus". In: *Preventive Veterinary Medicine* 53.1-2, pp. 83–101.
- Murray, J. (2003). *II. Spatial models and biomedical applications*. Springer.
- Mweu, M. M., Fournié, G., Halasa, T., Toft, N., and Nielsen, S. S. (2013). "Temporal characterisation of the network of Danish cattle movements and its implication for disease control: 2000–2009". In: *Preventive veterinary medicine* 110.3-4, pp. 379–387.
- Myerson, R. B. (1997). *Game Theory: Analysis of Conflict*. en. Harvard University Press. ISBN: 978-0-674-34116-6.
- Nagashima, K. and Tanimoto, J. (2019). "A stochastic Pairwise Fermi rule modified by utilizing the average in payoff differences of neighbors leads to increased network reciprocity in spatial prisoner's dilemma games". In: *Applied Mathematics and Computation* 361, pp. 661–669.
- Nåsell, I. (1999). "On the time to extinction in recurrent epidemics". en. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61.2, pp. 309–330. ISSN: 1467-9868. DOI: [10.1111/1467-9868.00178](https://doi.org/10.1111/1467-9868.00178).
- Natale, F., Giovannini, A., Savini, L., Palma, D., Possenti, L., Fiore, G., and Calistri, P. (2009). "Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread". In: *Preventive veterinary medicine* 92.4, pp. 341–350.
- Newman, M. (2010). *Networks: An Introduction*. en. OUP Oxford. ISBN: 978-0-19-150070-1.
- Newman, M. (2018). *Networks*. Oxford university press.
- Newton, J. (2018). "Evolutionary Game Theory: A Renaissance". en. In: *Games* 9.2. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, p. 31. DOI: [10.3390/g9020031](https://doi.org/10.3390/g9020031).
- Nigsch, A., Costard, S., Jones, B. A., Pfeiffer, D. U., and Wieland, B. (2013). "Stochastic spatio-temporal modelling of African swine fever spread in the European Union during the high risk period". en. In: *Preventive Veterinary Medicine*. Special Issue: SVEPM 2012 - The cutting-edge of animal disease control in a global environment 108.4, pp. 262–275. ISSN: 0167-5877. DOI: [10.1016/j.prevetmed.2012.11.003](https://doi.org/10.1016/j.prevetmed.2012.11.003).
- Ning, Y.-Z., Liu, X., Cheng, H.-M., and Zhang, Z.-Y. (2020). "Effects of social network structures and behavioral responses on the spread of infectious diseases". en. In: *Physica A: Statistical*

- Mechanics and its Applications* 539, p. 122907. ISSN: 0378-4371. DOI: [10.1016/j.physa.2019.122907](https://doi.org/10.1016/j.physa.2019.122907).
- Nöremark, M., Håkansson, N., Lewerin, S. S., Lindberg, A., and Jonsson, A. (2011). "Network analysis of cattle and pig movements in Sweden: measures relevant for disease control and risk based surveillance". In: *Preventive veterinary medicine* 99.2-4, pp. 78–90.
- Nowak, M. A. (2006). "Five rules for the evolution of cooperation". In: *Science* 314.5805, pp. 1560–1563.
- Nowzari, C., Preciado, V. M., and Pappas, G. J. (2015). "Analysis and Control of Epidemics: A survey of spreading processes on complex networks". en. In: *arXiv:1505.00768 [physics]*. arXiv: 1505.00768.
- Oraby, T. and Bauch, C. (2015). "Bounded rationality alters the dynamics of paediatric immunization acceptance". English. In: *Scientific Reports* 5. ISSN: 2045-2322. DOI: [10.1038/srep10724](https://doi.org/10.1038/srep10724).
- Otte, J. and Pica-Ciamarra, U. (2021). "Emerging infectious zoonotic diseases: The neglected role of food animals". en. In: *One Health* 13, p. 100323. ISSN: 2352-7714. DOI: [10.1016/j.onehlt.2021.100323](https://doi.org/10.1016/j.onehlt.2021.100323).
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). "Epidemic processes in complex networks". In: *Reviews of Modern Physics* 87.3. Publisher: American Physical Society, pp. 925–979. DOI: [10.1103/RevModPhys.87.925](https://doi.org/10.1103/RevModPhys.87.925).
- Pastor-Satorras, R. and Vespignani, A. (2001). "Epidemic Spreading in Scale-Free Networks". In: *Physical Review Letters* 86.14. Publisher: American Physical Society, pp. 3200–3203. DOI: [10.1103/PhysRevLett.86.3200](https://doi.org/10.1103/PhysRevLett.86.3200).
- Pastor-Satorras, R. and Vespignani, A. (2002). "Immunization of complex networks". en. In: *Physical Review E* 65.3. arXiv: cond-mat/0107066, p. 036104. ISSN: 1063-651X, 1095-3787. DOI: [10.1103/PhysRevE.65.036104](https://doi.org/10.1103/PhysRevE.65.036104).
- Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., and Trapman, P. (2015). "Eight challenges for network epidemic models". In: *Epidemics* 10, pp. 58–62.
- Perrings, C., Castillo-Chavez, C., Chowell, G., Daszak, P., Fenichel, E. P., Finnoff, D., Horan, R. D., Kilpatrick, A. M., Kinzig, A. P., Kuminoff, N. V., Levin, S., Morin, B., Smith, K. F., and Springborn, M. (2014). "Merging Economics and Epidemiology to Improve the Prediction and Management of Infectious Disease". en. In: *EcoHealth* 11.4, pp. 464–475. ISSN: 1612-9202, 1612-9210. DOI: [10.1007/s10393-014-0963-6](https://doi.org/10.1007/s10393-014-0963-6).
- Pingle, M. (1995). "Imitation versus rationality: An experimental perspective on decision making". en. In: *The Journal of Socio-Economics* 24.2, pp. 281–315. ISSN: 10535357. DOI: [10.1016/1053-5357\(95\)90023-3](https://doi.org/10.1016/1053-5357(95)90023-3).
- Pinior, B., Firth, C. L., Richter, V., Lebl, K., Trauffler, M., Dzieciol, M., Hutter, S. E., Burgstaller, J., Obritzhauser, W., Winter, P., and Käsbohrer, A. (2017). "A systematic review of financial and economic assessments of bovine viral diarrhoea virus (BVDV) prevention and mitigation activities worldwide". en. In: *Preventive Veterinary Medicine* 137, pp. 77–92. ISSN: 0167-5877. DOI: [10.1016/j.prevetmed.2016.12.014](https://doi.org/10.1016/j.prevetmed.2016.12.014).
- Poundstone, W. (1992). *Prisoner's Dilemma*. en. Doubleday. ISBN: 978-0-385-41567-5.
- Preciado, V. M., Zargham, M., Enyioha, C., Jadbabaie, A., and Pappas, G. (2013). "Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks". In: *52nd IEEE Conference on Decision and Control*, pp. 7486–7491. DOI: [10.1109/CDC.2013.6761078](https://doi.org/10.1109/CDC.2013.6761078).
- Probert, W. J., Lakkur, S., Fonnesbeck, C. J., Shea, K., Runge, M. C., Tildesley, M. J., and Ferrari, M. J. (2019). "Context matters: using reinforcement learning to develop human-readable, state-dependent outbreak response policies". In: *Philosophical Transactions of the Royal Society B* 374.1776, p. 20180277.
- Puterman, M. L. (2014). *Markov Decision Processes.: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.
- Qi, L., Beaunée, G., Arnoux, S., Dutta, B. L., Joly, A., Vergu, E., and Ezanno, P. (2019). "Neighbourhood contacts and trade movements drive the regional spread of bovine viral diarrhoea virus (BVDV)". In: *Veterinary Research* 50.1, p. 30. ISSN: 1297-9716. DOI: [10.1186/s13567-019-0647-x](https://doi.org/10.1186/s13567-019-0647-x).
- Rao, I. J. and Brandeau, M. L. (2021). "Optimal allocation of limited vaccine to control an infectious disease: Simple analytical conditions". In: *Mathematical biosciences* 337, p. 108621.

- Rat-Aspert, O. and Fourichon, C. (2010). "Modelling collective effectiveness of voluntary vaccination with and without incentives". In: *Preventive Veterinary Medicine* 93.4, pp. 265–275.
- Rautureau, S., Dufour, B., and Durand, B. (2011). "Vulnerability of Animal Trade Networks to The Spread of Infectious Diseases: A Methodological Approach Applied to Evaluation and Emergency Control Strategies in Cattle, France, 2005". en. In: *Transboundary and Emerging Diseases* 58.2, pp. 110–120. ISSN: 1865-1682. DOI: [10.1111/j.1865-1682.2010.01187.x](https://doi.org/10.1111/j.1865-1682.2010.01187.x).
- Reluga, T. C. and Galvani, A. P. (2011). "A general approach for population games with application to vaccination". en. In: *Mathematical Biosciences* 230.2, pp. 67–78. ISSN: 0025-5564. DOI: [10.1016/j.mbs.2011.01.003](https://doi.org/10.1016/j.mbs.2011.01.003).
- RL fundamentals - 2021 RL Virtual School (2022). URL: <https://rl-vs.github.io/rlvs2021/rl-fundamentals.html> (visited on 01/23/2022).
- Robinson, S. and Christley, R. (2007). "Exploring the role of auction markets in cattle movements within Great Britain". In: *Preventive Veterinary Medicine* 81.1. Lead papers of the eleventh symposium of the International Society for Veterinary Epidemiology and Economics (ISVEE), Cairns, Australia, pp. 21–37. ISSN: 0167-5877. DOI: <https://doi.org/10.1016/j.prevetmed.2007.04.011>.
- Rocha, L. E. and Masuda, N. (2014). "Random walk centrality for temporal networks". In: *New Journal of Physics* 16.6, p. 063023.
- Ross, R. (1910). *The prevention of malaria*. J. Murray.
- Rowthorn, R. E., Laxminarayan, R., and Gilligan, C. A. (2009). "Optimal control of epidemics in metapopulations". In: *Journal of the Royal Society Interface* 6.41, pp. 1135–1144.
- Ruget, A.-S., Rossi, G., Pepler, P. T., Beaunée, G., Banks, C. J., Enright, J., and Kao, R. R. (2021). "Multi-species temporal network of livestock movements for disease spread". en. In: *Applied Network Science* 6.1. Number: 1 Publisher: SpringerOpen, pp. 1–20. ISSN: 2364-8228. DOI: [10.1007/s41109-021-00354-x](https://doi.org/10.1007/s41109-021-00354-x).
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. John Wiley & Sons.
- Sankararaman, A., Ganesh, A., and Shakkottai, S. (2019). "Social Learning in Multi Agent Multi Armed Bandits". en. In: *arXiv:1910.02100 [cs, math, stat]*. arXiv: 1910.02100.
- Savage, L. J. (1972). *The Foundations of Statistics*. en. Courier Corporation. ISBN: 978-0-486-62349-8.
- Scaman, K., Kalogeratos, A., and Vayatis, N. (2014). "Dynamic Treatment Allocation for Epidemic Control in Arbitrary Networks". en. In: *Diffusion Networks and Cascade Analytics workshop*.
- Scaman, K., Kalogeratos, A., and Vayatis, N. (2015). "A greedy approach for dynamic control of diffusion processes in networks". In: *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, pp. 652–659.
- Scaman, K., Kalogeratos, A., and Vayatis, N. (2016). "Suppressing epidemics in networks using priority planning". In: *IEEE Transactions on Network Science and Engineering* 3.4, pp. 271–285.
- Schwartz, I. and Shaw, L. (2010). "Rewiring for adaptation". en. In: *Physics* 3, p. 17. ISSN: 1943-2879. DOI: [10.1103/Physics.3.17](https://doi.org/10.1103/Physics.3.17).
- Sent, E.-M. (2018). "Rationality and bounded rationality: you can't have one without the other". In: *The European Journal of the History of Economic Thought* 25.6, pp. 1370–1386. ISSN: 0967-2567. DOI: [10.1080/09672567.2018.1523206](https://doi.org/10.1080/09672567.2018.1523206).
- Sethi, S. P. and Staats, P. W. (1978). "Optimal Control of Some Simple Deterministic Epidemic Models". In: *Journal of the Operational Research Society* 29.2, pp. 129–136. ISSN: 0160-5682. DOI: [10.1057/jors.1978.27](https://doi.org/10.1057/jors.1978.27).
- Shafer, G. (2020). *A Mathematical Theory of Evidence*. en. Princeton University Press. ISBN: 978-0-691-21469-6.
- Shi, B., Liu, G., Qiu, H., Wang, Z., Ren, Y., and Chen, D. (2019). "Exploring voluntary vaccination with bounded rationality through reinforcement learning". In: *Physica A: Statistical Mechanics and its Applications* 515, pp. 171–182. ISSN: 0378-4371. DOI: <https://doi.org/10.1016/j.physa.2018.09.151>.
- Sok, J. and Fischer, E. A. J. (2020). "Farmers' heterogeneous motives, voluntary vaccination and disease spread: an agent-based model". In: *European Review of Agricultural Economics* 47.3, pp. 1201–1222. ISSN: 0165-1587. DOI: [10.1093/erae/jbz041](https://doi.org/10.1093/erae/jbz041).
- Sok, J., Lans, I. A. van der, Hogeveen, H., Elbers, A. R. W., and Oude Lansink, A. G. J. M. (2018). "Farmers' Preferences For Bluetongue Vaccination Scheme Attributes: An Integrated

- Choice and Latent Variable Approach". en. In: *Journal of Agricultural Economics* 69.2. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/1477-9552.12249>, pp. 537–560. ISSN: 1477-9552. DOI: [10.1111/1477-9552.12249](https://doi.org/10.1111/1477-9552.12249).
- Stauffer, D. and Aharony, A. (2018). *Introduction to percolation theory*. Taylor & Francis.
- Stott, A. W., Humphry, R. W., and Gunn, G. J. (2010). "Modelling the effects of previous infection and re-infection on the costs of bovine viral diarrhoea outbreaks in beef herds". en. In: *The Veterinary Journal* 185.2, pp. 138–143. ISSN: 1090-0233. DOI: [10.1016/j.tvjl.2009.05.020](https://doi.org/10.1016/j.tvjl.2009.05.020).
- Sturm, L. A., Mays, R. M., and Zimet, G. D. (2005). "Parental beliefs and decision making about child and adolescent immunization: from polio to sexually transmitted infections". eng. In: *Journal of developmental and behavioral pediatrics: JDBP* 26.6, pp. 441–452. ISSN: 0196-206X. DOI: [10.1097/00004703-200512000-00009](https://doi.org/10.1097/00004703-200512000-00009).
- Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- Sutton, R. S., McAllester, D. A., Singh, S. P., and Mansour, Y. (1999). "Policy Gradient Methods for Reinforcement Learning with Function Approximation". en. In: p. 7.
- Szabó, G. and Töke, C. (1998). "Evolutionary prisoner's dilemma game on a square lattice". In: *Physical Review E* 58.1, p. 69.
- Tago, D., Hammitt, J. K., Thomas, A., and Raboisson, D. (2016). "The impact of farmers' strategic behavior on the spread of animal infectious diseases". In: *PLOS One* 11.6, e0157450.
- Tanaka, M. and Tanimoto, J. (2021). "Proposal of an apposite strategy-updating rule for the vaccination game where hubs refer to hubs and lower-degree agents refer to lower-degree agents." In: *Biosystems* 209, p. 104532. ISSN: 0303-2647. DOI: <https://doi.org/10.1016/j.biosystems.2021.104532>.
- Tanimoto, J. (2021). "Evolutionary Game Theory: Fundamentals and Applications for Epidemiology". en. In: *Sociophysics Approach to Epidemics*. Ed. by Tanimoto, J. Evolutionary Economics and Social Complexity Science. Singapore: Springer, pp. 13–60. ISBN: 978-981-336-481-3. DOI: [10.1007/978-981-33-6481-3_2](https://doi.org/10.1007/978-981-33-6481-3_2).
- Tanner, M. W., Sattenspiel, L., and Ntaimo, L. (2008). "Finding optimal vaccination strategies under parameter uncertainty using stochastic programming". In: *Mathematical biosciences* 215.2, pp. 144–151.
- Taylor, D., Myers, S. A., Clauaset, A., Porter, M. A., and Mucha, P. J. (2017). "Eigenvector-based centrality measures for temporal networks". In: *Multiscale Modeling & Simulation* 15.1, pp. 537–574.
- Teitelbaum, J. and Zeiler, K. (2018). *Research handbook on behavioral law and economics*. eng. Research handbooks in law and economics. Northampton, MA: Edward Elgar Pub. ISBN: 978-1-84980-567-4.
- Terry, A. J. (2010). "Pulse vaccination strategies in a metapopulation SIR model". In: *Mathematical Biosciences & Engineering* 7.2, p. 455.
- Thompson, R. N., Gilligan, C. A., and Cunniffe, N. J. (2018). "Control fast or control smart: When should invading pathogens be controlled?" In: *PLoS computational biology* 14.2, e1006014.
- Tokic, M. and Palm, G. (2011). "Value-Difference Based Exploration: Adaptive Control between Epsilon-Greedy and Softmax". en. In: *KI 2011: Advances in Artificial Intelligence*. Ed. by Bach, J. and Edelkamp, S. Vol. 7006. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 335–346. DOI: [10.1007/978-3-642-24455-1_33](https://doi.org/10.1007/978-3-642-24455-1_33).
- Tomley, F. M. and Shirley, M. W. (2009). "Livestock infectious diseases and zoonoses". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, pp. 2637–2642.
- Tong, H., Prakash, B. A., Eliassi-Rad, T., Faloutsos, M., and Faloutsos, C. (2012). "Gelling, and melting, large graphs by edge manipulation". In: *Proceedings of the 21st ACM international conference on Information and knowledge management*. CIKM '12. New York, NY, USA: Association for Computing Machinery, pp. 245–254. ISBN: 978-1-4503-1156-4. DOI: [10.1145/2396761.2396795](https://doi.org/10.1145/2396761.2396795).
- Tuyls, K. and Nowé, A. (2005). "Evolutionary game theory and multi-agent reinforcement learning". en. In: *The Knowledge Engineering Review* 20.1. Publisher: Cambridge University Press, pp. 63–90. ISSN: 1469-8005, 0269-8889. DOI: [10.1017/S026988890500041X](https://doi.org/10.1017/S026988890500041X).
- Van Schepdael, A., Carlier, A., and Geris, L. (2016). "Sensitivity analysis by design of experiments". In: *Uncertainty in Biology*. Springer, pp. 327–366.

