

Risques d'infection liés aux échanges d'animaux d'élevage : l'apport de la théorie des graphes

L'augmentation de la puissance de calcul et des volumes de données disponibles a entraîné un foisonnement des travaux de modélisation concernant des phénomènes de réseaux (échanges d'informations ou de denrées, par exemple). Cette note, principalement méthodologique, décrit les apports de la théorie des graphes à la compréhension des liens entre échanges commerciaux d'animaux d'élevage et propagation des épizooties.

La théorie des graphes permet l'étude d'un grand nombre de phénomènes en interaction (agents, objets, concepts, etc.), constituant des réseaux dynamiques (réseaux sociaux, réseaux pair-à-pair, réseaux d'échanges commerciaux, etc.). La propagation de maladies, en particulier les maladies animales, en est un des domaines d'étude courants.

La crise de l'encéphalopathie spongiforme bovine (ESB), dans les années 1980 et surtout 1990, a confirmé aux instances sanitaires l'importance d'un système efficace de traçabilité pour suivre la propagation d'une maladie. La base de données nationale d'identification (BDNI) enregistre, dans ce but, chaque bovin naissant sur le territoire français ou y étant importé, ainsi que ses changements d'exploitation agricole jusqu'à sa mort (abattage ou équarrissage). Au-delà de son objectif de traçabilité pour répondre aux situations de crise, l'enregistrement de ces échanges commerciaux aide à savoir comment les acteurs interagissent, et ainsi à mieux comprendre les mécanismes concourant à la diffusion des maladies. Les outils de la théorie des graphes permettent de traiter ces données de façon originale, en dévoilant des dynamiques sous-jacentes peu identifiables autrement.

La première partie de cette note présente les bases de la théorie des graphes et ses principes d'identification des acteurs et interactions clés d'un système. La seconde partie illustre comment ces apports peuvent être utilisés dans le domaine de la santé animale.

1. Des graphes statiques aux approches temporelles

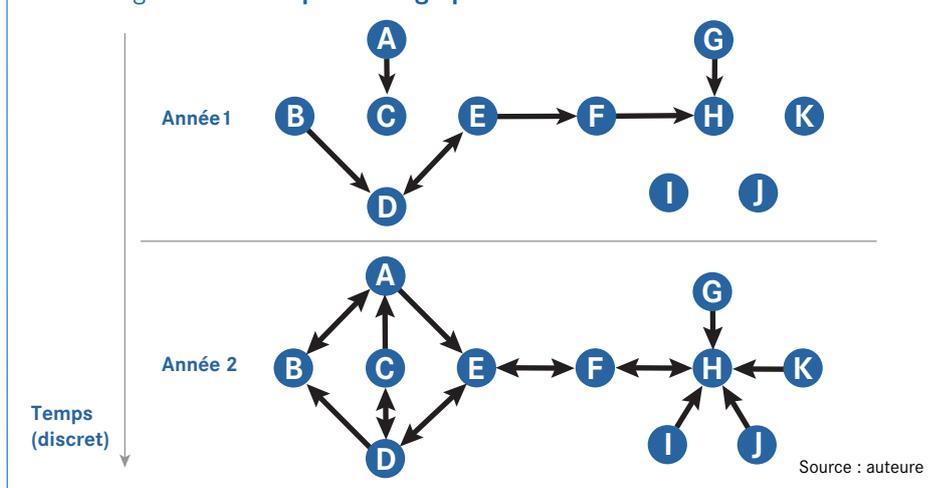
Graphes statique : une image synthétique des interactions

Un graphe est une représentation abstraite d'un réseau d'entités ou d'objets interagissants¹ : un ensemble de nœuds (les objets d'étude) sont reliés par des liens représentant leurs interactions. Cette conception, dite « statique », issue des mathématiques du XVIII^e siècle, ne conserve pas l'information temporelle : les liens ne sont pas datés.

Des représentations « temporelles » ont été développées ensuite², depuis une vingtaine d'années, pour étudier des phénomènes où la dynamique des interactions ne pouvait être négligée, comme c'est généralement le cas en épidémiologie.