- Vega-Redondo, F. (2003). *Economics and the Theory of Games*. en. Cambridge University Press. ISBN: 978-0-521-77590-8.
- Venkatramanan, S., Chen, J., Fadikar, A., Gupta, S., Higdon, D., Lewis, B., Marathe, M., Mortveit, H., and Vullikanti, A. (2019). "Optimizing spatial allocation of seasonal influenza vaccine under temporal constraints". en. In: *PLOS Computational Biology* 15.9. Publisher: Public Library of Science, e1007111. ISSN: 1553-7358. DOI: [10.1371/journal.pcbi.1007111](https://doi.org/10.1371/journal.pcbi.1007111).
- Vergu, E., Busson, H., and Ezanno, P. (2010). "Impact of the Infection Period Distribution on the Epidemic Spread in a Metapopulation Model". en. In: *PLOS ONE* 5.2. Publisher: Public Library of Science, e9371. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0009371](https://doi.org/10.1371/journal.pone.0009371).
- Vernon, M. C. (2011). "Demographics of cattle movements in the United Kingdom". In: *BMC veterinary research* 7.1, pp. 1–16.
- Vernon, M. C. and Keeling, M. J. (2009). "Representing the UK's cattle herd as static and dynamic networks". In: *Proceedings of the Royal Society B: Biological Sciences* 276.1656, pp. 469–476.
- Viet, A.-F., Fourichon, C., Seegers, H., and Guihenneuc-Jouyau, C. (2004). "A model of the spread of the bovine viral-diarrhea virus within a dairy herd". In: *Preventive veterinary medicine* 63, pp. 211–36. DOI: [10.1016/j.prevetmed.2004.01.015](https://doi.org/10.1016/j.prevetmed.2004.01.015).
- Viet, A.-F., Krebs, S., Rat-Aspert, O., Jeanpierre, L., Belloc, C., and Ezanno, P. (2018). "A modelling framework based on MDP to coordinate farmers' disease control decisions at a regional scale". en. In: *PLOS ONE* 13.6. Ed. by Zia, A., e0197612. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0197612](https://doi.org/10.1371/journal.pone.0197612).
- Voinson, M., Billiard, S., and Alvergne, A. (2015). "Beyond Rational Decision-Making: Modelling the Influence of Cognitive Biases on the Dynamics of Vaccination Coverage". en. In: *PLOS ONE* 10.11. Ed. by Bauch, C. T., e0142990. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0142990](https://doi.org/10.1371/journal.pone.0142990).
- Volkova, V., Howey, R., Savill, N., and Woolhouse, M. (2010). "Potential for transmission of infections in networks of cattle farms". In: *Epidemics* 2.3, pp. 116–122.
- Wallace, C. and Young, H. P. (2015). "Chapter 6 - Stochastic Evolutionary Game Dynamics". en. In: *Handbook of Game Theory with Economic Applications*. Ed. by Young, H. P. and Zamir, S. Vol. 4. Elsevier, pp. 327–380. DOI: [10.1016/B978-0-444-53766-9.00006-9](https://doi.org/10.1016/B978-0-444-53766-9.00006-9).
- Wang, J., He, J., Yu, F., Guo, Y., Li, M., and Chen, W. (2020). "Realistic decision-making process with memory and adaptability in evolutionary vaccination game". en. In: *Chaos, Solitons & Fractals* 132, p. 109582. ISSN: 09600779. DOI: [10.1016/j.chaos.2019.109582](https://doi.org/10.1016/j.chaos.2019.109582).
- Wang, Y., Chakrabarti, D., Wang, C., and Faloutsos, C. (2003). "Epidemic spreading in real networks: an eigenvalue viewpoint". In: *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings*. ISSN: 1060-9857, pp. 25–34. DOI: [10.1109/RELDIS.2003.1238052](https://doi.org/10.1109/RELDIS.2003.1238052).
- Wang, Z., Bauch, C. T., Bhattacharyya, S., d'Onofrio, A., Manfredi, P., Perc, M., Perra, N., Salathe, M., and Zhao, D. (2016). "Statistical physics of vaccination". In: *Physics Reports* 664, pp. 1–113.
- Wang, Z., Moreno, Y., Boccaletti, S., and Perc, M. (2017). "Vaccination and epidemics in networked populations—An introduction". In: *Chaos, Solitons & Fractals* 103, pp. 177–183. ISSN: 0960-0779. DOI: <https://doi.org/10.1016/j.chaos.2017.06.004>.
- Wang, Z., Szolnoki, A., and Perc, M. (2012). "If players are sparse social dilemmas are too: Importance of percolation for evolution of cooperation". In: *Scientific Reports* 2, p. 369.
- Wijayanto, A. W. and Murata, T. (2019). "Effective and scalable methods for graph protection strategies against epidemics on dynamic networks". In: *Applied Network Science* 4.1, pp. 1–31.
- Xia, S. and Liu, J. (2014). "A belief-based model for characterizing the spread of awareness and its impacts on individuals' vaccination decisions". In: *Journal of The Royal Society Interface* 11.94. Publisher: Royal Society, p. 20140013. DOI: [10.1098/rsif.2014.0013](https://doi.org/10.1098/rsif.2014.0013).
- Xuan, P. and Lesser, V. (2002). "Multi-agent policies: from centralized ones to decentralized ones". In: *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 3. AAMAS '02*. New York, NY, USA: Association for Computing Machinery, pp. 1098–1105. ISBN: 978-1-58113-480-3. DOI: [10.1145/545056.545078](https://doi.org/10.1145/545056.545078).
- Zaric, G. and Brandeau, M. (2001). "Resource allocation for epidemic control over short time horizons". In: *Mathematical Biosciences* 171.1, pp. 33–58. DOI: [10.1016/S0025-5564\(01\)00050-5](https://doi.org/10.1016/S0025-5564(01)00050-5).
- Zhan, J., Rafalski, T., Stashkevich, G., and Verenich, E. (2017). "Vaccination allocation in large dynamic networks". In: *Journal of Big Data* 4.1, pp. 1–17.

- Zhang, H.-F., Wu, Z.-X., Xu, X.-K., Small, M., Wang, L., and Wang, B.-H. (2013). "Impacts of subsidy policies on vaccination decisions in contact networks". en. In: *Physical Review E* 88.1, p. 012813. ISSN: 1539-3755, 1550-2376. DOI: [10.1103/PhysRevE.88.012813](https://doi.org/10.1103/PhysRevE.88.012813).
- Zhang, H., Fu, F., Zhang, W., and Wang, B. (2012). "Rational behavior is a 'double-edged sword' when considering voluntary vaccination". en. In: *Physica A: Statistical Mechanics and its Applications* 391.20, pp. 4807–4815. ISSN: 0378-4371. DOI: [10.1016/j.physa.2012.05.009](https://doi.org/10.1016/j.physa.2012.05.009).
- Zhang, Y. and Prakash, B. A. (2015). "Data-aware vaccine allocation over large networks". In: *ACM Transactions on Knowledge Discovery from Data (TKDD)* 10.2, pp. 1–32.
- Zhao, H. and Feng, Z. (2019). "Identifying optimal vaccination strategies via economic and epidemiological modeling". en. In: *Journal of Biological Systems* 27.04, pp. 423–446. ISSN: 0218-3390, 1793-6470. DOI: [10.1142/S0218339019400011](https://doi.org/10.1142/S0218339019400011).
- Zimmer, G. M., Van Maanen, C., De Goey, I., Brinkhof, J., and Wentink, G. H. (2004). "The effect of maternal antibodies on the detection of bovine virus diarrhoea virus in peripheral blood samples". eng. In: *Veterinary Microbiology* 100.3-4, pp. 145–149. ISSN: 0378-1135. DOI: [10.1016/j.vetmic.2004.03.008](https://doi.org/10.1016/j.vetmic.2004.03.008).



OPEN

Accounting for farmers' control decisions in a model of pathogen spread through animal trade

Lina Cristancho Fajardo^{1,2✉}, Pauline Ezanno² & Elisabeta Vergu¹

Accounting for individual decisions in mechanistic epidemiological models remains a challenge, especially for unregulated endemic animal diseases for which control is not compulsory. We propose a new integrative model by combining two sub-models. The first one for the dynamics of a livestock epidemic on a metapopulation network, grounded on demographic and animal trade data. The second one for farmers' behavior regarding the adoption of a control measure against the disease spread in their herd. The measure is specified as a protective vaccine with given economic implications, and the model is numerically studied through intensive simulations and sensitivity analyses. While each tested parameter of the model has an impact on the overall model behavior, the most important factor in farmers' decisions is their frequency, as this factor explained almost 30% of the variation in decision-related outputs of the model. Indeed, updating frequently local health information impacts positively vaccination, and limits strongly the propagation of the pathogen. Our study is relevant for the understanding of the interplay between decision-related human behavior and livestock epidemic dynamics. The model can be used for other structures of epidemic models or different interventions, by adapting its components.

Fighting livestock diseases spreading through animal trade is a major issue to guarantee sustainable farming, competitive agrifood chains and public health¹. Epidemic prevention and reduction of prevalence require improved methods of control and compliance of the actors, especially for non-regulated diseases for which control decisions are left to individual or collective initiatives². Mechanistic epidemiological models can provide a refined mathematical description and understanding of the complex system involved in pathogen spread, and be used to assess the effectiveness of control measures. They are complementary to observational or experimental approaches³. However, accounting for human behavior in such models in order to increase their predictive power remains a challenge^{4,5}, in particular for livestock diseases spreading through a trade network.

Indeed, most works on infectious diseases that consider the adoption of control measures usually do not account for human decision-making⁶ or they do not consider a structured population^{7,8}. In particular, in models based on a metapopulation over an explicit network⁹ such as¹⁰, control decisions are usually assumed to be taken at a centralized level^{11,12}. Furthermore, when human decision-making is explicitly taken into account, it generally focuses on the context of human diseases^{13–15}, but it has barely been applied to veterinary epidemiology yet¹⁶. However, in the context of animal diseases, the decision of implementing control measures can be much more influenced by economic considerations than for human diseases, an aspect that should be taken into account in the decision model. Finally, in the field of veterinary epidemiology, studies have been mostly focused on regulated diseases, so human behavior mainly consists in delaying the application of a central policy¹⁷. In the few works that investigate control measures for unregulated animal diseases, there is generally no dynamic feedback on decision due to epidemic evolution¹⁸. Additionally, some real-life aspects, such as farmers having limited rationality, free-riding or learning are generally missing^{19,20}. There is therefore a special need for models combining the dynamics of an epidemic process that takes place on a livestock trade network, and the behavior of farmers regarding the voluntary implementation of control measures^{21,22}.

We build an integrative model that can meet this need by coupling the dynamic spread of a livestock disease over a structured metapopulation, and the dynamics of the human decision-making process for applying a sanitary measure against the epidemic spread. To model the epidemic spread through a trade network we use a stochastic compartmental model that takes into account demographic dynamics and animal exchanges. The population structure of the model is calibrated using real data. Our decision model is inspired by previous studies^{13–15}, in which the result of a decision regarding the voluntary adoption of a control measure for a human

¹Université Paris-Saclay, INRAE, MalAGE, 78350 Jouy-en-Josas, France. ²INRAE, Oniris, BIOEPAR, 44307 Nantes, France. ✉email: lina.cristancho-fajardo@inrae.fr

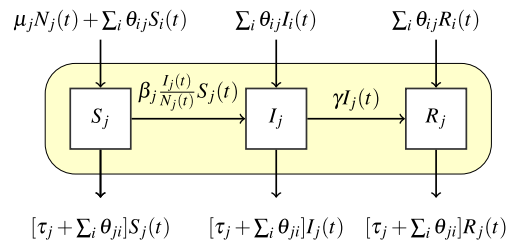


Figure 1. Schematic representation of the intra-herd epidemic-demographic dynamics for a herd j , without any control measure. Horizontal arrows represent transitions between health-related compartments, corresponding to the course of infection inside the herd (yellow rectangle), while vertical arrows represent population flows to and from the herd. The coefficients on the arrows are the transition rates. See main text in “Methods” for parameter definitions.

disease is evaluated after being applied, and preferences over the possible decisions are updated through time. We specifically formalize the dynamic decision problem that each farmer faces, and propose a mechanism that represents farmers’ decision-making process in such a context. Our formalization considers some real-life phenomena that can be present in the context of human decision-making: stochastic behavior, learning, and the emergence of imitation and free-riding²³.

This paper is structured as follows. First, “Methods” section presents the two components of the integrative model: the epidemic–demographic model, and the decision model. Then, we describe the specific control measure we consider, as well as its economic implications. We emphasize that the main contribution of our work is methodological, so this integrative model is a result in itself. At the end of this section, we describe the setting for simulations and sensitivity analyses we perform on the model, whose findings can be found in “Results” section. Finally, the model as well as the results of numerical explorations are examined in “Discussion” section.

Methods

In the following, we describe the two main components of our integrative model: the epidemic–demographic one, and the decision-making one. We then detail the integrative model by considering vaccination as a specific control measure. Finally, we describe the methodology used for the simulation and analysis of the model.

Epidemic model with demography in a metapopulation based on a trade network. For this work we place ourselves in the context of a hypothetical livestock infectious disease that is transmitted only through a contact network structure consisting in herds that exchange animals. This population structure is inspired by real data on animal movements, extracted from the French Cattle Identification Database (FCID). We assume this is a closed metapopulation, that is, we neglect exchanges with herds outside of it. This livestock trade network can be described as a directed weighted time-varying network, where nodes represent herds and links represent animal trade. The direction of each link is determined by the transfer’s direction, and its weight corresponds to the amount of animals exchanged. By nature, this network is time-varying since links may change over time. In fact, not only trade connections may appear or disappear, but the amount of animals exchanged can vary on a daily basis.

Given this trade network, we consider an infectious livestock disease that can potentially be spread on it, and that can only be directly transmitted between animals within the same herd. The disease is assumed to be spread between herds only by animal transfers, as can be observed for diseases such as paratuberculosis²⁴, bovine tuberculosis when there is no contact with wildlife²⁵, and porcine reproductive and respiratory syndrome virus²⁶. In addition, the infection risk and status are assumed independent of animal breed, age or sex. In the absence of any intervention, the intra-herd disease spread is described by a stochastic SIR model³ with demography, accounting for animal transfers over the trade network. In a compartmental SIR model, the population is divided into three compartments: Susceptible (S), Infected (I), and Recovered (R), according to their health status. The only two possible transitions in a basic SIR model correspond to infection ($S \rightarrow I$) and recovery ($I \rightarrow R$). The implicit modelling assumptions we make are the following: intra-herd homogeneous mixing, meaning that the contact rate is the same among all the animals in a given herd; absence of a latent period, i.e. animals become infectious as soon as they are infected; acquisition of immunity after recovery; no vertical transmission, i.e. no mother-to-child transmission during pregnancy or childbirth; frequency-dependent intra-herd transmission, i.e. the transmission rate depends on the proportion of infected animals in the herd, rather than on their number; variation in time of herd size due to births, deaths and animal transfers, which we assume are not affected by the disease prevalence.

Formally, we consider J herds in the population. Without any intervention, for each herd $j = 1, \dots, J$ the intra-herd transmission of the disease can be described by the scheme in Fig. 1. We note $S_j(t)$, $I_j(t)$ and $R_j(t)$ the number of susceptible, infected and recovered animals in herd j at time t . We suppose $S_j(0) > 0$ for all j , $I_j(0) > 0$ for at least one herd j , and $R_j(0) = 0$ for all j . We note as $N_j(t) := S_j(t) + I_j(t) + R_j(t)$ the size of herd j at time t . The parameters β_j , τ_j and μ_j are the herd specific daily rates of disease transmission, death and birth in herd j . As for γ , it is the recovery rate from the disease. Finally, θ_{ij} is the daily out rate from herd j to herd i , representing the mean daily proportion of animals of herd j going to herd i . We consider a continuous-time Markov chain model for the stochastic epidemic–demographic dynamics of each herd which we simulate through an Euler

discrete-time scheme using multinomial distributions for numerical efficiency (as described in²⁷ but with non random rates). Transition probabilities between compartments, corresponding to birth, death, infection, recovery and transfer events, are defined by $p_{XY} = [1 - \exp(-\sum_{X \neq Y} \eta_{XY})] \eta_{XY} / \sum_{X \neq Y} \eta_{XY}$, where η_{XY} is the daily rate concerning the transition from a compartment X to a compartment Y . See Supplementary Equations S1–S15 for further details.

Farmer’s decision-making model. We suppose that farmers can apply a sanitary measure that has a certain efficacy on the disease spread for a limited amount of time. Then, we assume they search to take the control decision that allows them to obtain an optimal value of an individual criterion, i.e. an expected cost resulting from the decision. To consider a simple and clear framework, we restrict ourselves to binary decisions (the measure is applied or not). Additionally, we make the assumption that decision times are synchronized, discrete, and equally spaced in time. This proves to be useful when considering the interaction of farmers’ decisions, and for evaluating the effect of the time length between successive decisions on the behavior of the integrative model. Formally, we suppose that each farmer $j = 1, \dots, J$ searches to solve:

$$\min_{a_j^t} \mathbb{E} \left[C_{a_j^t}^t(j) \right]; \quad t = \Delta_d, 2\Delta_d, 3\Delta_d \dots \tag{1}$$

where $t = \Delta_d, 2\Delta_d, 3\Delta_d \dots$ are the decision times, so Δ_d is the duration (in days) between any two consecutive decisions. It also determines the instant at which the first decisions are taken. The term $a_j^t \in \{0, 1\}$ refers to the control decision: if $a_j^t = 1$, the control-measure is applied in herd j at decision time t , otherwise it is not. $C_{a_j^t}^t(j)$ is the cost in herd j associated with the decision taken at time t . This constitutes a dynamic decision-making problem under uncertainty, this latter affecting the cost distribution associated with each possible decision.

To define the farmer’s decision-making process that attempts to solve this problem, we take an approach inspired by^{13–15}, in which farmers evaluate the result of a decision after its application, and update their preferences over time as a function of this result. In this particular context, this approach seems suitable for several reasons. First, as we mentioned earlier, costs over time not only depend on the epidemic and decision dynamics in the herd where the decision is made, but also on other herds. To exactly solve this optimization problem would imply that farmers integrate the actions and epidemic status of other herds. This is a very complex problem due to the dimensionality on the number of herds, and on the possible status of the system. Second, since we pose a dynamic decision-making problem, there is an effect of learning through repeated decisions. Indeed, we have supposed that the cost associated to a decision is observed before making the next decision. Then, it is natural to think that farmers learn from the costs they have obtained with their previous choices, to take their next decisions. Finally, through this approach we can easily consider social dynamics such as imitation effects between farmers. In our context, this consists in considering a stochastic decision mechanism where the probability of applying the measure is updated through the costs each farmer observes over time, and the costs observed by his/her neighbors.

Algorithm 1 Exponential weighting stochastic mechanism with imitation	
Input:	2 options = $\{0, 1\}$, $p_1^{\Delta_d}(j) := p_1^{init} \forall j$, $\kappa \geq 0$, $\rho \geq 0$, $B(j) = \{i; \theta_{ij} \neq 0 \text{ or } \theta_{ji} \neq 0\}; j = 1, \dots, J$.
For:	$t = \Delta_d, 2\Delta_d, 3\Delta_d \dots$ (at each decision time):
	For: $j = 1, \dots, J$ (each farmer):
	- $a_j^t \leftarrow \text{Bernoulli}(p_1^t(j))$ (takes a decision using his/her current probability of applying the measure)
	- $C_{a_j^t}^t(j)$ (observes the cost related to his/her decision)
	- $j^* \leftarrow \text{Unif}(B(j))$ (selects one of his/her neighbors in the trade network)
	- $(a_{j^*}^t, C_{a_{j^*}^t}^t(j^*))$ (observes the decision taken by j^* and his/her observed cost)
	- (updates the probability of applying the measure):
	$P_1^{t+\Delta_d}(j) = \frac{p_1^t(j) e^{-\kappa C_1^t(j) - \rho C_1^t(j^*)}}{p_1^t(j) e^{-\kappa C_1^t(j) - \rho C_1^t(j^*)} + (1 - p_1^t(j)) e^{-\kappa C_0^t(j) - \rho C_0^t(j^*)}} \tag{2}$
	where the costs of the non taken options are equal to 0, i.e. for $k = 0, 1$:
	* $C_k^t(j) = C_{a_j^t}^t(j)$ if $k = a_j^t$, 0 otherwise.
	* $C_k^t(j^*) = C_{a_{j^*}^t}^t(j^*)$ if $k = a_{j^*}^t$, 0 otherwise.

The mechanism we propose (Algorithm 1) works by updating the probability of applying the measure, proportionally to an exponential weight that takes into account the last decision taken by the farmer and that taken by one of his/her neighbors, through a weighted sum of the associated costs. Then at each decision time, each farmer $j = 1, \dots, J$ takes a decision a_j^t using his/her current probability of applying the measure $p_1^t(j)$. We assume that this probability is initially the same for all farmers, and equal to a value p_1^{init} , and that each farmer observes the cost related to his/her decision, and the decision and associated cost observed by one of his/her neighbors in the trade network, who is randomly chosen by the farmer. A neighbor of j in the trade network is a farmer with whom j exchanges animals according to the daily trade rates, i.e. a farmer j^* such that $\theta_{jj^*} \neq 0$, or $\theta_{j^*j} \neq 0$. In the algorithm, we note as $B(j)$ the set of neighbors of j in the trade network.

The update in the probability is then given by Eq. (2). The parameter κ represents farmer’s “sensitivity” to his/her own observed costs. A κ close to zero implies that farmers are not very sensitive to their own observed costs, and therefore mostly rely on their initial probability of applying the measure, whereas a large κ represents

the situation in which farmers are very sensitive to their own observed costs for updating their probability of applying the measure. For considering an imitation effect, we introduce the parameter ρ that works analogously to κ , but on the cost observed by the chosen neighbor. The parameters κ and ρ act then as weights to the farmer's and the neighbor's observed cost, respectively. In our model, farmer's next decision can be updated considering any of his/her neighbors, regardless of what the neighbor has decided in the previous step. Finally, for updating the probabilities, it is natural that these are set so that the decision with a smaller sum of weighted costs receives higher probability. Although there are many ways to turn the sum of weighted costs into probabilities, a simple and popular method is to use an exponential weighting scheme. This scheme quickly reduces the probability of the decision that has resulted to be very bad (high sum of weighted costs). This form is found in the Pairwise Fermi (PW-Fermi) rule, which has been previously used in similar contexts, as its stochastic behavior is similar to real-life human decision-making²⁸. In this update, the cost associated to non-taken decisions are zero, i.e. either $C_1^t(j)$ or $C_0^t(j)$ is zero, and either $C_1^t(j^*)$ or $C_0^t(j^*)$ is zero. The non-zero costs define the final form of the probability update. In order to see the effect of the decisions and the associated observed costs in this update, we remark that since we consider binary decisions, Eq. (2) can be rewritten as an update on the odds of applying the measure:

$$odds_1^{t+\Delta_d}(j) = odds_1^t(j) \times e^{(1-2a_j^t)\kappa C_{a_j^t}^t(j) + (1-2a_{j^*}^t)\rho C_{a_{j^*}^t}^t(j^*)} \tag{3}$$

where $odds_1^t(j) := p_1^t(j)/(1 - p_1^t(j))$; $\forall t = \Delta_d, 2\Delta_d, \dots$. From this we can see that the odds are reinforced or decreased as a result of the farmer's and the neighbor's decision and cost. If they both apply the measure at time t , the term in the exponential is negative since costs are positive or zero, so the odds for j applying the measure decrease. Analogously, if neither of them applies the measure at time t , the term in the exponential is positive and the odds of applying it increase. Finally, if they do not make the same decision at time t it is the comparison between $\kappa C_{a_j^t}^t(j)$ and $\rho C_{a_{j^*}^t}^t(j^*)$ that determines the direction of the update.

Additionally, we explore an extension of the model where each farmer considers the decisions and costs observed by all of his/her neighbors at each decision time. To update his/her probability of vaccinating, he/she takes into account the costs observed by his/her neighbors who did not vaccinate at the previous decision time, and the costs observed by those who vaccinated, as described in Supplementary Algorithm S1.

An epidemic control measure. For the control measure that can be applied to manage the spread of the disease, we specifically consider a vaccine that can reduce the rate of disease transmission towards a susceptible vaccinated animal. We assume this is the only effect the vaccine has. We make the assumption that the vaccine maintains a constant efficacy during a certain time period, whose duration is the same as the decision time-step. Then, if the vaccine is applied on a susceptible animal in herd j at time t , the rate of transmission towards that susceptible animal during the period $[t; t + \Delta_d]$ will be $\beta_j^v = \beta_j(1 - e_v)$, where $0 \leq e_v \leq 1$ is the protection efficacy of the vaccine.

An economic–epidemiological cost function. We assume that the farmers are able to assess the economic impact that their decisions have on the disease spread in their herd. Therefore, we define the costs on the basis of a simple economic cost function, related to the epidemiological consequences of the decision taken at t in herd j . We define it in particular for the considered control measure, a protective vaccine, but it can easily be modified for a control measure with a different impact on the epidemic transition rates. The cost function we considered is:

$$C_{a_j^t}^t(j) := \frac{[CF_v + CU_v N_j(t)] a_j^t + \phi r N_{S_j \rightarrow I_j}(t, t + \Delta_d)}{N_j(t, t + \Delta_d)} \tag{4}$$

where in the numerator the first term refers to the cost farmers pay to apply the vaccine, and the second one to the economic impact of the epidemic consequences of the vaccine. Precisely, in the first term a_j^t equals 1 if the vaccine is applied on herd j at decision time t , and it equals 0 otherwise. CU_v is the unitary cost of the vaccine per animal, and CF_v defines a fixed cost of applying vaccination per herd. This would typically correspond to the cost of a veterinary visit. In the second term, r is the monetary value of a healthy animal, and $0 \leq \phi \leq 1$ is the rate of reduction of this value if the animal gets infected. So ϕr is the cost of an infection, that is, the loss in the monetary value of an animal if it gets infected. $N_{S_j \rightarrow I_j}(t, t + \Delta_d)$ is the number of new infections in the herd, from the moment decision is taken until the next decision time. Therefore, the benefit of having healthy animals is implicitly given by the animal not reducing its value due to an infection. We remark that we make the assumption that each farmer perfectly observes the number of new infections that occurred during the decision period, or at least the global loss in the monetary value of the herd $\phi r N_{S_j \rightarrow I_j}(t, t + \Delta_d)$ related to these new infections. However, farmers can not identify which animals are infected, which is why we assume they choose to vaccinate the whole herd if vaccination is decided. Finally, in order to account for differences in the costs that may only be related to the variation of the herd size over the period, we standardize the cost by the sum of the daily herd size during the concerned period, which we note as $N_j(t, t + \Delta_d) = N_j(t) + N_j(t + 1) + N_j(t + 2) + \dots + N_j(t + \Delta_d)$. This is equivalent to standardizing by $\Delta_d \bar{N}_j(t, t + \Delta_d)$, where $\bar{N}_j(t, t + \Delta_d)$ is the mean daily herd size over the period.

The scheme of the integrative model for vaccination can be found in Fig. 2. It shows the feed-back loop between the epidemic–demographic dynamics, and the decision dynamics. The epidemic–demographic process takes place for a period of length Δ_d , until a new decision is taken. This decision is itself a function of economic and epidemic consequences of the previous decision.

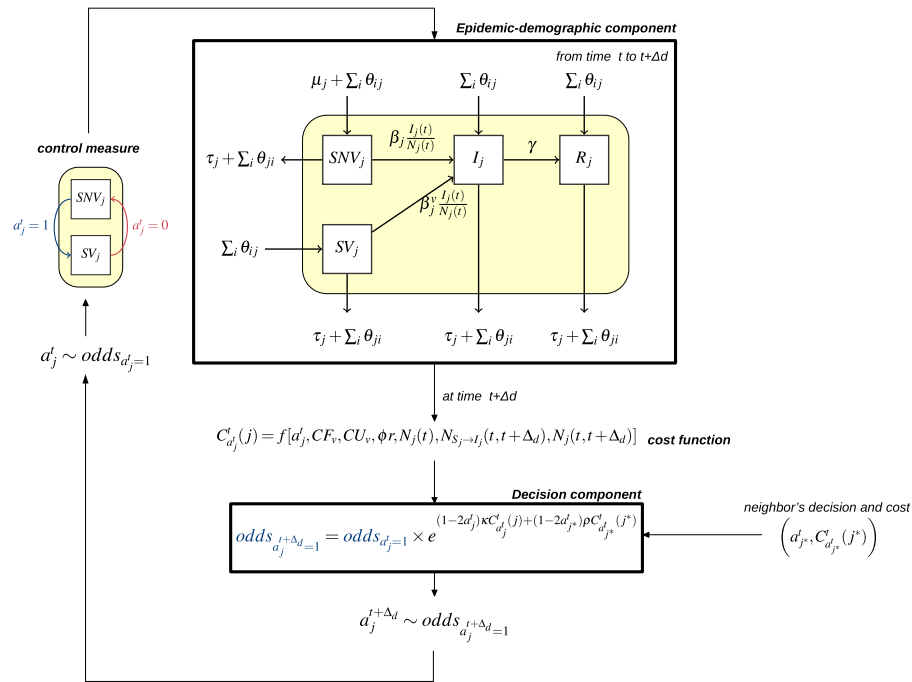


Figure 2. Representation of the integrative epidemic-decision dynamical model for a herd j , accounting for vaccinating decisions with a protective effect ($\beta_j^v < \beta_j$). See main text in “Methods” for parameter definitions.

The flow between the non vaccinated compartment, $SNV_j(t)$, and the vaccinated one, $SV_j(t)$, is deterministic once the decision is taken. Indeed, if $a_j^t = 1$, all susceptible animals in herd j will enter the $SV_j(t)$ compartment for the next decision period. If $a_j^t = 0$, they will be in $SNV_j(t)$. If decided, vaccination is then applied only once per decision time. Indeed, in livestock diseases, as opposed to what happens for human diseases, it is not customary that farmers vaccinate newborns or the animals they buy after they have already vaccinated the herd, since each vaccination would then imply a cost for a veterinary visit. Furthermore, if the herd is vaccinated, farmers generally rely on herd immunity to indirectly protect susceptible animals in the herd.

Simulation setting. For our simulation study, the population structure is set close to the one observed in real data obtained from the FCID. Furthermore, demographic parameters are fixed close to real-life values. Details on the values used for these parameters and the procedures to generate population structure are specified in the Supplementary Methods. In particular, the simulated trade network is scale-free, as the one observed in the real-life animal movements, then the in-degree and out-degree distributions follow a power law. That is, the majority of herds only buy (sell) animals to a few other herds, and very few herds buy (sell) to many different herds, which are known as hubs. We simulate this network through the configuration model, using degree sequences generated from a power law. We consider $J = 5000$ herds, roughly the number of herds in the Finistère region in France, a cattle densely populated region, which we follow during $T = 1095$ days (i.e. 3 years). The values of the epidemic, economic and decision-related parameters used in the simulation study are given in Table 1.

We remark that these are set close to realistic values, having in mind a standard SIR endemic disease. In particular, we consider the same transmission rate across herds, so $\beta_j = \beta; \forall j = 1, \dots, J$. As for the duration of the decision it is chosen to be 180 days, which is a reasonable assumption in practice. The values for κ and ρ are chosen so as to have two potentially contrasted decision scenarios.

Sensitivity analyses. Sensitivity analysis is useful to study how much the variation in each parameter of the model contributes to the variation of the model outputs²⁹. In our sensitivity analyses we consider 13 input parameters in total. Other parameters, in particular the demographic ones, are fixed as specified earlier. We consider eight outputs corresponding to the three model components: epidemic, economic and decision-related, and one additional output that combines epidemic and decision-related elements. These outputs are defined in Table 2.

The values of the inputs used in the sensitivity analyses are chosen using Fractional Factorial design³⁰ with 5 equally spaced levels, which results in 625 combinations of parameters. To obtain this design we use the R package PLANOR³¹. Since the model is stochastic, we run 50 simulations for each combination, and we consider the mean and the variance of each output over runs. Table 1 contains the values considered for each input in the full sensitivity analysis. Since we use a IV-resolution design, we are able to estimate main effects unconfounded by two-factor interactions, while limiting the number of runs required for the analysis. With this design, we can also estimate two-factor interaction effects, even if these may be confounded, i.e. can not be estimated independently

	Parameters	Definition	Standard value	Values tested in the sensitivity analysis
Epidemic	β/γ	Transmission rate per herd \times average duration of infection	2	[1.1, 2.07, 3.05, 4.02, 5]
	$1/\gamma$	Average duration of infection (in days)	90	[10, 32.5, 55, 77.5, 100]
	p_{herds}^0	Initial proportion of infected herds	0.10	[0.01, 0.22, 0.43, 0.64, 0.85]
	p_{anim}^0	Initial proportion of infected animals in infected herds	0.15	[0.01, 0.25, 0.50, 0.75, 1]
Economic	r	Monetary value of a healthy animal (in euros)	2000	[1000, 1500, 2000, 2500, 3000]
	ϕ	Reduction in the monetary value of an animal if it gets infected	0.8	[0.01, 0.25, 0.50, 0.75, 1]
	CU_v	Unitary cost of the vaccine per animal (in euros)	5	[1, 4.5, 8, 11.5, 15]
	CF_v	Fixed cost of applying vaccination per herd (in euros)	50	[1, 25.75, 50.5, 75.25, 100]
Decision-related	e_v	Protection efficacy of the vaccine on susceptible animals	1	[0.01, 0.25, 0.50, 0.75, 1]
	Δ_d	Duration of the decision (time between two consecutive decisions). It also determines the time of the first decision, and is equal to the duration efficacy of the vaccine (in days)	180	[30, 114, 198, 281, 365]
	p_v^{init}	Farmers' initial probability of vaccinating	0.01	[0.01, 0.25, 0.5, 0.74, 0.99]
	κ	Farmers' sensitivity to their own observed cost	0.5. or 12.5	[0.5, 3.5, 6.5, 9.5, 12.5]
	ρ/κ	Farmers' sensitivity to a neighbor's cost over farmers' sensitivity to his/her own observed cost	0.5	[0, 0.25, 0.5, 0.75, 1]

Table 1. Parameters of the integrative model: description, standard values and values tested in the full sensitivity analysis.

Group	Output	Definition
Epidemic	p_{herds}^T (final inter-herd prevalence rate)	Final proportion of infected herds: $\frac{1}{J} \sum_{j=1}^J \mathbb{1}_{I_j(T)>0}$
	\bar{p}_{anim}^T (final intra-herd mean prevalence rate)	Mean over final infected herds of the final proportion of infected animals: $\left(\sum_{j=1}^J \frac{I_j(T)}{N_j(T)} \right) / \left(\sum_{j=1}^J \mathbb{1}_{I_j(T)>0} \right)$
	$p_{herds}^{[0,T]}$ (inter-herd cumulative incidence rate)	Cumulative proportion of newly infected herds (i.e. herds with new infections): $\frac{1}{J} \sum_{j=1}^J \mathbb{1}_{\sum_{t=0}^T N_{S_j \rightarrow I_j(t)} > 0}$
	$\bar{p}_{anim}^{[0,T]}$ (mean cumulative intra-herd incidence rate)	Mean cumulative proportion of new infected animals over susceptible animals, for newly infected herds: $\left(\sum_{j=1}^J \frac{1}{T} \sum_{t=0}^T \frac{N_{S_j \rightarrow I_j(t)}}{S_j(t)} \right) / \left(\sum_{j=1}^J \mathbb{1}_{\sum_{t=0}^T N_{S_j \rightarrow I_j(t)} > 0} \right)$
	$\bar{I}_{anim}^{[0,T]}$ (mean cumulative intra-herd incidence)	Mean cumulative number of new infected animals for new infected herds: $\left(\sum_{j=1}^J \sum_{t=0}^T N_{S_j \rightarrow I_j(t)} \right) / \left(\sum_{j=1}^J \mathbb{1}_{\sum_{t=0}^T N_{S_j \rightarrow I_j(t)} > 0} \right)$
Economic	$C^{[0,T]}$ (total economic cost of the disease)	Sum of the non standardized cumulative disease-related costs (costs of vaccination and costs of new infections): $\sum_{j=1}^J \left[\sum_{n=0}^{\lfloor T/\Delta_d \rfloor} C_{a_j}^{n\Delta_d}(j) \Delta_d \bar{N}_j(t, t + \Delta_d) \right]$. Counts costs even before the first decision and after the last one
Decision-related	$\bar{p}_v^{[0,T]}$ (mean vaccination proportion)	Mean proportion of herds that vaccinate over the different decision times except the first one: $\left(\sum_{n=2}^{\lfloor T/\Delta_d \rfloor} \frac{1}{J} \sum_{j=1}^J \mathbb{1}_{a_j^{n\Delta_d}=1} \right) / (\lfloor T/\Delta_d \rfloor - 1)$
	Aggregated vaccination patterns	Vector consisting in three proportions: of herds that never vaccinate, of herds that vaccinate at least once and at most half of the time, and of herds that vaccinate more than half of the time but not always. Without taking the first decision into account
Epidemic-decision related	Mean cumulative intra-herd incidence rate by aggregated vaccination pattern	Vector of the mean cumulative intra-herd incidence rate (see output $\bar{p}_{anim}^{[0,T]}$) of herds grouped by the aggregated vaccination pattern: herds that never vaccinate, herds that vaccinate at least once and at most half of the time, and herds that vaccinate more than half of the time but not always. Without taking the first decision into account

Table 2. Description of the outputs of the sensitivity analyses.

to each other³². We study the outputs individually, by groups regarding the nature of the outputs, and by considering all outputs together. For the multivariate analyses, we use PCA (Principal Component Analysis) to reduce the dimension of the output space, before using Analysis of variance (ANOVA) for the computation of Global Sensitivity Indices (GSI), which are weighted means of the sensitivity indices over the retained dimensions in the PCA, as described in³³. More precisely, ANOVA is particularly suited for analyzing the outcome of a factorial

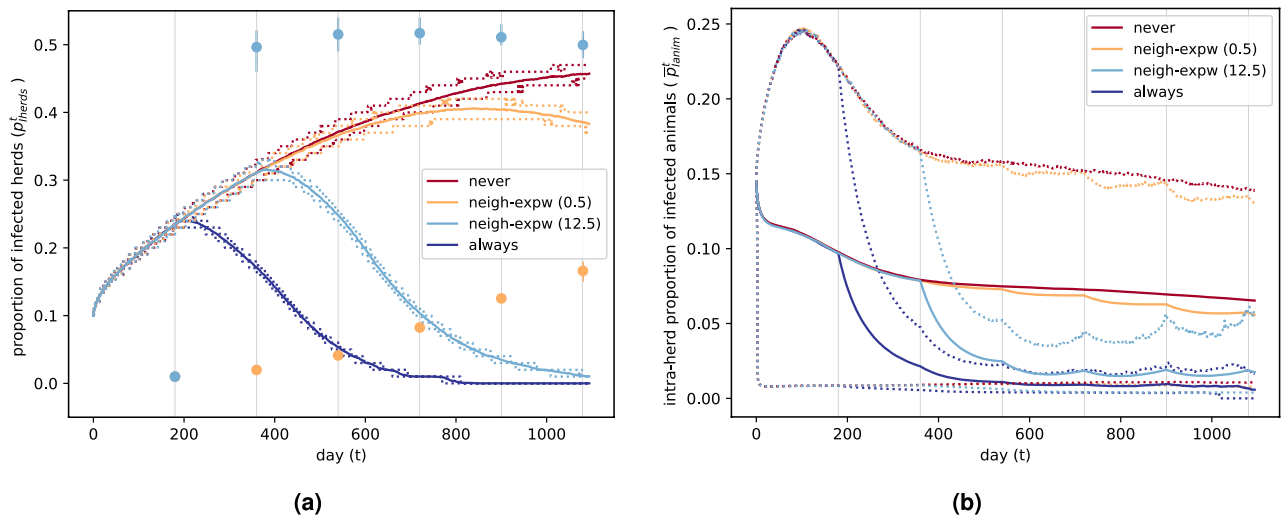


Figure 3. Temporal dynamics of the epidemic spread for each vaccination scenario over 50 runs. Each decision instant is represented by a vertical grey line. **(a)** Inter-herd prevalence. Mean over runs (solid lines), 10th and 90th percentiles over runs (dotted lines). Mean proportion of herds that vaccinate at each decision-time in each *neigh-expw* scenario (light blue and orange dots), and its variation over runs (from the 10th to the 90th percentile in light blue and orange vertical lines). **(b)** Intra-herd prevalence for infected herds. Mean over runs of the means over infected herds (solid lines), 10th percentile over runs of the 10th percentiles over infected herds, and 90th percentile over runs of the 90th percentiles over infected herds (dotted lines).

design³⁴. For all the sensitivity analyses we use the R package *multisensi*³⁵. In the PCAs the means are centered and scaled, and the dimension is selected as the smallest value that keeps at least 95% of the total variability. Among the many experiments, we retain the results of the three following ones:

- (i) First experiment: all 13 inputs. The means and variances of all outputs: by group, and all outputs simultaneously.
- (ii) Second experiment: all inputs except the four epidemic parameters (fixed to their standard values in Table 1). Means and variances of all outputs.
- (iii) Third experiment: all inputs except the two epidemic parameters (p_{herds}^0 and β/γ) and the two decision-related parameters (Δ_d and p_v^{init}), fixed to their standard values in Table 1. Means and the variances of decision-related outputs.

Results

Model predictions for different decision scenarios. Results regarding the inter-herd prevalence, and the intra-herd prevalence for infected herds are provided for four different scenarios (Fig. 3): no farmer ever vaccinates, *never* scenario; every farmer vaccinates at every decision-time, *always* scenario; farmers vaccinate following the proposed decision-making mechanism (Algorithm 1) using $\kappa = 0.5$, *neigh-expw(0.5)* scenario; and the same mechanism using $\kappa = 12.5$, *neigh-expw(12.5)* scenario.