Une première façon de prendre en compte la dimension temporelle fut d'introduire les « séquences de graphes » (figure 1) : un graphe statique est construit pour chaque intervalle de temps défini, comme une séquence de photographies. Illustrons cette approche avec la BDNI. En 2014, 182 452 élevages, marchés et centres de rassemblement ont échangé

Figure 1 - Une séquence de graphes annuels



1. De ce fait, le terme « réseau » sera privilégié dans cette note pour désigner les interactions entre agents, quand « graphe » renverra à l'objet mathématique théorique.

2. Holme P., Saramäki J., 2012, « Temporal networks », *Physics reports*, 519(3), pp 97-125

au moins un animal (par vente, achat, prêt, etc.). Ces exploitations sont dites « actives ». Le graphe de 2014 de la BDNI comporte donc 182 452 nœuds, représentant les exploitations. Lorsque deux exploitations ont interagi, c'est-à-dire échangé au moins un animal, un lien les relie dans le graphe. Dans le cas de la BDNI, les liens sont orientés, c'est-à-dire que le sens du lien indique le point de départ des animaux et leur point d'arrivée. En 2014, le graphe comporte 1 060 354 liens orientés, que l'on peut également considérer comme 1 060 354 couples de nœuds ayant interagi. En 2015, des échanges différents ont eu lieu entre des exploitations différentes ; le graphe correspondant comporte 176 771 nœuds et 1 037 465 liens. Le nombre de nœuds et de liens actifs a ainsi diminué, ce qui s'expliquerait³ par le phénomène de fusion des exploitations.

Une structure impactant la propagation de maladies

Différentes métriques et notions peuvent être étudiées et calculées, sur chaque graphe d'une séquence, pour comprendre l'organisation de ces graphes. Les centralités et la connexité en sont deux exemples.

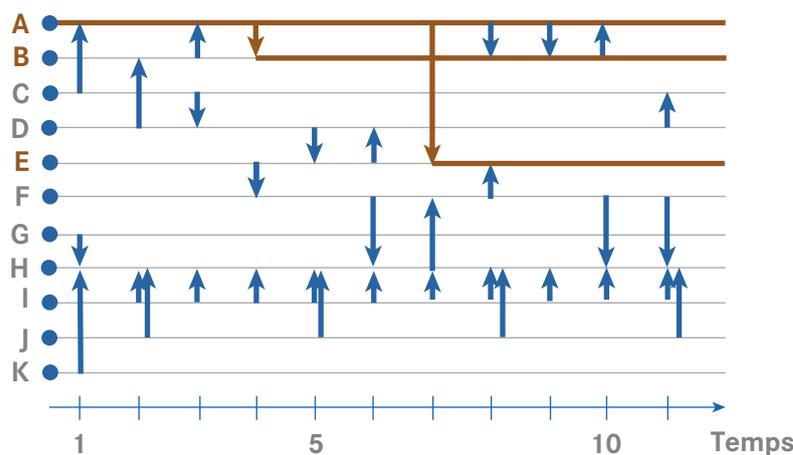
Les centralités sont une famille de mesures qui permettent de déterminer l'importance d'un nœud au sein du graphe. Par exemple, le degré se calcule en comptant le nombre de liens impliquant un nœud donné : le nœud A de la figure 1 a un degré 1 la première année et un degré 3 la seconde⁴. Cette métrique permet d'identifier les hubs du réseau. Dans la BDNI, les 100 nœuds de plus fort degré sont tous de type centre de rassemblement ou marché, ce qui confirme leur capacité de concentration des échanges.

Second exemple de notion, un graphe est dit « connexe » lorsqu'il existe une succession de liens permettant de relier chaque nœud à un autre, indépendamment de l'orientation des liens. Par exemple, le graphe de la figure 1 est connexe la seconde année. La première année, les nœuds I, J et K et le couple de nœuds A et C sont déconnectés de B, D, E, F, G et H, le graphe est donc non connexe. Cette notion permet d'approximer le potentiel de diffusion dans un système.

Des interactions, supports des modèles de diffusion

Un graphe est par nature un ensemble d'interactions entre paires de nœuds. Cette information peut servir à modéliser la propagation d'une maladie animale. L'étude des graphes et celle des phénomènes de diffusion sont donc très liées.