As expected, the worst and best case scenarios are the scenario where farmers never vaccinate, and the one where they all vaccinate at each decision time. We remark that the vaccination gain particularly affects inter-herd prevalence, but is still observable for intra-herd prevalence. In the intermediate scenarios, farmers' sensitivity to observed costs determines the changes in the proportion of herds that vaccinate over time, and therefore in the control of the pathogen spread. Indeed, in the scenario with higher farmers' sensitivity to costs (*neigh-expw(12.5)* scenario), the proportion of farmers that vaccinate quickly increases after the first decision, generating a mean inter-herd and intra-herd prevalence dynamics rather close to the best case scenario. On the contrary, the scenario with smaller farmers' sensitivity to costs (*neigh-expw(0.5)* scenario) exhibits a slow increase in the proportion of herds that vaccinate, which gives rise to a prevalence behavior close to the one observed for the worst case scenario, even if around 2 years it starts to decline. Model predictions over a longer time horizon (9 years) can be found in Supplementary Fig. S8. The scenarios concerning vaccination exhibit some peaks in the intra-herd prevalence roughly at each decision time. For intra-herd prevalence dynamics this behavior is firstly explained by the fact that we consider this prevalence only for infected herds at each time, so the concerned herds are not the same over the whole trajectory. Furthermore, since we consider a perfect vaccine, when a herd is vaccinated all its susceptible animals are completely protected, so that the number of animals that can actually get infected drops instantaneously to zero, until there are births or imports of non-vaccinated susceptible animals (see Supplementary Figs. S5–S7 for an exploration of this behavior). The dynamics of the total number of infected animals (Supplementary Fig. S8(c)) is an alternative quantity to study. Yet, as evidenced by the figure, it is highly correlated to the proportion of infected herds.

Additionally, Fig. 4 presents the temporal dynamics of the vaccination decisions of the two intermediate scenarios (for a single run as an example). In the *neigh-expw(0.5)* scenario most herds never vaccinate (67%).

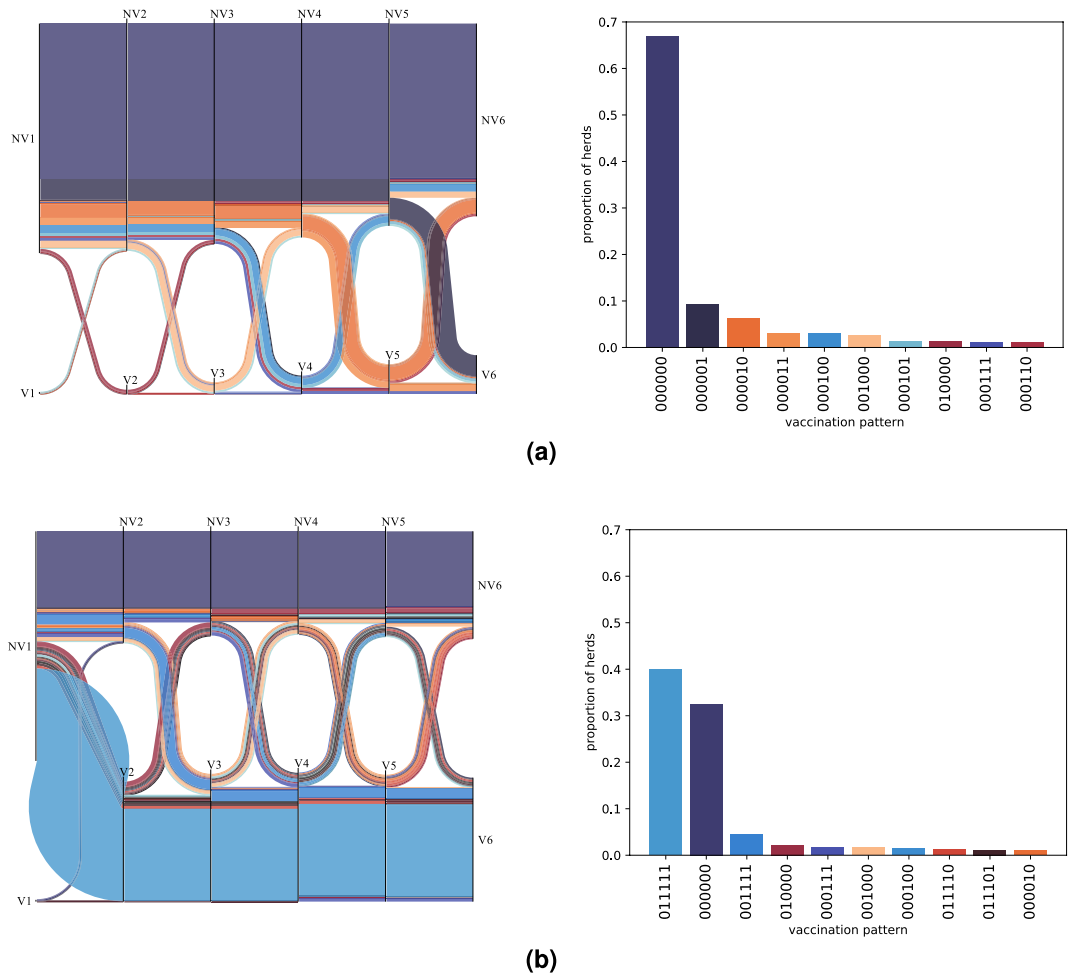


Figure 4. Temporal dynamics of the vaccination decisions using the decision mechanism defined in Algorithm 1 with $\kappa = 0.5$ (a), and $\kappa = 12.5$ (b). Results for one run. NV and 0 stand for not vaccinating, while V and 1 for vaccinating. Each color represents a different vaccination pattern, defined by the sequence of vaccination decisions at each of the six decision times. So the pattern 001111 (or equivalently [NV1, NV2, V3, V4, V5, V6]) concerns herds that do not vaccinate at the two first decision times, and always vaccinate afterwards. In the left plots, each vertical black line represents a decision time, and the width of the flows between decisions is proportional to the frequency of the pattern. In the right plots, the histogram of the patterns with a frequency $\geq 1\%$ is plotted. Hence, in (a), 67% of herds never vaccinate (pattern 000000). In (b), 39% of farms always vaccinate except in the first instant (pattern 011111), and 33% never vaccinate (pattern 000000).

They are followed by herds that only vaccinate at the last decision time, which are in turn followed by those that only vaccinate at the next to last decision time, etc. Only 28 out of the 64 possible patterns (over 6 decision times) are observed in this scenario. On the other hand, in the *neigh-expw(12.5)* scenario the most frequent behavior (39%) is to not vaccinate at the initial decision and to always vaccinate afterwards. However, this vaccination pattern is closely followed by the one where herds never vaccinate (33%). We also observe a higher variety of behaviors than in the *neigh-expw(0.5)* scenario, 44 out of the 64 possible patterns, which translates into less frequent patterns. Nevertheless, some of them stand out: the one where herds vaccinate from the third decision time, the one where herds only vaccinate at the second decision time, and the one where herds vaccinate from the fourth decision time.

Results concerning the alternative decision rule, where the information on the costs related to decisions is available for all trade neighbors, slightly differ (Supplementary Figs. S9–S10). For both scenarios where $\kappa = 0.5$, and $\kappa = 12.5$, there are slightly less and hence more frequent vaccination patterns with respect to the scenarios with the same parameter values but considering only one neighbor. For $\kappa = 0.5$, the proportion of herds that vaccinate increases and stabilizes more rapidly to a smaller value. The highest proportion of infected herds is slightly smaller, but afterwards it decreases less rapidly. For $\kappa = 12.5$, the proportion of herds that vaccinate increases more rapidly at the beginning and then it continues to decrease. The prevalence of the disease decreases only slightly faster than when using only one neighbor, the epidemic behavior being almost the same.

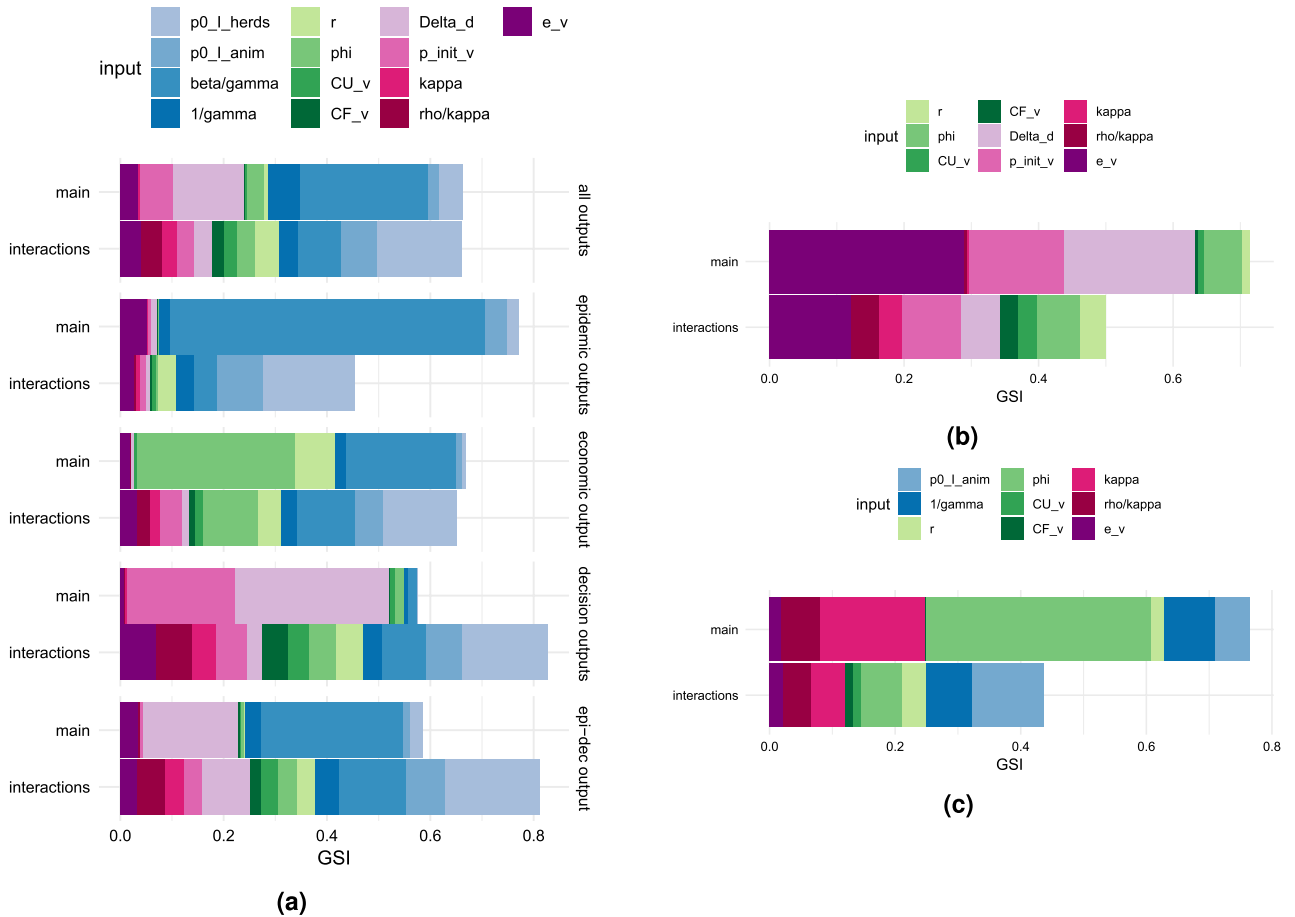


Figure 5. Global Sensitivity Indices (GSI) for the means over runs of the outputs considered in each experiment. Sensitivities are split in main effect and two-factor interactions. Blue colors correspond to epidemic parameters, green colors to economic parameters, and pink colors to decision-related parameters. **(a)** GSI for the means of all outputs, and by group of outputs in experiment (i). **(b)** GSI for the means of all outputs in experiment (ii). **(c)** GSI for the means of decision outputs in experiment (iii). See Table 1 for parameters definition, and Table 2 for outputs definition.

Key determinant parameters to decision-making and epidemiological dynamics. We present the results for the sensitivity analyses on the means over runs for the concerned outputs in each experiment in Fig. 5. Results regarding the variances over runs can be found in Supplementary Fig. S11. Overall, in Fig. 5(a) we see that according to experiment (i), the most influential parameters of the model are the epidemic parameter β/γ , which contributes to 25% of the variation of the means, and the decision-related parameter Δ_d , which contributes to 14%. So only these two parameters account for more than 38% of the variation. They are followed by other epidemic and decision-related parameters, as well as by an economic parameter. Furthermore, for each group of outputs, the parameters with the highest main effects on the means are of the same nature as the outputs (epidemic parameters have the greatest influence on epidemic outputs, economic parameters on the economic output, etc.). For the epidemic outputs, the most influential parameter, β/γ , has a contribution of 61% to the variation of the means. As expected, the exploration of simulation results evidences that this contribution translates into an increase in the propagation of the pathogen.

When focusing on the mean of decision-related outputs, even if interactions have the strongest effect, the most influential main effect is Δ_d , i.e. the duration between two consecutive decisions, which contributes 30% of the variation. It is followed by the initial probability of vaccinating, which contributes 21% to the variation. We remark that Δ_d has an overall negative influence on vaccination of herds, as it determines if control decisions are taken at early stages of the epidemic, and is therefore associated with a higher spread of the pathogen. The initial probability of vaccinating has, on the contrary, a positive effect on the vaccination and on the limitation of the epidemic spread. Concerning the interaction effects, epidemic parameters have the highest influence on the means of each group of outputs, and when considering the means of all outputs together. In particular, $p_{l_{herds}}^0$ is for each group the most influential parameter through its interaction effects. It mostly interacts with other epidemic parameters such as $p_{l_{anim}}^0$, but it has smaller interactions with other parameters as well.

For experiment (ii), Fig. 5(b) shows that when fixing epidemic parameters, overall the greatest main effects are those of the vaccine efficacy e_v (29% contribution), the duration of the decision Δ_d (19% contribution), and the initial probability of vaccinating p_v^{init} (14% contribution). They are followed by the main effect of the economic

parameter: ϕ (6% contribution). Concerning higher order effects, we mainly observe interactions between the first three parameters: e_v , Δ_d , and p_v^{init} . Overall, e_v has the greatest interaction influence. Finally, Fig. 5(c) shows that in experiment (iii) the parameters ϕ and κ manage to explain about 50% of the variability of the means through their main effects, having a 35% and a 16% contribution, respectively. Each of the other parameters explains less than 10% of the variation.

Discussion

In this paper we present a new integrative model for the epidemic spread of a livestock disease on a trade network, accounting for farmers' dynamic decisions concerning the adoption of a control measure in their herd. The model consists of an epidemic–demographic and a decision-making components that are interlinked through a feedback loop. On the one hand, control decisions have consequences on the epidemic spread, both at the intra-herd and the inter-herd levels. On the other hand, the epidemic spread has an impact on the control decisions that farmers subsequently take. For the epidemic–demographic component we use a stochastic compartmental model with demography on a trade network, that accounts for intra-herd population changes, in particular those that concern animal transfers. For the decision-making component we assume the same dynamic decision problem for each farmer, and we propose a mechanism that represents their decision-making strategy.

Whereas most epidemiological models found in literature do not consider the voluntary adoption of a control measure for the spread of a disease⁶, or consider an exogenous probability of applying the measure in order to only study the observed effects of decisions¹⁸, we propose a decision model that considers strategic interactions and cognitive considerations in the decision-making process. Our model can therefore be considered as a game-theoretical or a psychological model, according to the conceptual classification of behavioral epidemiological models found in⁶. The decision-mechanism we propose takes into account different phenomena such as learning, stochastic behavior, and imitation dynamics. To our knowledge, these elements are not present in the few existing models that have aimed at dynamically integrating the epidemic and decision-making processes of a livestock unregulated disease¹⁹. We remark that the basic structure of the decision-problem and the decision-mechanism can be found in different fields, particularly in the field of online optimization (such as multi-armed bandits³⁶). However, we do not seek to find an asymptotically optimal algorithm, which is often the goal in that area, but rather to describe farmers' decision-making process for the application of a control measure such as vaccination. More precisely, we consider an update of the probability of a farmer applying the measure, that is based on self-obtained results and on neighbors' results.

In our model, farmer's next decision is based on a neighbor regardless of what the neighbor has decided in the previous step. This is not the case in similar models focused on human diseases^{13–15}, in which a person only considers other people's observations if they have taken the opposite decision. In particular, this allows to always decrease the odds of a farmer vaccinating if both the farmer and his/her neighbor have previously vaccinated. Together with the use of the trade network as the information network in our model, this can amplify the emergence of strategic behaviors, as the farmer can search to benefit from the vaccination of one of their neighbors, while avoiding the cost of the vaccine. The behavior where individuals (consciously or not) benefit from the actions of others without having to bear the cost, is known as free-riding, and has been previously addressed within vaccination decision-making models for human diseases⁶. In particular³⁷, shows it is possible that individuals will consciously free-ride when making vaccinating decision.

Overall, our integrative model can be considered as an SIR model with pulsed vaccination³⁸ in a metapopulation³⁹, but where the pulse vaccination is asynchronous among sub-populations, and non equally spaced in time for each population, since the decision to vaccinate is not made at each decision time by each farmer. Our formalization of the integrative model is presented as general as possible so it can potentially be adapted to more complex epidemiological models or to other decision-making mechanisms that may be more relevant for specific contexts. Similar models have been previously proposed for human diseases^{13–15}, yet none truly establishes the model in a generic manner in order to facilitate its adaptation for other diseases, or control measures. Even if the economic cost we propose is associated with vaccination and the consequences of an SIR model, its basic structure could take into account the epidemiological and economic consequences of other measures, for example a treatment that would increase the recovery rate of infected animals. In particular, if the epidemic model was aged-structured, the cost on which farmers base their decisions could be refined to take into account the age of the animals. The real-life farmers decision-making being undoubtedly complex, the decision model we propose is reductive. Yet it provides a complete and adaptive framework with respect to state-of-the-art methods in veterinary epidemiology. In the presence of detailed information on farmer's real-life behavior, our model could be run with other parameter values, or it could be modified to stick closer to reality if observations on farmers' decisions denote a different decision-making process.

Among the methodological extensions to consider, we believe that the model could mostly benefit from a relaxation of some hypotheses in the decision-making mechanism. First, we consider that farmers perfectly observe the costs associated to their control decisions, as well as the decisions and costs of their neighbors, which is not completely realistic. Actually, farmers may observe costs with some error, or neighbors may not communicate their true actions or costs. Second, we assume that the trade network is the information network through which farmers share their observations. But farmers may be informed about other herds control practices in a more aggregated way, or only from geographical neighbors. Furthermore, from an economic point of view, in our decision-model we consider only the financial results of the farmer's decision, which is in principle a good indicator of what interests him/her. We remark however that farmers may have social, personal or environmental motivations for taking decisions related to animal welfare⁴⁰. For example, farmers may have a non-use value for their animals, that is, a value related to the animal well-being independently of the use the farmer may make of the animal⁴¹. Even if some refinement could be made in this direction, this does not seem straightforward from

a mathematical modelling perspective. However, our decision model can implicitly integrate this information through the values of the initial probability of vaccinating and the parameters κ and ρ . In addition, considering other types of farmers' behavior can be of interest in this context. For example, the adaption of the exchange network as a function of other farmers' health state. This intervention is known as network rewiring⁴², and is generally appropriate for regulated diseases for which there is aggregated information on the health status of each herd, i.e. neglecting the intra-herd epidemic dynamics. Yet, even with this aggregated information, the adaptation of the network can be quite complex. Lastly, an exploration on the emergence of collaborative behaviors^{43–45}, in particular through network reciprocity⁴⁶, can be an interesting perspective for a deeper understanding of the observed decision dynamics.

Regarding model's predictions, simulations evidence the retroactive effect between the dynamics of the epidemic spread, and the dynamics of the vaccination decisions. A deeper examination of the model through sensitivity analysis confirms that decision-parameters play a role in the model's behavior. Apart from the epidemic parameters, the time between two consecutive decisions has the highest impact overall, and is the main driver in decision-related outputs. Indeed, the shorter the time between decisions, the more frequently farmers evaluate their information on epidemic spread, and the fastest they start vaccinating if necessary. A constantly updated local information on the disease spread regularly helps updating farmers' vaccination decisions from the beginning of the epidemic, and is therefore crucial for limiting the disease spread. This is consistent with observations from models for human diseases, where health information can produce the eradication of the disease if there is a rapid diffusion of this information and if individuals decide to act based on this information⁴⁷. Furthermore, it has been documented that the impact of locally spreading information is amplified if information and disease transmission networks overlap⁴⁸, as it is the case in our model. Finally, an extension of the model where each farmer considers all of his/her neighbors decisions in the decision-making process, evidenced small differences with respect to the model considering only one neighbor per decision instant. In particular, when farmers have a small sensitivity to costs, taking into account all of their neighbors seems to be slightly better in the short term for controlling the epidemic diffusion, but not in the long term. When farmers have a very high sensitivity to costs, considering all of their neighbors does not significantly change the course of the epidemics with respect to the scenario where they consider only one neighbor.