Figure 2 - Exemple de flot de liens



Lecture : On voit que C échange au moins un animal avec A au temps 1. Si A est infecté, il a une certaine probabilité de contaminer B et E. Si l'information temporelle est agrégée pour transformer ce flot de liens en un graphe statique, on retrouve le graphe de l'année 2 de la figure 1, où A pouvait potentiellement infecter 6 autres nœuds. Ignorer le temps implique ici une surestimation du risque de contamination du réseau par A.

Source : auteur

En quantifiant l'importance d'un nœud selon certains critères, les centralités permettent d'estimer quel risque ferait courir au système la contamination de nœuds donnés. Par exemple, un nœud de fort degré aura la capacité de transmettre une maladie à un grand nombre d'autres nœuds. Dès lors, les nœuds peuvent être classés selon leur score de centralité décroissant, pour obtenir une priorisation des nœuds selon leur niveau de risque. Ces nœuds seront par exemple à cibler en priorité lors de campagnes de contrôle ou de vaccination. Par ailleurs, plus il existe un grand nombre de liens reliant les nœuds les uns aux autres, plus le graphe sera sensible à la propagation d'une maladie. Dans la figure 1, si A est infecté la première année, seul C risque d'être contaminé à son tour, alors que la seconde année, A peut infecter 6 autres nœuds. Nous verrons ci-dessous que la dimension temporelle est importante pour estimer ce potentiel d'infection.

Un graphe est un support de la modélisation de la propagation d'une maladie : à partir d'un nœud ou d'un groupe de nœuds infectés, on peut simuler une contamination de proche en proche (de lien en lien), selon un modèle de maladie donné. Une façon de procéder est de recourir à un modèle à compartiments, où les nœuds sont catégorisés selon leur statut sanitaire. Par exemple, le « modèle susceptible-infecté » distingue les nœuds sains des nœuds infectés, ces derniers pouvant contaminer les premiers en cas de contact (lien) selon une probabilité donnée, et sans guérison possible.

Des usages orientant le choix d'une représentation

La dimension temporelle n'est prise en compte que partiellement dans une « séquence de graphes » : au sein d'une fenêtre temporelle et du graphe correspondant, l'ordre des interactions est perdu. Par exemple, dans la figure 1, on ne sait pas si B et D ont interagi avant D et E la première année. Or, cette information est essentielle pour modéliser la propagation d'une maladie, car elle indique quels nœuds peuvent avoir été en contact direct ou indirect avec un nœud malade, selon sa date d'infection et leur date d'interaction (exemple figure 2). Utiliser une séquence de graphes entraîne ainsi la surestimation de l'ampleur des épizooties⁵, en comptabilisant des nœuds qui ont interagi avec les nœuds contaminés avant leur infection. Les séquences de graphes

3. Dutta B. L., Ezanno P. et Vergu, E., 2014, « Characteristics of the spatio-temporal network of cattle movements in France over a 5-year period », *Preventive veterinary medicine*, 117(1), 79-94.

4. Dans le cas d'un graphe orienté, on peut distinguer le degré entrant et le degré sortant, c'est-à-dire le nombre de liens entrant dans un nœud et le nombre de liens en sortant, respectivement. Dans l'exemple de A l'année 2 (figure 1), A a un degré entrant de 2 et un degré sortant de 2 également.

5. Dubé C., Ribble C., Kelton D. et 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 », *Transboundary and emerging diseases*, 55(9-10), 382-392.

devraient donc être utilisées lorsque cette approximation a des conséquences limitées, ou lorsqu'il existe un pas de temps caractéristique dans les échanges. Par exemple, pour les échanges commerciaux de denrées alimentaires, observer les échanges par mois ou même par année fait sens.

Les représentations qui intègrent l'ensemble de l'information temporelle s'étant développées récemment, les métriques, par exemple les centralités, n'y sont pas toutes définies, et lorsque c'est le cas, elles ne le sont pas toujours de manière consensuelle. De ce fait, les séquences de graphes sont encore souvent utilisées en épidémiologie, malgré leurs limites, pour bénéficier de la riche littérature scientifique disponible et consolidée des graphes statiques. De plus, ajouter une dimension, le temps, complique grandement les calculs, ce qui rend plus délicat le traitement de gros volumes de données.

Vers une représentation dynamique des échanges

Pour conserver la structure des interactions et leur ordre chronologique, les « flots de liens » ont été introduits⁶. Ils consistent en une liste d'interactions datées (figure 2). Avec cette représentation, la BDNI comporte toujours, en 2015, 176 771 nœuds actifs (ayant interagi au moins une fois), pour 1 037 465 couples de nœuds ayant interagi (liens au sens statique), mais elle compte 2 652 131 liens temporels (triplets date-couple-interagissant). Ainsi, la structure des flots de liens peut être étudiée (ex. quel acteur est important ?), tout comme leur temporalité (ex. les acteurs importants le demeurent-ils ?).

Dans ce cas également, la liste des interactions datées peut servir à simuler la propagation de maladies par l'utilisation d'un modèle à compartiments. Il faudra alors prendre en compte dans les calculs la date de l'interaction (figure 2).

2. Application en santé animale

Quelques exemples à différentes échelles

L'étude d'un réseau d'interactions nécessite de détenir des données les plus exhaustives possibles sur les interactions ayant lieu. Les exemples d'application concernent ainsi essentiellement des analyses d'événements ayant pu être enregistrés (comment s'organisent les échanges commerciaux entre pays, comment circule une information, une maladie ou tout autre flux, etc.). Si le système est relativement stable au cours du temps, les conclusions tirées pourront

être supposées valables dans l'ensemble, et servir à éclairer la mise en place d'actions. À noter que de nombreux chercheurs se penchent sur la question de la prédiction de liens⁷, qui pourrait permettre à terme d'anticiper les évolutions d'un réseau.

La transmission de maladies animales se fait par de nombreuses voies : contacts inter-troupeaux au pâturage, introduction d'un animal malade sur une exploitation, contacts avec une espèce sauvage hôte, utilisation sur une exploitation saine d'un matériel agricole contaminé, etc. En fonction des caractéristiques de la maladie, toutes ces voies ne jouent pas un rôle équivalent dans la diffusion. Étudier un réseau d'interactions peut aider à déterminer lesquelles sont les plus probables. Ainsi, en analysant les réseaux d'échanges des différents acteurs de la filière porcine (exploitations, fournisseurs, transporteurs), dans le but d'estimer les voies d'introduction probables de la diarrhée épidémique porcine au Canada, en 2014, une étude⁸ montre que le réseau de livraison des aliments sur les exploitations est associé au plus fort risque d'introduction de la maladie. Les échanges d'animaux (introduction d'un animal étranger) ou la contamination au cours de leur transport ne seraient pas des routes prépondérantes. À l'inverse, la propagation de la brucellose bovine en Italie se serait faite principalement par les échanges d'animaux⁹, et le contact entre différents troupeaux au pâturage aurait été un facteur aggravant. Les auteurs de cette étude estiment que leur méthode d'analyse pourrait servir aux acteurs de terrain pour évaluer le risque de persistance de la brucellose sur leur territoire.

Les approches réseaux sont également mobilisables pour étudier les contacts au sein d'une population, lorsque les données existent pour ce faire. Par exemple, l'utilisation de graphes statiques, un pour chaque type de contact entre bovins laitiers (reniflement, léchage, frottement), a montré¹⁰ que la grande majorité des animaux ont peu de contacts avec les autres, et qu'il y a quatre fois plus de chances qu'une vache ait des contacts avec un congénère avec lequel elle a déjà interagi par le passé, par rapport à un individu avec lequel elle n'a jamais interagi. Ce comportement aurait tendance à ralentir la propagation de maladies. Toutefois, quelques animaux entrent en contact avec beaucoup d'autres individus : ce sont notamment des vaches en chaleur et des veaux mâles. Ces animaux favoriseraient la propagation de maladies et seraient donc à surveiller et à traiter en priorité.

Analyser un réseau d'interactions permet aussi d'évaluer les impacts de modifications externes sur les interactions d'un système, comme par exemple l'évolution du cadre réglementaire. Par exemple, suite à l'épizootie de fièvre aphteuse de 2001, le Royaume-Uni a instauré une période d'immobilisation obligatoire des animaux introduits sur une exploitation, avant de pouvoir les déplacer à nouveau. L'objectif était de réduire les échanges entre exploitations et par suite le potentiel de propagation des maladies. Cependant, une étude¹¹ a montré, en utilisant une séquence de graphes hebdomadaires et en analysant leur connexité, que la vulnérabilité du secteur s'était au contraire accrue : le réseau s'est optimisé pour faciliter les échanges sous cette contrainte et, de ce fait, il a généré de plus nombreuses voies de propagation potentielles pour une maladie.