Overall, we conclude that our model effectively integrates the dynamics of the decision process regarding the voluntary adoption of a sanitary measure in each herd, and the dynamics of the epidemic spread over a structured population of herds in a trade network. Hence, we make a significant step towards accounting for human decision-making in mechanistic epidemiological models, in particular for endemic animal diseases. Given its integrative structure, its flexibility and stability in results, our model can be well adapted for simulation studies concerning specific real-life diseases or other control measures.

Code availability

The Python simulation code is available at <https://github.com/CristanchoLina/IntegrativeEpiDecisionModel>. The R packages used for the sensitivity analysis are referenced in "Methods".

Received: 10 December 2020; Accepted: 12 April 2021

Published online: 05 May 2021

References

- Tomley, F. M. & Shirley, M. W. Livestock infectious diseases and zoonoses. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 2637–2642 (2009).
- Carlsake, D. *et al.* Endemic cattle diseases: Comparative epidemiology and governance. *Philos. Trans. R. Soc. B Biol. Sci.* **366**(1573), 1975–1986 (2011).
- Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2011).
- Hidano, A., Enticott, G., Christley, R. M. & Gates, M. C. Modeling dynamic human behavioral changes in animal disease models: Challenges and opportunities for addressing bias. *Front. Vet. Sci.* **5**, 137 (2018).
- Manfredi, P. & D'Onofrio, A. *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases* (Springer, 2013).
- Wang, Z. *et al.* Statistical physics of vaccination. *Phys. Rep.* **664**, 1–113 (2016).
- Morin, B. R., Fenichel, E. P. & Castillo-Chavez, C. SIR dynamics with economically driven contact rates. *Nat. Resour. Model.* **26**(4), 505–525 (2013).
- Perrings, C. *et al.* Merging economics and epidemiology to improve the prediction and management of infectious disease. *EcoHealth* **11**(4), 464–475 (2014).
- Keeling, M. J. & Eames, K. T. Networks and epidemic models. *J. R. Soc. Interface* **2**(4), 295–307 (2005).
- Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *J. Theor. Biol.* **251**(3), 450–467 (2008).
- Scaman, K., Kalogeratos, A. & Vayatis, N. Suppressing epidemics in networks using priority planning. *IEEE Trans. Netw. Sci. Eng.* **3**(4), 271–285 (2016).
- Braunstein, A., Dall'Asta, L., Semerjian, G. & Zdeborova, L. Network dismantling. *Proc. Natl. Acad. Sci.* **113**(44), 12368–12373 (2016).
- Fu, F., Rosenbloom, D. I., Wang, L. & Nowak, M. A. Imitation dynamics of vaccination behaviour on social networks. *Proc. R. Soc. B Biol. Sci.* **278**(1702), 42–49 (2010).
- Wang, Z., Szolnoki, A. & Perc, M. If players are sparse social dilemmas are too: Importance of percolation for evolution of cooperation. *Sci. Rep.* **2**, 369 (2012).
- Kuga, K., Tanimoto, J. & Jusup, M. To vaccinate or not to vaccinate: A comprehensive study of vaccination-subsidizing policies with multi-agent simulations and mean-field modeling. *J. Theor. Biol.* **469**, 107–126 (2019).
- Horan, R. D., Fenichel, E. P., Wolf, C. A. & Gramig, B. M. Managing infectious animal disease systems. *Annu. Rev. Resour. Econ.* **2**(1), 101–124 (2010).
- Tago, D., Hammitt, J. K., Thomas, A. & Raboisson, D. The impact of farmers' strategic behavior on the spread of animal infectious diseases. *PLoS One* **11**(6), e0157450 (2016).

18. Rat-Aspert, O. & Fourichon, C. Modelling collective effectiveness of voluntary vaccination with and without incentives. *Prev. Vet. Med.* **93**(4), 265–275 (2010).
19. Krebs, S., Picault, S. & Ezanno, P. Modelisation multi-agents pour la gestion individuelle et collective d'une maladie infectieuse. In *HAL* (ed. Picard, G.) (Christophe Lang, Nicolas Marilleau, 2018).
20. Shi, B., Liu, G., Qiu, H., Wang, Z., Ren, Y. & Chen, D. Exploring voluntary vaccination with bounded rationality through reinforcement learning. In *Physica A: Statistical Mechanics and its Applications* **515** (2019), pp. 171–182. ISSN: 0378-4371. <https://doi.org/10.1016/j.physa.2018.09.151>. <https://www.sciencedirect.com/science/article/pii/S0378437118312664>.
21. Heesterbeek, H. *et al.* Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347**, 6227 (2015).
22. Brooks-Pollock, E., De Jong, M., Keeling, M., Klinkenberg, D. & Wood, J. Eight challenges in modelling infectious livestock diseases. *Epidemics* **10**, 1–5 (2015).
23. Hershey, J. C., Asch, D. A., Thumasathit, T., Meszaros, J. & Waters, V. V. The roles of altruism, free riding, and bandwagoning in vaccination decisions. *Organ. Behav. Hum. Decis. Process.* **59**(2), 177–187 (1994).
24. Beaunee, G., Vergu, E. & Ezanno, P. Modelling of paratuberculosis spread between dairy cattle farms at a regional scale. *Vet. Res.* **46**, 111 (2015).
25. Cousins, D. Mycobacterium BOVIS infection and control in domestic livestock. *Rev. Sci. Tech.* **20**(1), 71–85 (2001).
26. Mortensen, S. *et al.* Risk factors for infection of sow herds with porcine reproductive and respiratory syndrome (PRRS) virus. *Prev. Vet. Med.* **53**12, 83–101 (2002).
27. Breto, C., He, D., Ionides, E. L. & King, A. A. Time series analysis via mechanistic models. *Ann. Appl. Stat.* **3**(1), 319–348 (2009).
28. Nagashima, K. & Tanimoto, J. A stochastic Pairwise Fermi rule modified by utilizing the average in payoff differences of neighbors leads to increased network reciprocity in spatial prisoner's dilemma games. *Appl. Math. Comput.* **361**, 661–669 (2019).
29. Saltelli, A. *et al.* *Global Sensitivity Analysis: The Primer* (Wiley, 2008).
30. Gunst, R. F. & Mason, R. L. Fractional factorial design. *Wiley Interdiscip. Rev. Comput. Stat.* **1**(2), 234–244 (2009).
31. Kobilinsky, A., Bouvier, A. & Monod, H. *PLANOR: An R package for the automatic generation of regular fractional factorial designs*. R package version 1.5-3. INRA, MIA, Jouy en Josas, France, 2020.
32. Montgomery, D. C. *Design and Analysis of Experiments* (Wiley, 2017).
33. Lamboni, M., Monod, H. & Makowski, D. Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliabil. Eng. Syst. Saf.* **96**(4), 450–459 (2011).
34. Van Schepdael, A., Carlier, A. & Geris, L. Sensitivity analysis by design of experiments. In *Uncertainty in Biology* 327–366 (Springer, 2016).
35. Bidot, C., Lamboni, M. & Monod, H. *multisensi: Multivariate Sensitivity Analysis*. R package version 2.1-1. 2018. <https://CRAN.R-project.org/package=multisensi>.
36. Auer, P., Cesa-Bianchi, N. & Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.* **47**(2–3), 235–256 (2002).
37. Ibuka, Y., Li, M., Vietri, J., Chapman, G. B. & Galvani, A. P. Free-riding behavior in vaccination decisions: An experimental study. *PLoS One* **9**1, e87164 (2014).
38. Choisy, M., Guegan, J.-F. & Rohani, P. Mathematical modeling of infectious diseases dynamics. In: *Encyclopedia of Infectious Diseases: Modern Methodologies*, 379 (2007).
39. Terry, A. J. Pulse vaccination strategies in a metapopulation SIR model. *Math. Biosci. Eng.* **7**(2), 455 (2010).
40. Hansson, H. & Lagerkvist, C. J. Identifying use and non-use values of animal welfare: Evidence from Swedish dairy agriculture. *Food Policy* **50**, 35–42 (2015).
41. Lagerkvist, C. J., Hansson, H., Hess, S. & Hoffman, R. Provision of farm animal welfare: Integrating productivity and non-use values. *Appl. Econ. Perspect. Policy* **33**(4), 484–509 (2011).
42. Kiss, I. Z. *et al.* *Mathematics of Epidemics on Networks* 598 (Springer, 2017).
43. Wang, Z., Moreno, Y., Boccaletti, S. & Perc, M. "Vaccination and epidemics in networked populations—an introduction". In: *Chaos Solitons Fractals* **103**, 177–183 (2017). ISSN: 0960-0779. <https://doi.org/10.1016/j.chaos.2017.06.004>. <https://www.sciencedirect.com/science/article/pii/S0960077917302515>.
44. Karlsson, C.-J. & Rowlett, J. Decisions and disease: A mechanism for the evolution of cooperation. *Sci. Rep.* **10**(1), 1–9 (2020).
45. Cardillo, A., Reyes-Suarez, C., Naranjo, F. & Gomez-Gardenes, J. Evolutionary vaccination dilemma in complex networks. *Phys. Rev. E* **88**(3), 032803 (2013).
46. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**(5805), 1560–1563 (2006).
47. Kiss, I. Z., Cassell, J., Recker, M. & Simon, P. L. The impact of information transmission on epidemic outbreaks. *Math. Biosci.* **225**(1), 1–10 (2010).
48. Funk, S., Gilad, E., Watkins, C. & Jansen, V. A. The spread of awareness and its impact on epidemic outbreaks. *Proc. Natl. Acad. Sci.* **106**(16), 6872–6877 (2009).

Acknowledgements

This work was carried out with the financial support of the French Research Agency (ANR), project ANR-16-CE32-0007 (CADENCE). We are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi: 10.15454/1.5572390655343293E12) for providing computing resources. We thank the French minister of agriculture for providing the access to the FCID.

Author contributions

L.C.F., P.E., and E.V. conceived and designed the study. L.C.F. conducted the analyses, prepared the figures and wrote the manuscript. P.E. and E.V. advised model construction and model analyses, advised epidemiological details, discussed results and edited manuscript. All authors revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-88471-6>.

Correspondence and requests for materials should be addressed to L.C.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021, corrected publication 2021

Research



Cite this article: Cristancho-Fajardo L, Ezanno P, Vergu E. 2022 Dynamic resource allocation for controlling pathogen spread on a large metapopulation network. *J. R. Soc. Interface* **19**: 20210744.
<https://doi.org/10.1098/rsif.2021.0744>

Received: 23 September 2021
 Accepted: 10 February 2022

Subject Category:
 Life Sciences–Mathematics interface

Subject Areas:
 biomathematics

Keywords:
 infectious disease control, stochastic modelling, optimization, greedy scores

Author for correspondence:
 Lina Cristancho-Fajardo
 e-mail: lina.cristancho-fajardo@inrae.fr

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5872682>.

Dynamic resource allocation for controlling pathogen spread on a large metapopulation network

Lina Cristancho-Fajardo^{1,2}, Pauline Ezanno² and Elisabeta Vergu¹

¹Université Paris-Saclay, INRAE, MaLAGE, Jouy-en-Josas 78350, France

²INRAE, Oniris, BIOEPAR, Site de la Chantrerie, CS40706, Nantes 44307, France

LC-F, 0000-0001-7216-4037; PE, 0000-0002-0034-8950; EV, 0000-0001-6470-7898

To control the spread of an infectious disease over a large network, the optimal allocation by a social planner of a limited resource is a fundamental and difficult problem. We address this problem for a livestock disease that propagates on an animal trade network according to an epidemiological–demographic model based on animal demographics and trade data. We assume that the resource is dynamically allocated following a certain score, up to the limit of resource availability. We adapt a greedy approach to the metapopulation framework, obtaining new scores that minimize approximations of two different objective functions, for two control measures: vaccination and treatment. Through intensive simulations, we compare the greedy scores with several heuristics. Although topology-based scores can limit the spread of the disease, information on herd health status seems crucial to eradicating the disease. In particular, greedy scores are among the most effective in reducing disease prevalence, even though they do not always perform the best. However, some scores may be preferred in real life because they are easier to calculate or because they use a smaller amount of resources. The developed approach could be adapted to other epidemiological models or to other control measures in the metapopulation setting.

1. Introduction

Infectious disease spread is a problem that can have important social, sanitary and economic consequences. Like for human diseases, this is a major public health concern for animal diseases, for guaranteeing animal welfare and food security [1]. In this context, epidemiological models, together with other relevant mathematical approaches, can help in the description and understanding of the mechanisms involved in disease propagation, as well as in assessing the effectiveness of control measures [2]. An approach for controlling a disease spreading on a population, from a social planner’s point of view, is the allocation of a resource that has an effect on this spread [3]. Many questions can arise in this context: how much resource is needed to restrain the disease propagation to a certain level [4–6], when should it be allocated [7] and where. In this work, we are interested in the third question. More specifically, we are concerned with the problem of dynamically deciding where to allocate a limited available resource in an optimal manner, in order to minimize disease spread on a large animal metapopulation network.

On the one hand, most of the research addressing the issue of resource optimal allocation on a large network [8–12] does not focus on metapopulation networks, i.e. does not account for infection-related dynamics within each sub-population represented by a node of the network. Even more, works relying on mean-field theory [13] do not consider structured populations, in particular as a network.

On the other hand, the existing studies addressing the resource allocation problem on metapopulation networks are based on techniques that lack of scalability. For example, optimal control [14] and reinforcement learning [15], which would theoretically give an optimal strategy, cannot be used in the context of very large networks due to the dimensionality in the state space [16,17]. Even if we consider a sub-population as healthy or infected (only two possible states per sub-population), a network with J sub-populations would have 2^J possible health states, which yields an asymptotically intractable optimal allocation when the network is very large. Hence, authors that study resource allocation on metapopulation networks generally build and evaluate their approach in a small number of sub-populations, usually less than 50 [18–25]. Recently, [26] proposed a framework built upon optimal control theory that is able to deal with the dynamic allocation problem in a network of hundreds of sub-populations thanks to several simplifications, among which is considering only a subset of edges for the optimization. Yet, such a scale does not allow one to capture the structural characteristics of complex networks, such as the animal trade network we consider here. In particular, animal trade networks are in general scale-free [27] (most herds have few trading partners while a few herds have many) and dynamic, as the amount of exchanged animals can vary over time [28].

Finally, published studies assume the resource to be in general distributed only once, before or at the beginning of an outbreak [21,29]. Therefore, the resource allocation problem is static. However, the allocation problem can be intrinsically dynamic if it is studied in the long run. For example, if the available resource is a vaccine, this can have a limited effect in time, so there is need for several vaccination campaigns.

Given the intractability of the optimal strategy, in this work we restrict ourselves to score-based strategies, i.e. strategies that consist of allocating the resource according to a scoring function (or indicator), up to the limit given by the available quantity of the resource. Furthermore, given the practical importance of dynamical aspects of the allocation problem, we include this view in the present study, i.e. we assume that the resource allocation is dynamic.

The contribution of this work is twofold. First, from a methodological perspective, by adapting the greedy approach in [30], we obtain analytic scores for controlling disease spread on a large animal trade network, where the disease propagation is represented by a stochastic SIR model that accounts for demography and trade (model introduced in §2.1). The approach consists of finding the scoring function that minimizes a short-term approximation of a given objective function (§§2.2, 2.3.1 and 3.1). Our generalization is mainly driven by the metapopulation framework, which implies that a herd is not only infected or healthy, but that it has an internal infection and demographic dynamics. In particular, this allows for the possibility of needing different amounts of resources for different herds. Furthermore, we extend this approach for two different types of resources (vaccination and treatment) and two different objective functions (the number of infected animals and the number of infected herds). Second, regarding real-life disease control for livestock diseases, in §3.2, we evaluate the performances of the analytically obtained strategies along with the one of several heuristic strategies (introduced in

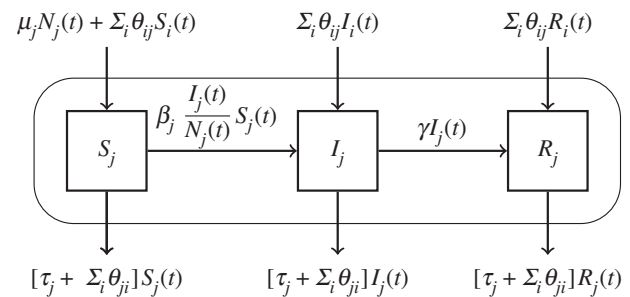


Figure 1. Schematic representation of the intra-herd infection and demographic dynamics for a herd j , without resource allocation. Horizontal arrows represent transitions between health-related compartments, corresponding to the course of infection inside the herd (curved rectangle), while vertical arrows represent animal flows to and from the herd. Coefficients on the arrows are transition rates. See main text in §2.1 for parameter definitions.

§2.3.2) that can be relevant for this context. Finally, in §4, we extend the interpretation of the analytically found scores and discuss their suitability in a metapopulation context in the light of simulation-based results. We also consider several perspectives, based either on the development of other greedy scores or on new simulation studies that could use the explorations performed in this work.

2. Methods

2.1. SIR stochastic epidemiological model with demography in a metapopulation based on a trade network

We considered a livestock disease that spreads on a large animal trade network composed of J herds. We supposed that the disease introduction in a herd could only be due to animal transfers, and that it could only be transmitted between animals of the same herd. This livestock trade network underlies a metapopulation network where nodes represent herds and links represent animal transfers from one herd to another. For representing this system, we used the stochastic intra-herd SIR epidemiological-demographic model described in [31], which takes into account animal exchanges. The model is summarized in figure 1. $S_j(t)$, $I_j(t)$ and $R_j(t)$ are the number of susceptible, infected and recovered animals in herd j at time t . Parameters β_j , μ_j and τ_j are the daily rates of disease transmission, birth and death in herd j , assuming newborns are all susceptible. As for γ , it is the daily recovery rate from the infection, reasonably assumed equal for all herds. Finally, θ_{ji} is the daily out rate of animals going from herd j to herd i , assuming animals in any health state can be exchanged. We denote as $N_j(t)$ the size of herd j at time t . The model was specified as a continuous-time Markov chain, and its simulation was built on an Euler discrete-time scheme using multinomial distributions, as described in [32]. Details can be found in the electronic supplementary material of [31].

2.2. Dynamic resource allocation problem in the metapopulation framework

We supposed that there is a central social planner seeking to minimize the disease propagation on the animal trade network, by distributing a limited amount of a resource among the herds in the network, dynamically with a given decision time-step. This

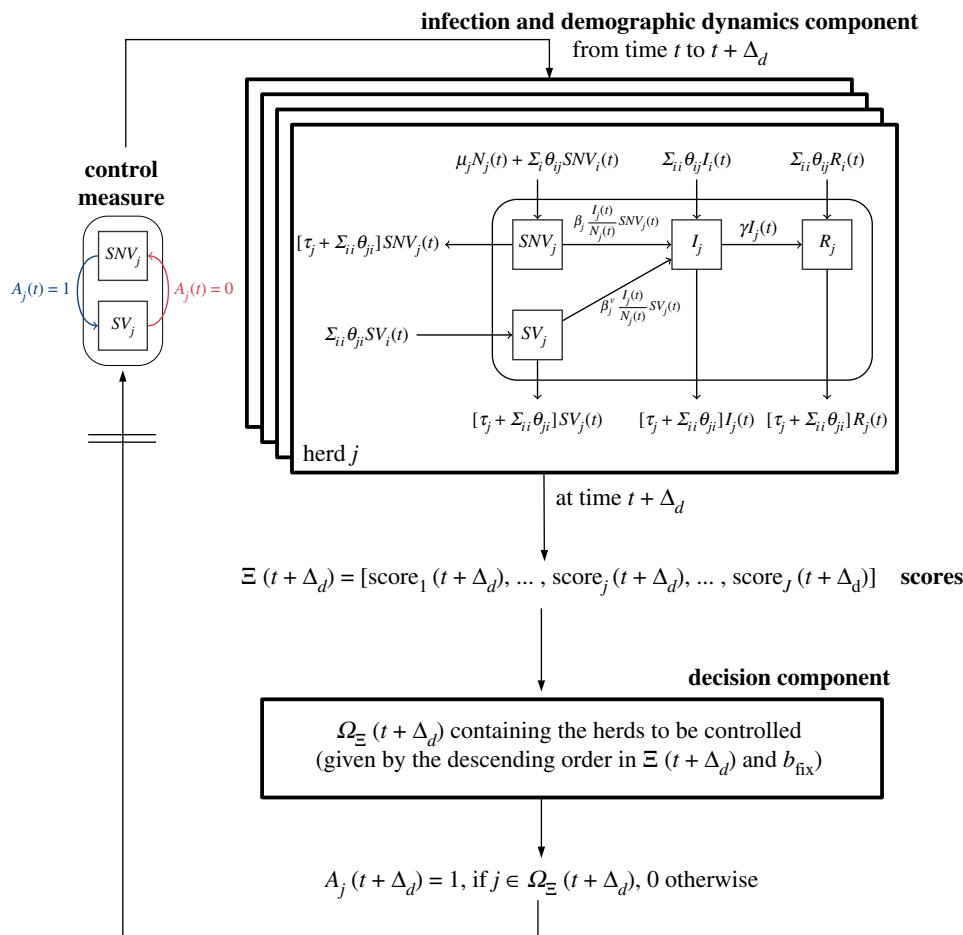


Figure 2. Representation of the modelling and optimization framework of the score-based dynamic resource allocation under constraint, applied to a metapopulation and vaccine allocation. See main text in §§ 2.1 and 2.2 for parameter definitions.

dynamic resource allocation problem was formulated as:

$$\min_A F(A) \quad \text{subject to} \quad \sum_{j=1}^J b_j(t) A_j(t) \leq b_{\text{fix}}, \quad (2.1)$$

$$\forall t = \Delta_d, 2\Delta_d, \dots$$

$F(A)$ in equation (2.1) is the function that the planner has to minimize, which depends on the allocation strategy A , i.e. the function that determines the decisions $A_j(t)$, $\forall j = 1, \dots, J$, $\forall t = \Delta_d, 2\Delta_d, \dots$, where j denotes the herd, and Δ_d is the decision time-step. We assumed binary allocation decisions for each herd, i.e. $A_j(t) = 1$ if the resource is allocated to herd j at time t , $A_j(t) = 0$ otherwise. In the condition of equation (2.1), b_{fix} is the quantity of resource that is available at each decision time and $b_j(t)$ is the quantity of the resource that would be needed for herd j at time t if this herd was selected for the allocation.

We supposed that the resource to allocate could be either a vaccine or a treatment. For the vaccine, we assumed that if applied to a susceptible animal at time t , the disease transmission rate towards this susceptible animal becomes $\beta^v = \beta(1 - e_v)$ during the period $[t; t + \Delta_d]$, where $0 \leq e_v \leq 1$ is the protective efficacy of the vaccine. That is, Δ_d is also the duration of the vaccine's efficacy. We underline that even if all susceptible animals in herd j are vaccinated at time t , infections can occur within j from time t to $t + \Delta_d$ if new susceptible animals enter the herd, through births or imports of susceptible animals from other herds. For the treatment, we supposed that it increases the recovery rate of treated infected animals by an additional factor γ' , i.e. reducing the mean duration of the infectious period for these animals to $(\gamma + \gamma')^{-1}$.

Regarding the resource constraint, for vaccination we assumed it concerned the number of available doses, and if a

herd j was selected at time t all the animals in the herd would be vaccinated. So in the condition in equation (2.1), $b_j(t) = N_j(t)$ (the size of herd j at time t). For the treatment, the resource constraint was on the number of herds in which animals were treated at each decision time, so $b_j(t) = 1$. The choice to consider that the constraint for the treatment involves the number of herds was mainly motivated by analytical considerations discussed in §2.3.1.

2.3. Score-based strategies

Score-based strategies consist of ordering herds according to a certain scoring function Ξ and selecting the top herds, up to the limit given by the condition in equation (2.1). Let $\Omega_{\Xi}(t)$ be the set that contains the selected herds according to $\Xi(t)$ (the score values at time t) and b_{fix} (the available quantity of resource per decision time-step). Then, $A_j(t) = 1$ if $j \in \Omega_{\Xi}(t)$, 0 otherwise. Figure 2 represents the modelling and optimization framework of the dynamic resource allocation of vaccines under a score-based strategy. The treatment allocation differs only in the infection and demographic dynamics component, and in the control measure component. That is, for any herd j , there is an additional compartment T_j for treated infected animals, where animals go from I_j to T_j at decision time t , if $A_j(t) = 1$. Unlike vaccination, the transition from T_j to I_j is not possible even if $A_j(t) = 0$, i.e. treated animals can only recover.

Regarding the scoring function, it can be either optimized or heuristic. In the following, we address the two possibilities.

2.3.1. Greedy scores

First, following the approach in [30] we searched for optimized scoring functions. This approach consists of finding a scoring function that minimizes a short-term approximation of the

objective function

$$F(A) := \int_0^{\infty} e^{-at} \mathbb{E}[J_I^A(t)] dt,$$

where $J_I^A(t)$ is a function of the infection-related state of the network at time t . The objective establishes an infinite time horizon, and a parameter $a \geq 0$ which reduces the long-term impact of $J_I^A(t)$. This means that the larger a , the more we are interested in the short-term behaviour of the infection dynamics, i.e. the more weight we put on the initial stochasticity of the disease diffusion.

Let us denote by $X(t)$ the infection-related state of the network at time t . In the SIR metapopulation framework, $X(t)$ is a $J \times 3$ matrix where each row $j \in \{1, \dots, J\}$ contains the values of $S_j(t)$, $I_j(t)$ and $R_j(t)$ at time t . Since the allocation strategy has no impact on the initial state, and since the process is Markovian, the minimization problem (equation (2.1)) is equivalent to

$$\min_A \int_{u=0}^{\infty} e^{-au} \mathbb{E}[J_I^A(t+u) | X(t) = X] du, \quad (2.2)$$

for all decision times t and for all network possible states X , under the same constraint in equation (2.1). Then, focusing on the short-term behaviour of the system, it is possible to obtain an approximation of the objective function, which is based on a Taylor expansion of $\mathbb{E}[J_I^A(t+u) | X(t) = X]$ for a small value of u . Finally, one must find the scoring function that minimizes this approximation. This approach is therefore called greedy, as it yields locally (i.e. short term) optimal allocation decisions.

We stress that the form of the scoring function will not depend on the parameter a in equation (2.2), i.e. we do not need to fix a value for this parameter when exhibiting the scoring function. Yet, the higher is a , the lower is the impact of the approximation accuracy. See the electronic supplementary material, S1, for more details on the approach.

We adapted this greedy approach to the metapopulation framework by considering, for each one of the two types of resource (vaccination and treatment), the minimization of two different objective functions in equation (2.2). A function on the number of infected animals

$$J_I^A(t) = \sum_{j=1}^J I_j(t), \quad (2.3)$$

and a function on the number of infected herds

$$J_I^A(t) = \sum_{j=1}^J \mathbf{1}_{I_j(t) > 0}. \quad (2.4)$$

That is, we treated four different cases, depending on the type of resource (vaccine or treatment) and the objective to be minimized: the number of infected animals (equation (2.3)) or the number of infected herds (equation (2.4)). In particular, the derivation of the score for the allocation of a treatment under the objective of equation (2.4) required to consider that the resource constraint of equation (2.1) was expressed in terms of the number of attainable herds ($b_j(t) = 1$). This allowed the total number of treated herds to be formulated as the minimum between b_{fix} and the number of herds that have exactly one infected animal (see electronic supplementary material, S1.2.1, for details).

2.3.2. Heuristic scores

In addition, we considered three types of heuristic scores based on: the topology of the static aggregated network; the demographic changes in the network; and the dynamic infection-related state of the network. Table 1 contains the list of the 16 heuristic scoring functions we tested for the metapopulation framework: five topological ones (in-strength, out-strength, closeness, betweenness and pagerank), five demographic ones

($N_j(t)$, $\text{purchases}_j(0, t)$, $\text{purchases}_j(t - \Delta_d, t)$, $\text{sales}_j(0, t)$, $\text{sales}_j(t - \Delta_d, t)$), 5 epidemiological ones ($s_j(t)$, $i_j(t)$, $r_j(t)$, $i_j(t - \Delta_d, t)$, $r_j(t - \Delta_d, t)$), and a random scoring function. All the topological scoring functions are classical centrality measures in networks [33].

2.4. Simulation setting and numerical explorations

The simulation setting was described in detail in the electronic supplementary material of [31]. The metapopulation structure was set close to real data, extracted from the French Cattle Identification Database (FCID): animal movements correspond to the Finistère administrative area in western France, which is densely populated with cattle. The trade parameters (θ_{ji} , $\forall j, i = 1, \dots, J$) underlying the network structure were set to values based on these data. The network is scale-free and consists of $J = 5000$ herds. The initial herd size distribution can be found in the electronic supplementary material, figure S1a. Details on the trade parameters and on the herd size distribution can be found in electronic supplementary material, fig. S3 of [31].

Without loss of generality, the values of demographic and epidemiological parameters were set to the same value for all herds. In particular, death and birth rates were $\tau_j = \tau$, $\mu_j = \mu$, $\forall j = 1, \dots, J$, where $\tau = 0.0009 \text{ days}^{-1}$ and $\mu = 2\tau$. So, based on field reality, animals have a mean life time of approximately three years, and on average an animal gives birth to a calf every 1.5 years. As herd sizes could vary not only through births and deaths but also through animal movements, we established a constraint on the capacity of each herd, $N_j(t) \leq 1.5 N_j(0)$, so that these remain relatively stable over time. Electronic supplementary material, figure S1b, shows the final herd size distribution for a run of the epidemiological–demographic model (without resource allocation) on the simulated trade network. Regarding the epidemiological parameters, we set $\beta_j = \beta$, $\forall j = 1, \dots, J$, and considered a disease with moderate immediate impact and long-lasting development ($\beta/\gamma = 2$ and $1/\gamma = 90$ days). A second numerical setting, corresponding to a disease with higher early peak and smaller infection duration ($\beta/\gamma = 4$, $1/\gamma = 30$ days), was explored in the electronic supplementary material, S2. Finally, for the available resource, we assumed it could either be a perfectly effective protective vaccine, i.e. $e_v = 1$, or a treatment that greatly reduces the infectious period of infected animals (but which is not perfect in order to avoid instantaneous recovery, an unrealistic assumption). More specifically, the duration of the infectious period with treatment was assumed to be 3% of the duration without treatment, i.e. the mean duration of the infectious period for a treated infected animal is $(\gamma + \gamma')^{-1} = 0.03(1/\gamma) = 2.7$ days.

2.4.1. Setting for the exploration of infection-related dynamics with score-based resource allocation

Given these parameter values, we simulated the infection-related dynamics of the metapopulation during 3 years in 74 ($([16 + 3] + [16 + 2]) \times 2$) cases characterized by the type of resource, the score according to which it is allocated (16×3 for vaccination and 16×2 for treatment) and the scenario. This last one can be: an *epidemic* scenario, where initially 10% of the herds (chosen completely at random) had a random subset of 15% of their animals infected; or an *endemic* scenario, where the initial state was given by the state at roughly 3 years (1080 days) without resource allocation departing from the epidemic scenario. Indeed, electronic supplementary material, figure S9, shows that if the simulations are extended beyond 3 years, the total proportion of infected animals remains rather stable, and that there is only a 10% reduction in the proportion of infected herds between levels attained at 3 and 9 years. Hence, although the infection dynamics after 3 years of simulation did not reach a steady state rigorously speaking, this date was chosen as the initial point of the endemic scenario. Indeed, on the one hand,

Table 1. Heuristic scoring functions for herd j at time t . Dependence on t means the score is dynamic in time, otherwise it is static.

	scoring function for herd j	description
topological	in-strength $_j$	weighted in-degree, i.e. the sum of the weights of all edges going to j . Weights \equiv daily mean trade rates. Measures how much j purchases from other herds
	out-strength $_j$	weighted out-degree, i.e. the sum of the weights of all edges going from j . Weights \equiv daily mean trade rates. Measures how much j sells to other herds
	closeness $_j$	inverse of the sum of distances to other herds, based on the weighted shortest paths going from/to j . Weights \equiv daily mean trade rates. Measures how 'close' j is to other herds
	betweenness $_j$	ratio of the weighted shortest paths that pass through j over all such paths in the network. Weights \equiv daily mean trade rates. Measures the influence of j on the animal exchanges of the network
	pagerank $_j$	stationary density of a discrete-time random walk. Weights \equiv daily mean trade rates. Measures the accessibility to j
demographic	$N_j(t)$	current herd size
	purchases $_j(0, t)$	number of animals bought by j over the period $]0, t]$ and $]t - \Delta_d, t]$, respectively
	purchases $_j(t - \Delta_d, t)$	
	sales $_j(0, t)$	number of animals sold by j over the period $]0, t]$ and $]t - \Delta_d, t]$, respectively
sales $_j(t - \Delta_d, t)$		
epidemiological	$s_j(t)$	current proportion of susceptible animals in j : $\frac{S_j(t)}{N_j(t)}$
	$i_j(t)$	current proportion of infected animals in j : $\frac{I_j(t)}{N_j(t)}$
	$r_j(t)$	current proportion of recovered animals in j : $\frac{R_j(t)}{N_j(t)}$
	$i_j(t - \Delta_d, t)$	increment of the proportion of infected animals in j between times $t - \Delta_d$ and t : $\frac{I_j(t)}{N_j(t)} - \frac{I_j(t - \Delta_d)}{N_j(t - \Delta_d)}$
	$r_j(t - \Delta_d, t)$	increment of the proportion of recovered animals in j between times $t - \Delta_d$ and t : $\frac{R_j(t)}{N_j(t)} - \frac{R_j(t - \Delta_d)}{N_j(t - \Delta_d)}$
	random $_j(t)$	assigns a random score for j at each decision time t

Table 2. Parameter values in the allocation problem depending on the type of resource.

resource	parameter	definition	values
vaccine	b_{fix}	number of available doses at each decision time (as a % of the initial total number of animals)	$(25\% \times \sum_{j=1}^J N_j(0))^a$ and $([5\% - 90\%] \times \sum_{j=1}^J N_j(0))^b$
	Δ_d	decision step (in days)	180
treatment	b_{fix}	number of attainable herds at each decision time	25^a and $[5 - 100]^c$
	Δ_d	decision step (in days)	15

^aValues for infection-related dynamics explorations.

^bValues for percolation analysis of vaccination. From 5% to 30%: by 0.5%. From 40% to 90%: by 10%.

^cValues for percolation analysis of treatment. From 5 to 75: by 5. From 80 to 100: by 10.

at this date the pathogen had widely spread in the metapopulation (electronic supplementary material, figure S9), and on the other hand, considering 3 years limited the simulation cost. In each case, we explored the dynamics of the proportion of infected herds and of the total number of infected animals for a fixed value of the available quantity of resource, b_{fix} . The values of b_{fix} and Δ_d parameters can be found in table 2. We supposed that vaccination decisions were more spaced in time since vaccines are preventive and tend to have long-lasting effect.

Meanwhile, we supposed that treatment decisions were more frequent as they are more prone to being applied in a critical situation.

We stress that in all simulations for the vaccine allocation, only herds with at least one susceptible animal were eligible, and that the treatment allocation only considered herds with at least one infected animal.

Furthermore, we evaluated how other quantities of interest, raw proxies of 'costs', varied with the chosen scoring function

Table 3. Greedy scoring functions studied in the numerical explorations. All the greedy scores are dynamic.

resource	$J_j^A(t)$	scoring function for herd j	scoring function name
vaccine	inf. animals (equation (2.3))	$(I_j(t)/N_j(t))S_j(t)$	<i>greedyV_infanimals</i>
	inf. herds (equation (2.4))	$(I_j(t)/N_j(t))S_j(t)(1_{I_j(t)=1}(\gamma + \tau + \sum_{i \neq j} \theta_{ij}) + \sum_{i \neq j} \theta_{ij} 1_{I_i(t)=0})$	<i>greedyV_infherds</i>
	inf. herds (equation (2.4))	$(I_j(t)/N_j(t))S_j(t)(1_{0 < I_j(t) < 20}(\gamma + \tau + \sum_{i \neq j} \theta_{ij}) + \sum_{i \neq j} \theta_{ij} 1_{I_i(t)=0})$	<i>greedyV_infherds_threshold</i>
treatment	inf. animals (equation (2.3))	$I_j(t)$	<i>greedyT_infanimals</i>
	inf. herds (equation (2.4))	$[-\sum_{i \neq j} \theta_{ij} 1_{I_i(t) > 0}]_{j: I_j(t)=1}$	<i>greedyT_infherds</i>
	inf. herds (equation (2.4))	$[(-\sum_{i \neq j} \theta_{ij} 1_{I_i(t) > 0}) 1_{0 < I_j(t) < 20}]$	<i>greedyT_infherds_threshold</i>

and the amount of available resource. We stress that the quantity of resource that is actually used may be less than the available quantity (b_{fix}). For vaccination, we explored the mean proportion of wasted doses over time, i.e. the total number of vaccine doses allocated to infected or recovered animals, out of the number of available doses. For the treatment, where we assumed that only infected animals were treated, we evaluated the mean number of used doses over time, and we explored the mean size of the target population, i.e. the total number of animals in treated herds.

2.4.2. Setting for percolation analysis

We performed percolation experiments by assessing the final (after 3 years) proportion of infected herds when the allocation was done using a given scoring function, and a fixed quantity of resource. We also included in this analysis an exploration of the raw proxies of costs. The range of values tested for b_{fix} (cf. table 2) was chosen so as to exacerbate differences in the infection-related dynamics by scoring function. The objective was to analyse the dependence of the relative performances of the scoring functions on limiting disease spread with respect to the available quantity of the resource. To avoid increasing computation time, we chose to run the percolation analysis only for a subset of scoring functions. This subset was determined from the results of the complete exploration setting described in §2.4.1.

Finally, we investigated the sensibility of the infection-related results of percolation analysis to the value of the vaccine efficacy e_v . More specifically, we evaluated through simulations two cases where vaccine efficacy is not perfect: $e_v = 0.9$ and $e_v = 0.7$. Lower values for e_v were not considered, since it is very unlikely that a vaccine for a livestock disease with an efficacy below 0.7 is even considered in the field, as its implementation could induce greater economic costs than epidemiological benefits. With each of these values we ran the percolation analysis described above, for three values of the available number of doses: [25%, 40%, 70%] of the initial total number of animals in the metapopulation.

3. Results

3.1. Greedy scoring functions

We obtained four different analytic scoring functions by considering the minimization problem in equation (2.2) for each resource and for each objective with the greedy approach. The scoring functions for an objective on the number of infected herds (*greedyV_infherds* and *greedyT_infherds*) were obtained through a second-order Taylor expansion, while the other two scoring functions (*greedyV_infanimals* and *greedyT_infanimals*) were obtained using a first-order approximation. Details on how the greedy scoring functions were computed in each case can be found in the electronic supplementary material, S1.

As mentioned in the previous section, the values of β_j , τ_j , μ_j were set equal for all herds in our simulations, so we present in table 3 the scores in this setting. The greedy scoring functions found for the generic framework, where these values can be different across herds, can be found in electronic supplementary material, table S1. In simulations, five scoring functions (directly or indirectly issued from the greedy approach) were used. Table 3 includes in total six scoring functions: the four obtained by optimization, an additional scoring function for vaccination (*greedyV_infherds_threshold*) and the scoring function *greedyT_infherds_threshold* for the treatment, which replaced *greedyT_infherds* in our simulations.

To minimize the number of infected animals by distributing a vaccine, the *greedyV_infanimals* scoring function privileges herds with a large within-herd incidence rate ($\beta I_j(t) S_j(t) / N_j(t)$), i.e. many infected animals and a large proportion of susceptible animals. For the treatment, with *greedyT_infanimals* the allocation would be made only as a function of the number of infected animals by herd ($I_j(t)$).