Les méthodes d'analyse des réseaux s'adaptent ainsi à de nombreux contextes et champs d'application. Cependant, les possibilités d'utilisation des représentations temporelles sont en pratique limitées par la nécessité d'implémenter soi-même les algorithmes que l'on veut tester, alors que des packages sont disponibles pour l'étude des « graphes statiques » dans les langages informatiques les plus courants (R, Python, etc.). À l'heure actuelle, la théorie des graphes est donc souvent limitée au domaine de la recherche. Elle permet l'analyse de phénomènes d'intérêt public mais pas, à ce stade, une utilisation opérationnelle de ces modèles.

6. Latapy M., Viard T. et Magnien C., 2018, « Stream graphs and link streams for the modeling of interactions over time », *Social Network Analysis and Mining*, 8(1), 61.

7. Arnoux T., Tabourier L. et Latapy M., 2018, « Predicting interactions between individuals with structural and dynamical information » : arXiv preprint arXiv:1804.01465.

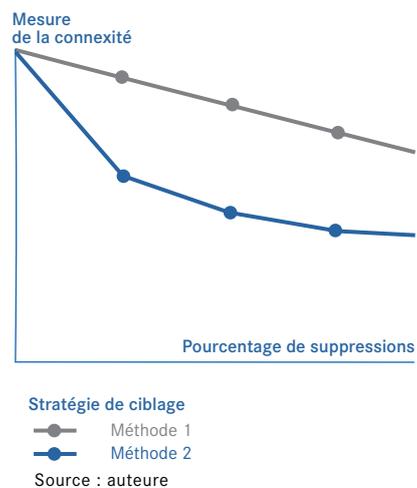
8. Perri A. M., Poljak Z., Dewey C., Harding J. C. et O'Sullivan T. L., 2019, « Network analyses using case-control data to describe and characterize the initial 2014 incursion of porcine epidemic diarrhea (PED) in Canadian swine herds », *Preventive veterinary medicine*, 162, 18-28.

9. Darbon A., Valdano E., Poletto C., Giovannini A., Savini L., Candeloro L. et Colizza V., 2018, « Network-based assessment of the vulnerability of Italian regions to bovine brucellosis », *Preventive veterinary medicine*, 158, 25-34.

10. Freslon I. de, Martínez-López B., Belkhiria J., Strappini A. et Monti G., 2019, « Use of social network analysis to improve the understanding of social behaviour in dairy cattle and its impact on disease transmission », *Applied Animal Behaviour Science*, 213, 47-54.

11. Robinson S. E., Everett M. G. et Christley R. M., 2007, « Recent network evolution increases the potential for large epidemics in the British cattle population », *Journal of the Royal Society Interface*, 4(15), 669-674.

Figure 3 - Méthode de comparaison des stratégies de ciblage selon les scores de centralité décroissants des nœuds ou des liens

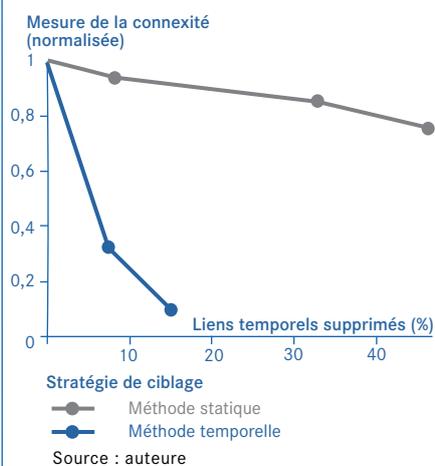


La priorisation des exploitations à traiter en cas de crise

En cas de crise sanitaire, on recourt parfois à un abattage préventif des animaux d'exploitations localisées dans une zone géographique proche du foyer infectieux. Une étude¹², qui modélise et compare cette stratégie de ciblage géographique avec des méthodes basées sur le concept des « centralités », a montré que ces dernières sont plus efficaces pour sélectionner les exploitations à cibler et ainsi lutter contre la propagation d'une épizootie. Elles permettent en effet une réduction plus efficace du potentiel d'infection des nœuds du réseau.