Regarding the minimization of the number of infected herds, for vaccination it led to a scoring function, *greedyV_infherds*, favouring two types of herds: either herds that have a large within-herd incidence rate, and that send to many healthy herds (large $\sum_{i \neq j} \theta_{ij} 1_{I_i(t)=0}$); or herds with only one infected animal, also presenting a large proportion of susceptible animals and which sell many animals (large $\sum_{i \neq j} \theta_{ij}$). With the same objective for treatment, the *greedyT_infherds* scoring function only concerns herds with exactly one infected animal, and among these, the priority is on herds that send the smallest flows to infected buyers (small $\sum_{i \neq j} \theta_{ij} 1_{I_i(t) > 0}$).

We remark that the minimization of the function on the number of infected animals for vaccination led to scoring functions that only depend on the epidemiological state of herd j , but not on the states of other herds, and in particular not on the topology of the network. This is due to the use of a first-order Taylor development for approximating the objective function (see electronic supplementary material, S1, for details).

The two additional scoring functions, considered on the basis of the analytically obtained scores, were built in the following way. For vaccination, the additional scoring function *greedyV_infherds_threshold* consists of replacing $1_{I_j(t)=1}$ by $1_{0 < I_j(t) < 20}$ in *greedyV_infherds*. This intends to avoid that the first term of the sum in the scoring function becomes 0 for herds that have few infected animals but not necessarily just one. Similarly, for the treatment, we replaced *greedyT_infherds* by a scoring function with a softer condition on the number of infected animals. The condition in

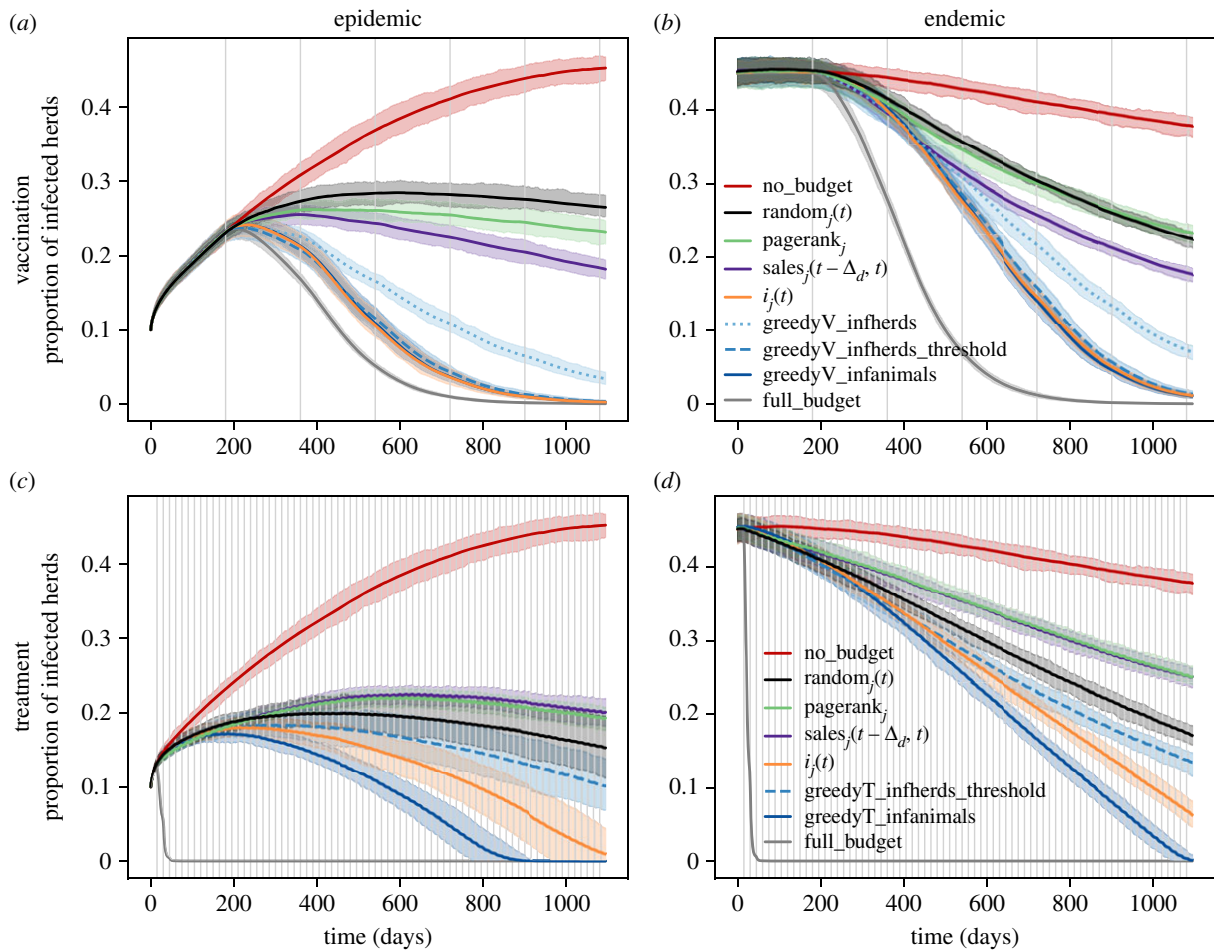


Figure 3. Dynamics of infection spread (proportion of infected herds) under resource allocation based on different scoring functions. Epidemic (left) and endemic (right) scenarios, for vaccination (*a,b*) and treatment (*c,d*). *no_budget* (red) represents the case where there was no resource allocated, and *full_budget* (grey) the case where the resource was not limited. For other curves, each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange) and greedy scoring functions (blue tones). For vaccination, the amount of available doses by decision time was $b_{\text{fix}} = 25\%$ of the initial total number of animals in the metapopulation. For treatment, the number of attainable herds by decision time was $b_{\text{fix}} = 25$ herds. Grey vertical lines represent resource allocation times. Mean results and 90% confidence bands over 50 runs.

greedyT_inferds on having exactly one infected animal for a herd to be eligible appeared to be too restrictive. Indeed, if the quantity of available treatment exceeded the number of herds that satisfy this condition, the rest of the treatment would not be allocated to any herd. Yet, allocating the exceeding treatment to herds with more than one infected animal could only be beneficial for limiting disease spread and would satisfy the constraint on the quantity of available treatment. Hence, the greedyT_inferds_threshold scoring function considers all herds that are potentially eligible with a non-negligible probability, yet favours herds that have few infected animals ($1_{0 < I_j(t) < 20}$).

3.2. Results of numerical explorations

3.2.1. Infection-related dynamics following score-based resource allocation

Figure 3 presents the results for the dynamics of the proportion of infected herds, under the setting described in §2.4.1 for a subset of the scoring functions. In particular, since the results for the topological scores were very similar, we present only the results of the pagerank_{*j*} score. This was also the case for demographic scores, so we chose the sales over the decision period, $\text{sales}_j(t - \Delta_d, t)$, as the representative score for this group. For the epidemiological scores, we present only the

results for the best performing score, $i_j(t)$. In addition, we included as the best and worst reference cases, results for cases where there was sufficient resource for all herds (*full_budget*), and where there was no resource to allocate (*no_budget*). Electronic supplementary material, figure S2, presents the complete results for cases by scenario (epidemic or endemic), type of available resource (vaccine or treatment) and the heuristic or greedy score according to which the allocation was performed. We also included in electronic supplementary material, figure S2, results for the dynamics of the total number of infected animals, yet we remark they were similar to the ones found for the proportion of infected herds.

One of the main remarks emerging from the analysis of figure 3 and electronic supplementary material, figure S2, is that topological and demographic scoring functions were the groups that performed less well for limiting the disease spread, both for vaccination and for the treatment. This is in comparison with the group formed by the scoring functions that take into account the health statuses of the herds, i.e. the greedy scoring functions and the best performing epidemiological functions, in particular $i_j(t)$.

Furthermore, electronic supplementary material, figure S2, shows that for vaccination, the scoring function $s_j(t)$ (the proportion of susceptible animals) performed the worst for

allocating the available quantity of vaccines (equal to 25% of the initial number of animals in the metapopulation) in both scenarios. The group of topological scoring functions performed better than the random scoring function in the epidemic scenario, but was not very different from it in the endemic scenario. Demographic scoring functions performed in general better than the topological ones, in particular $\text{sales}_j(0, t)$ and $\text{sales}_j(t - \Delta_{dr}, t)$. Apart from $s_j(t)$, the epidemiological scoring functions had a good performance, except for the proportion of recovered animals ($r_j(t)$) in the endemic scenario, where it was just as good as the random score.

The second important remark (figure 3) is that the best performing scores were the proportion of infected animals ($i_j(t)$), and greedyV_infanimals. Electronic supplementary material, figure S2 shows that this applies both for reducing the total number of infected animals and the proportion of infected herds. They were closely followed by greedyV_infherds_threshold in both scenarios. Meanwhile, greedyV_infherds was not as good as these three, particularly in the endemic scenario.

Figure 3 and electronic supplementary material, figure S2, also present results when the resource was a treatment and b_{fix} was equal to 25 herds. In particular, electronic supplementary material, figure S2, shows that in the epidemic scenario, $s_j(t)$ performed badly when compared to the other scoring functions. Yet $r_j(t)$ arrived to perform worse at the end of the 3 years in this scenario, and was the worst-performing scoring function in the endemic scenario. Even more, these two scoring functions, the topological and demographic ones, and the difference in the proportion of recovered animals ($r_j(t - \Delta_{dr}, t)$), all performed worse than the random score in both scenarios. Regarding the other scoring functions, greedyT_infanimals (allocating according to the number of infected animals by herd, $I_j(t)$) had the best performance in both scenarios. It was followed by the score $i_j(t)$, i.e. the proportion of infected animals by herd, though this last one did not manage to eradicate the disease before the 3 years. The next best performance was given by $i_j(t - \Delta_{dr}, t)$, which was followed by greedyT_infherds_threshold. To sum up, for the treatment allocation, only the greedy scores and two epidemiological scores ($i_j(t)$ and $i_j(t - \Delta_{dr}, t)$) performed better than the random score. In particular, greedyT_infanimals was the only one that eradicated the disease within the 3 years (figure 3).

The dynamics of raw proxies of costs can be found in electronic supplementary material, figure S3. In particular, it is shown that the proportion of herds that were vaccinated varied according to the allocation scoring function. Indeed, topological and demographic scoring functions led to vaccinating slightly fewer herds than the epidemiological and the greedy scores (excluding $s_j(t)$). Unsurprisingly, $s_j(t)$ led to the highest proportion of vaccinated herds and did not waste any doses, i.e. only vaccinated herds without infected or recovered animals, while $r_j(t)$ wasted the highest proportion of available doses. For the treatment, the topological and demographic scoring functions led to the smallest number of used doses and to the highest size of target population, contrary to the epidemiological and greedy scores.

Finally, electronic supplementary material, figure S4, shows the relationship between allocation decisions among different scoring functions at a given decision time. Both for vaccination and treatment, decisions according to the topological and demographic indicators were very similar (in terms of targeted herds) according to the Jaccard index [34]. For vaccination, decisions according to epidemiological and

greedy scoring functions were similar at the first decision time (six months), yet this similarity diminished over time. Regarding decisions through time for a given scoring function (electronic supplementary material, figure S5), topological functions tended to allocate the resource to the same herds over time. This was also the case for demographic functions, except for the sales scoring functions in vaccination, for which the first decisions were less and less similar to decisions at the following decision times. On the contrary, for $s_j(t)$ the similarity between consecutive vaccination decisions seemed to increase over time. And for each of the other epidemiological and greedy scoring functions, vaccination decisions were in general less similar over time. Treatment decisions according to each epidemiological and greedy scoring function were very different over time as long as the disease was not eradicated, except for decisions according to $r_j(t)$.

3.2.2. Percolation analysis results

Figure 4 shows results of the percolation analysis in the endemic scenario for each type of resource, using a selected subset of scoring functions. Results in the epidemic scenario can be found in electronic supplementary material, figure S6. We ran this analysis using all the epidemiological and greedy scores, since the infection-related dynamics results in §3.2.1 were quite different for the scores within each of these groups. By contrast, because the results of the topological scores were very similar, as were the results of the demographic scores, we considered only one of each type: $\text{sales}_j(t - \Delta_{dr}, t)$ for the demographic scoring functions and pagerank_j for the topological scores. This figure (also figure 4) confirms the main observations made in §3.2.1: certainly, the best-performing scoring functions for reducing disease prevalence, for almost every quantity of resource that we tested, were the greedy scores along with some epidemiological scores.

For vaccination, figure 4a shows that when the proportion of vaccinated herds was at least 0.10 of the initial number of animals in the metapopulation, the scoring functions that led to the lowest final proportion of infected herds were greedyV_infanimals, greedyV_infherds_threshold and $i_j(t)$. They were followed by the greedyV_infherds scoring function. The topological and demographic scoring functions, pagerank_j and $\text{sales}_j(t - \Delta_{dr}, t)$, were just as good as the greedy scores only when the available quantity of resource was very low (particularly when the proportion of vaccinated herds was less than 0.05), but did not perform well as this proportion was higher. The $r_j(t - \Delta_{dr}, t)$ and $i_j(t - \Delta_{dr}, t)$ scoring functions had a bad performance for such low levels of vaccinated herds, performed better for medium levels of vaccination, and performed worse when the proportion of vaccinated herds was higher than 0.2. In particular, $i_j(t - \Delta_{dr}, t)$ was better than greedyV_infherds when the proportion of vaccinated herds was less than 0.2. On the contrary, as the proportion of vaccinated herds was higher, $r_j(t)$ performs better. Indeed, when less than 25% of herds were vaccinated, it was just as bad as the random scoring function, which was the second worse score after $s_j(t)$. But for a proportion of vaccinated herds higher than 0.45, its performance was close to the one of the best-performing greedy scores. Even when this proportion was more than 0.6, only $r_j(t)$, $i_j(t)$ and the greedy scores arrived to eradicate the disease.

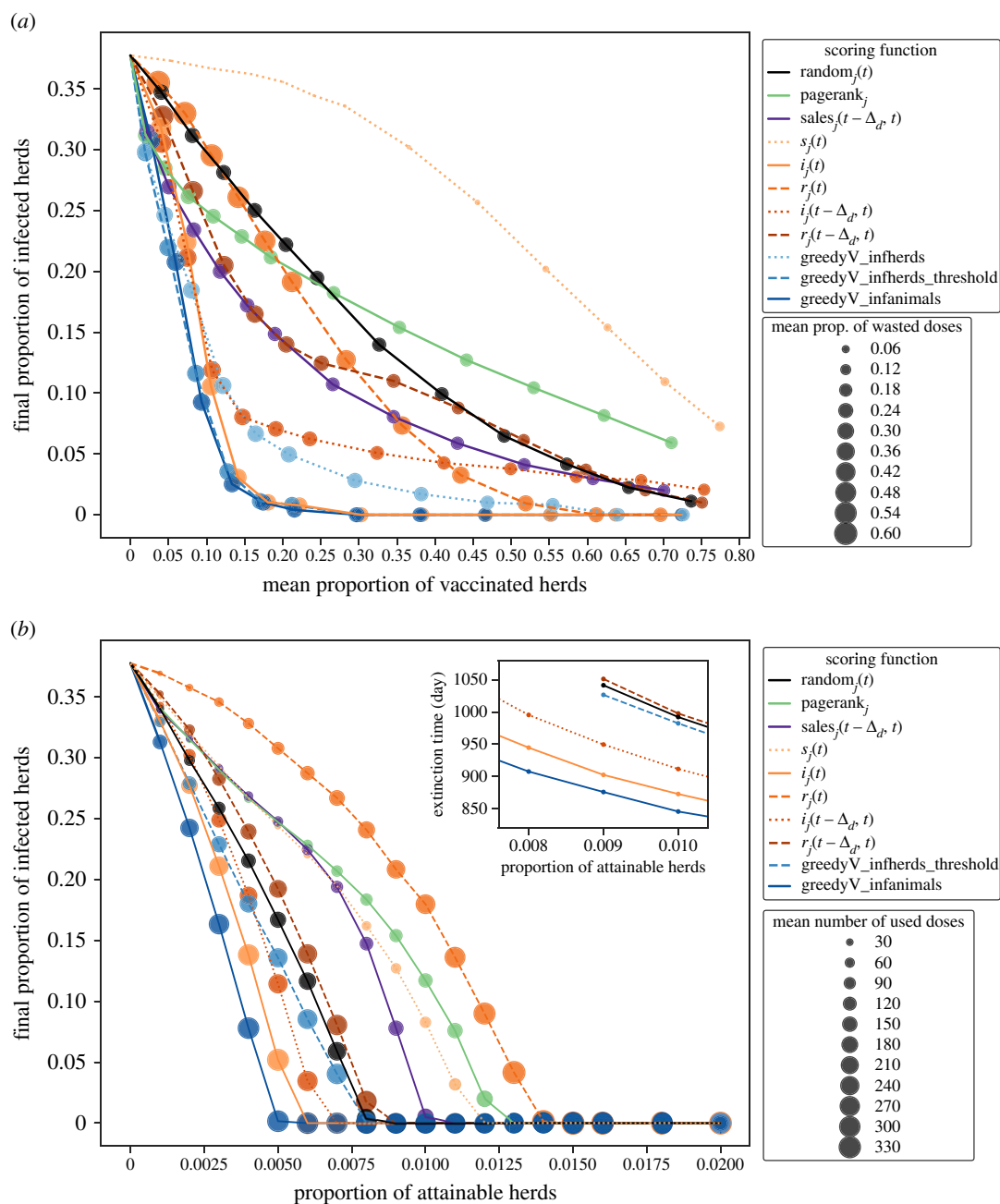


Figure 4. Percolation experiments results in the endemic scenario. Each colour represents the scoring function according to which the allocation decisions were taken: random (black), topological (green), demographic (purple), epidemiological (orange tones) and greedy scoring functions (blue tones). For (a) vaccination, final (after 3 years) mean (over runs) proportion of infected herds as a function of the mean (over time and over runs) proportion of vaccinated herds. Each point corresponds to a value of b_{fix} , and its size represents the mean (over time and over runs) proportion of wasted doses. For (b) treatment, final (after 3 years) mean (over runs) proportion of infected herds as a function of the proportion of attainable herds at each decision time. Each point corresponds to a value of b_{fix} , and its size represents the mean (over time and over runs) number of used doses. The inset shows the mean (over runs) extinction time with each scoring function, for a restricted range of the proportion of attainable herds. Results for 50 runs.

Regarding the mean proportion of wasted doses (over time and over runs), it was almost zero when the vaccine allocation was done using $s_j(t)$ as criterion, irrespective of the proportion of vaccinated herds. The other scoring functions led to quite similar levels of wasted doses, except for $r_j(t)$ and $r_j(t - \Delta_d, t)$, which led to the highest proportion of wasted doses, particularly when the proportion of vaccinated herds was lower than 0.15.

In figure 4b, we observe that when the resource to allocate was a treatment, $\text{greedyT_inffanimals}$ managed to eradicate the disease when it was possible to treat at least 25 herds per decision time-step, i.e. 0.5% of the total number of herds. As for $i_j(t)$, it could eradicate the disease when this percentage was higher than 0.6%, and $i_j(t - \Delta_d, t)$ when it

was at least 0.7% herds. When b_{fix} was equal to 0.9% of the total number of herds, $\text{greedyT_inffherds_threshold}$, the random scoring function and $r_j(t - \Delta_d, t)$ also eradicated the disease, in that order. The other epidemiological scoring functions and the topological and demographic ones performed worse than the random score for all values of b_{fix} . In particular, $\text{sales}_j(t - \Delta_d, t)$ only eradicated the disease if b_{fix} was higher than 1% of the total number of herds, $s_j(t)$ could only do it when this percentage was higher than 1.2%, and the other two scoring functions ($r_j(t)$ and pagerank_j) eradicated it when it was possible to treat at least 1.3% and 1.4% of the total number of herds, respectively.

As for the number of used doses of treatment, they were in general lower for bad-performing scoring functions, and

higher for those that performed the best. The exception was $r_j(t)$, which performed badly and led to a high number of used doses. When b_{fix} was high, it used even more doses on average than the best scoring functions for reducing disease prevalence.

Results in the epidemic scenario (electronic supplementary material, figure S6) were quite similar. Yet, for the treatment, only greedyT_inanimals, $i_j(t)$ and $i_j(t - \Delta_d, t)$ were always better than the random score. The greedyT_inherds_threshold indicator was better than the random score only when b_{fix} was less than 0.8% of the total number of herds. Additionally, electronic supplementary material, figure S7, shows that the variability of the percolation results were moderate between runs, with no or little overlapping 90% confidence intervals.

In addition, electronic supplementary material, figure S8, shows that when vaccine efficacy is decreased, the loss of performance is not the same for all scores. In particular, the scores that are not the best but perform rather well (epidemiological scores and greedyV_inherds) are the ones for which performance is most depreciated. This results in $\text{sales}_j(t - \Delta_d, t)$ performing better than $i_j(t - \Delta_d, t)$ when the average proportion of vaccinated herds was 55% and vaccine efficacy was 0.9 or less. Yet, despite the overall loss in efficiency, two of the greedy scores and one epidemiological score ($i_j(t)$) still performed the best.

Finally, electronic supplementary material, figures S9 and S10, show that a disease with higher early peak and smaller infection duration ($\beta/\gamma = 4$, $1/\gamma = 30$ days) spreads and fades out rapidly at the intra-herd level. At the same time, recurrent outbreaks are observed which are likely due to reintroduction of infected animals by trade. Electronic supplementary material, figure S11, shows that, in such a case, allocating vaccines according to the historic sales, in particular according to $\text{sales}_j(t - \Delta_d, t)$, was among the best strategies in both the epidemic and endemic scenarios.

4. Discussion

To control an infectious disease that spreads in a metapopulation network, allocating a limited resource is a fundamental yet difficult question, especially for large networks. In this study, we considered this resource allocation problem for a livestock disease that spreads over a large animal trade network, where the intra-herd infection and demographic dynamics was specified as an SIR stochastic model taking into account animal movements and demography.

The problem of resource allocation in networks had been previously addressed from several perspectives such as optimal control [23] and reinforcement learning [35], but mostly for networks where each node is an individual [10], or where the network is rather small [36]. Yet, in the context of a very large network these methods lack scalability for tracking the optimal solution [30]. In this work, we chose to concentrate on strategies based on scoring functions, heuristic and optimized, which consist of ordering the nodes of the network according to their score and allocating the resource to the top of the ranking, up to the limit given by the available resource.

First, following the greedy approach in [30], we provided new analytic scoring functions for controlling the disease spread over the animal metapopulation network by

optimizing approximated objective functions. The scoring functions we derived depend on the infection-related state of the herd, and some are also dependent on the topology of the metapopulation network. They differ according to the objective of the control (minimizing the number of infected animals versus minimizing the number of infected herds) and the type of available resource (a protective vaccine or a treatment that reduces the infectious period). Meanwhile, most similar existing approaches for other population structures derive strategies solely for distributing a vaccine [5,11], or are concerned with only one objective to be optimized [10,13].

Through intensive simulations, we observed that these analytically obtained scoring functions can be optimal for reducing disease prevalence in the metapopulation, though this is not always the case. For example, even if greedyV_inanimals (the score for greedily minimizing a function on the total number of infected animals in the metapopulation) showed the greatest reduction in disease prevalence through vaccination, allocating vaccines according to the proportion of infected animals by herd, $i_j(t)$, can be just as good when the number of available vaccine doses equalled 25% of the initial number of animals in the metapopulation (figure 3). Even more, we observed that this was the case as long as the number of available doses was more than 15% of the initial number of animals in the metapopulation (figure 4a).

For the treatment, most of the scoring functions, in particular topological and demographic ones, were counterproductive in the sense that they performed worse than randomly allocating the resource among the infected herds (figures 3 and 4b). We explain this by the fact that the infected herds which were central in the network were not the most infected ones (in terms of the proportion of infected animals). Indeed, electronic supplementary material, figure S3b, shows that the random allocation among infected herds also targeted herds with many infected animals, while the scoring functions that performed badly only targeted high sized infected herds but generally with few infected animals.

Furthermore, we noticed that irrespective of the resource type, the optimized scoring function for an objective on the number of infected herds was outperformed by the optimized scoring function for an objective on the number of infected animals. Even if for vaccination, a slightly modified version of greedyV_inherds provided results almost as good as the ones of greedyV_inanimals (figures 3 and 4a), this was not the case for the treatment. Indeed, the allocation implemented using the greedyT_inanimals scoring function, i.e. the number of infected animals per herd, yielded undoubtedly the best results (figures 3 and 4b). This is probably due to the fact that the scoring functions for minimizing an objective on the number of infected herds only focus on the fast recovery of slightly infected herds ($0 < I_j(t) < 20$), for vaccination, or on avoiding that completely healthy herds receive infectious animals ($I_j(t) > 0$), for the treatment. Although this is the best way to have a small incremental number of infected herds from one instant to another according to these scoring functions, it does not take into account new animal infections, which only occur at the intra-herd level once the herd is infected. Our interpretation is that a scoring function obtained with the greedy approach (which consists of focusing on the short-term behaviour of the objective function) performs better for limiting the disease spread if the objective function it is built on directly captures the intra-herd aspect of the disease dynamics. Hence,

minimizing a first-order approximation for an objective that directly concerns the number of infected animals can provide more performing scoring functions than that obtained by minimizing a higher-order approximation for an objective that does not.

Numerical investigations also allowed evidencing that intra-herd health information can be crucial for optimally controlling the propagation of a slowly spreading disease such as the one considered in this work. In most combinations, given by the resource's type and available quantity, even if topology-based scoring functions managed to limit disease spread, scoring functions based on the infection-related state of herds performed better (figure 4; electronic supplementary material, figure S6). This observation can give some insight into why most control strategies implemented in real systems might fail to eradicate livestock diseases in areas that lack this kind of information.

A final interesting remark is that the best scoring functions for reducing disease prevalence can induce a higher number of wasted vaccine doses or a higher number of used treatment doses compared to other scores that performed less well for controlling the disease spread (figure 4; electronic supplementary material, figures S3 and S6). However, for vaccination, they are not necessarily those that vaccinate the highest proportion of herds.

To our knowledge, our work is one of the few studies that explores dynamic resource allocation in a metapopulation network for many allocation scoring functions (16 heuristics and at least 2 optimized scores by measure), while varying the available quantity of the resource in two different scenarios. Despite the fact that the performances of the scores could have been different if the network had been static or not scale-free, score-based resource allocation can be a relevant approach for controlling pathogen spread in other cases, as the complexity of the problem is mainly due to the large dimension of the network. Furthermore, we stress that the scores we found by optimization using the greedy approach are the same irrespective of the network topology, as they are based on fixed trade parameters representing the static aggregated network. Indeed, they are built on a static view of the network, irrespective of whether it is actually static or dynamic (as it is in our study). We believe it would be possible to use the same approach to obtain optimized scores that take into account the dynamic nature of the network, i.e. scores that are function of time-dependent trade parameters, although this requires a new formal analysis. Of course, these new scores could be of a different form from the ones we found, and hence their performance might also be different.

We assessed the robustness of the results regarding vaccine allocation by considering realistic values for the vaccine efficacy. This showed a limited impact on the relative performance of the different scoring functions (electronic supplementary material, figure S8). We note that although in reality vaccines are rarely perfectly effective, and also take some time to be effective, it did not seem straightforward to determine an appropriate time frame for the vaccine to have an effect. More importantly, it seems unlikely that a slight delay in the effect of the vaccine would have a significant impact on our results, given that we considered a pathogen that spreads rather slowly.

Regarding the limitations of our work, we emphasize that in the context of a fast-spreading pathogen, the current framework

is not really appropriate. In such a case, other decision factors should be taken into account. For example, if the disease is zoonotic or has a strong economic impact, the social planner may consider more radical options, such as mass culling. In this case, the question of resource allocation thus becomes irrelevant. In particular, we showed that for a disease with higher early peak and shorter infection duration ($\beta/\gamma=4$, $1/\gamma=30$ days) an intra-herd epidemic extinguishes before a new disease introduction occurs (electronic supplementary material, figures S9 and S10). Therefore, it is not surprising that vaccinating herds that sell many animals appears as a good strategy for limiting disease propagation (electronic supplementary material, figure S11). So, the resource allocation problem seems more straightforward in such a scenario and does not necessarily require an optimization procedure.

Additionally, a parameter that could impact our conclusions is the decision step, Δ_{d_i} , for which we considered a fixed heuristic value. Although the assumption of regular vaccination decisions defined by the duration of the protection conferred by the vaccine appears to be a realistic hypothesis relative to field practice, a more versatile assumption could be considered to determine the frequency of allocation decisions. Indeed, the decision step could be determined in an adaptive manner by the social planner, for example by taking into account the stability in disease prevalence, or some external input such as the farmers' demand for accelerating resource allocation. A second option would be to determine the decision step by optimization. Yet, these are essentially different problems from the one we addressed in this article: determining when to allocate instead of where to allocate a limited resource. Optimizing both aspects at the same time is a more complex problem that, to the best of our knowledge, has only been addressed by heuristic approaches [37]. In particular, it does not seem straightforward to address with the approach of this study.

Finally, we stress that the performance of the epidemiological and the greedy scoring functions can be counterbalanced by their difficulty of access. Indeed, having updated knowledge on the epidemiological state of all the herds of the network is a strong hypothesis in real life, as this kind of information can be hard to gather for most livestock diseases [38]. For example, a scoring function calculated as the increment in the proportion of infected animals in a herd over a certain period can be observed through changes in the herd's seroprevalence between two time points, which incurs into increased logistics, can be observed with error and not in real time. Furthermore, having such updated and detailed health-related information can be costly, and this cost should be taken into account in the constrained optimization problem for the allocation. Among the possible perspectives of this work, the previous point opens an important one: combining scoring functions for improved performances, and above all for yielding a scoring function that can be useful in practice. This could be achieved, for example, through the (linear) combination of scoring functions, or through the selection of herds at the top of the ranking given by several scoring functions that do not allocate the resource in a similar way. Additionally, for cases when the value of the score is the same for many herds, the allocation could be done using a second scoring function that would take different allocation decisions. As a second, more methodological perspective, the greedy scoring functions built on first-order approximations could be eventually constructed using higher-order approximations. This could lead to analytic

scores that also depend on the network topology, and could improve their performance. Although it has to be stressed that they were already among the best-performing scores for reducing disease prevalence in all cases. Finally, despite the fact that we focused on a protective vaccine and a treatment that increases the recovery rate, other types of resources could be studied with the same approach. However, it might not be straightforward to derive the analytic expression of the greedy scoring function in such cases. For example, the effect of the restriction of animal movements, which is a relevant control measure in this context, lies on the connections of the herd rather than on the intra-herd level, which could further complicate the derivation of the scoring function.

Data accessibility. Additional figures, details on the greedy approach and on data generation are provided in electronic supplementary

material [39]. The Python simulation code for generating data and figures is available at <https://github.com/CristanchoLina/DRAAnimalMetapop>.

Authors' contributions. L.C.: conceptualization, data curation, formal analysis, investigation, methodology, software, visualization, writing—original draft; P.E.: conceptualization, methodology, project administration, supervision, writing—review and editing; E.V.: conceptualization, funding acquisition, methodology, project administration, supervision, writing—review and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. This work was carried out with the financial support of the French Research Agency (ANR), project ANR-16-CE32-0007 (CADENCE).

Acknowledgements. We thank the French minister of agriculture for providing access to the FCID.

References

- Tomley FM, Shirley MW. 2009 Livestock infectious diseases and zoonoses. *Phil. Trans. R. Soc. B* **364**, 2637–2642. (doi:10.1098/rstb.2009.0133)
- Heesterbeek H *et al.* 2015 Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347**, aaa4339. (doi:10.1126/science.aaa4339)
- Preciado VM, Zargham M, Enyioha C, Jadbabaie A, Pappas G. 2013 Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks. In *52nd IEEE Conf. on Decision and Control, Firenze, Italy, 10–13 December 2013*, pp. 7486–7491. (doi:10.1109/CDC.2013.6761078)
- Holme P, Litvak N. 2017 Cost-efficient vaccination protocols for network epidemiology. *PLoS Comput. Biol.* **13**, e1005696. (doi:10.1371/journal.pcbi.1005696)
- Tanner MW, Sattenspiel L, Ntairo L. 2008 Finding optimal vaccination strategies under parameter uncertainty using stochastic programming. *Math. Biosci.* **215**, 144–151. (doi:10.1016/j.mbs.2008.07.006)
- Enayati S, Özalpin OY. 2020 Optimal influenza vaccine distribution with equity. *Eur. J. Oper. Res.* **283**, 714–725. (doi:10.1016/j.ejor.2019.11.025)
- Thompson RN, Gilligan CA, Cuniffe NJ. 2018 Control fast or control smart: when should invading pathogens be controlled? *PLoS Comput. Biol.* **14**, e1006014. (doi:10.1371/journal.pcbi.1006014)
- Zhang Y, Aditya Prakash B. 2015 Data-aware vaccine allocation over large networks. *ACM Trans. Knowl. Discov. Data (TKDD)* **10**, 1–32 (doi:10.1145/2803176)
- Lorch L, De A, Bhatt S, Trouleau W, Upadhyay U, Gomez-Rodriguez M. 2018 Stochastic optimal control of epidemic processes in networks. (<https://arxiv.org/abs/1810.13043>)
- Wijayanto AW, Murata T. 2019 Effective and scalable methods for graph protection strategies against epidemics on dynamic networks. *Appl. Netw. Sci.* **4**, 1–31. (doi:10.1007/s41109-019-0122-7)
- Zhan J, Rafalski T, Stashkevich G, Verenich E. 2017 Vaccination allocation in large dynamic networks. *J. Big Data* **4**, 1–17. (doi:10.1186/s40537-016-0062-3)
- Scaman K, Kalogeratos A, Vayatis N. 2016 Suppressing epidemics in networks using priority planning. *IEEE Trans. Netw. Sci. Eng.* **3**, 271–285. (doi:10.1109/TNSE.2016.2600029)
- Forster GA, Gilligan CA. 2007 Optimizing the control of disease infestations at the landscape scale. *Proc. Natl Acad. Sci. USA* **104**, 4984–4989. (doi:10.1073/pnas.0607900104)
- Lenhart S, Workman JT. 2007 *Optimal control applied to biological models*. Boca Raton, FL: CRC Press.
- Sutton RS, Barto AG. 2018 *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.
- Dvijotham K, Todorov E. 2013 Linearly solvable optimal control. In *Reinforcement learning and approximate dynamic programming for feedback control*, vol. 17, pp. 119–141. Hoboken, NJ: Wiley.
- Andriotis CP, Papakonstantinou KG. 2019 Managing engineering systems with large state and action spaces through deep reinforcement learning. *Reliab. Eng. Syst. Saf.* **191**, 106483. (doi:10.1016/j.res.2019.04.036)
- Ndeffo Mbah ML, Gilligan CA. 2011 Resource allocation for epidemic control in metapopulations. *PLoS ONE* **6**, e24577. (doi:10.1371/journal.pone.0024577)
- Matrajt L, Halloran ME, Longini Jr IM. 2013 Optimal vaccine allocation for the early mitigation of pandemic influenza. *PLoS Comput. Biol.* **9**, e1002964. (doi:10.1371/journal.pcbi.1002964)
- Long EF, Nohdurft E, Spinler S. 2018 Spatial resource allocation for emerging epidemics: a comparison of greedy, myopic, and dynamic policies. *Manuf. Serv. Oper. Manage.* **20**, 181–198. (doi:10.1287/msom.2017.0681)
- Rao IJ, Brandeau ML. 2021 Optimal allocation of limited vaccine to control an infectious disease: simple analytical conditions. *Math. Biosci.* **337**, 108621. (doi:10.1016/j.mbs.2021.108621)
- Carli R, Cavone G, Epicoco N, Scarabaggio P, Dotoli M. 2020 Model predictive control to mitigate the COVID-19 outbreak in a multi-region scenario. *Annu. Rev. Control* **50**, 373–393. (doi:10.1016/j.arcontrol.2020.09.005)
- Bussell EH, Dangerfield CE, Gilligan CA, Cuniffe NJ. 2019 Applying optimal control theory to complex epidemiological models to inform real-world disease management. *Phil. Trans. R. Soc. B* **374**, 20180284. (doi:10.1098/rstb.2018.0284)
- Chernov AA, Kelbert MY, Shemendyuk AA. 2020 Optimal vaccine allocation during the mumps outbreak in two SIR centres. *Math. Med. Biol.: J. IMA* **37**, 303–312. (doi:10.1093/imammb/dqz012)
- Duijzer LE, van Jaarsveld WL, Wallinga J, Dekker R. 2018 Dose-optimal vaccine allocation over multiple populations. *Prod. Oper. Manage.* **27**, 143–159. (doi:10.1111/poms.2018.27.issue-1)
- Lemaitre JC, Pasetto D, Zanon M, Bertuzzo E, Mari L, Miccoli S, Casagrandi R, Gatto M, Rinaldo A. 2021 Optimizing the spatio-temporal allocation of COVID-19 vaccines: Italy as a case study. *medRxiv*. (doi:10.1101/2021.05.06.21256732)
- Rautureau S, Dufour B, Durand B. 2011 Vulnerability of animal trade networks to the spread of infectious diseases: a methodological approach applied to evaluation and emergency control strategies in cattle, France, 2005. *Transbound. Emerg. Dis.* **58**, 110–120. (doi:10.1111/jva.2011.58.issue-2)
- Brooks-Pollock E, de Jong MCM, Keeling MJ, Klinkenberg D, Wood JLN. 2015 Eight challenges in modelling infectious livestock diseases. *Epidemics* **10**, 1–5. (doi:10.1016/j.epidem.2014.08.005)
- Dangerfield CE, Vyska M, Gilligan CA. 2019 Resource allocation for epidemic control across multiple sub-populations. *Bull. Math. Biol.* **81**, 1731–1759. (doi:10.1007/s11538-019-00584-2)
- Scaman K, Kalogeratos A, Vayatis N. 2015 A greedy approach for dynamic control of diffusion processes in networks. In *2015 IEEE 27th Int. Conf. on Tools with Artificial Intelligence (ICTAI), Vietri sul Mare, Italy, 9–11 November 2015*, pp. 652–659. (doi:10.1109/ICTAI.2015.99)
- Cristancho Fajardo L, Ezanno P, Vergu E. 2021 Accounting for farmers' control decisions in a model

- of pathogen spread through animal trade. *Sci. Rep.* **11**, 1. (doi:10.1038/s41598-020-79139-8)
32. Bretó C, He D, Ionides EL, King AA. 2009 Time series analysis via mechanistic models. *Ann. Appl. Stat.* **3**, 319–348 (doi:10.1214/08-AOAS201)
33. Newman M. 2018 *Networks*. Oxford, UK: Oxford University Press.
34. Levandowsky M, Winter D. 1971 Distance between sets. *Nature* **234**, 34–35. (doi:10.1038/234034a0)
35. Probert WJM, Lakkur S, Fonnesbeck CJ, Shea K, Runge MC, Tildesley MJ, Ferrari MJ. 2019 Context matters: using reinforcement learning to develop human-readable, state-dependent outbreak response policies. *Phil. Trans. R. Soc. B* **374**, 20180277. (doi:10.1098/rstb.2018.0277)
36. Rowthorn RE, Laxminarayan R, Gilligan CA. 2009 Optimal control of epidemics in metapopulations. *J. R. Soc. Interface* **6**, 1135–1144. (doi:10.1098/rsif.2008.0402)
37. Venkatramanan S, Chen J, Fadikar A, Gupta S, Higdon D, Lewis B, Marathe M, Mortveit H, Vullikanti A. 2019 Optimizing spatial allocation of seasonal influenza vaccine under temporal constraints. *PLoS Comput. Biol.* **15**, e1007111. (doi:10.1371/journal.pcbi.1007111)
38. Carslake D, Grant W, Green LE, Cave J, Greaves J, Keeling M, McEldowney J, Weldegebriel H, Medley GF. 2011 Endemic cattle diseases: comparative epidemiology and governance. *Phil. Trans. R. Soc. B* **366**, 1975–1986. (doi:10.1098/rstb.2010.0396)
39. Cristancho-Fajardo L, Ezanno P, Vergu E. Dynamic resource allocation for controlling pathogen spread on a large metapopulation network. Figshare.