Tout un pan de la littérature sur les réseaux consiste à raffiner les notions de « centralité » pour les rendre plus adaptées au contexte étudié (ex : un système évoluant significativement au cours du temps), puis à comparer leur efficacité pour cibler les nœuds (élevages, centres de rassemblement, marchés, etc.) ou les liens (transferts d'animaux, etc.) les plus importants dans un contexte donné. Pour pouvoir les comparer, une méthode consiste à classer les nœuds selon leur score de centralité décroissant, puis à les retirer du réseau progressivement selon cet ordre (figure 3). La stratégie de ciblage la plus efficace est celle qui permet de retirer le plus faible nombre de nœuds ou de liens tout en déconnectant le plus fortement le réseau (ex : la méthode 2 de la figure 3). Autrement dit, si les nœuds ou les liens ciblés, une fois retirés, diminuent fortement le potentiel de propagation d'une maladie, alors ces éléments ont de plus fortes chances, que les autres éléments du réseau, de servir de point de passage pour une maladie, par rapport aux autres

Figure 4 - Comparaison de l'efficacité d'un ciblage temporel des voies de propagation (liens) par rapport à un ciblage statique, sur l'année 2015 de la BDNI



éléments du réseau, et ils devraient donc être ciblés préférentiellement par des mesures de sensibilisation, de prophylaxie, etc.

Comme abordé plus haut, de nombreux efforts sont fournis pour proposer des notions de centralité adaptées aux réseaux temporels (ex. pour les flots de liens⁶). La méthode de « comparaison des stratégies », présentée ci-dessus, peut aider à vérifier si le coût de calculs supplémentaires, dû à l'ajout de la dimension temporelle, est justifié par une efficacité accrue du ciblage des nœuds ou des liens à contrôler. Dans des travaux menés sur la BDNI¹³, le recours à une notion « d'importance temporelle », permise par l'utilisation de flots de liens pour représenter les échanges, permet d'améliorer significativement l'efficacité de l'identification des voies de propagation les plus à risque (*i.e.* détection des liens importants), par rapport aux méthodes d'identification statiques (figure 4). En effet, lorsque 10 % des liens les plus importants sont supprimés, le potentiel de propagation d'une maladie sur le réseau décroît de près de 80 %, alors que la stratégie de ciblage statique équivalente réduit ce potentiel de moins de 10 % pour une même quantité de liens temporels (*i.e.* triplet date et couple de nœuds en interaction) supprimés. Ainsi, utiliser cette stratégie temporelle permettrait d'optimiser le ciblage des contrôles ou des mesures de lutte, par rapport aux méthodes statiques plus couramment utilisées.

*

Le ciblage des voies de propagation (liens), selon une stratégie intégrant l'information temporelle, semble prometteur pour identifier le niveau de risque des interactions entre

exploitations agricoles, marchés et centres de rassemblement. Une mesure concrète pourrait consister à cibler les structures prenant part aux échanges à plus haut risque, lors de campagnes de sensibilisation, de prophylaxie ou de contrôle. Les niveaux de risque obtenus pourraient également être croisés avec des informations géographiques sur les exploitations, afin d'identifier les zones sur lesquelles les moyens de surveillance, de contrôle ou de lutte devraient se concentrer.

Avant de songer à mettre en œuvre de telles mesures, il convient de souligner que les réseaux ne sont jamais figés dans le temps : toute contrainte sur les échanges entraînera des modifications, qui nécessitent de revoir la pertinence de l'action en cours. Autrement dit, si certaines voies d'échange sont fermées, d'autres se créeront si aucune autre action n'est prise, annulant le bénéfice de la mesure mise en œuvre. C'est ce que montre par exemple l'étude sur l'immobilisation obligatoire des ovins au Royaume-Uni, pour lutter contre la fièvre aphteuse, présentée dans cette note. Étudier les possibles reconfigurations ou les moyens de les contrôler est donc indispensable avant de penser toute mise en œuvre sur le terrain.

La poursuite du développement des méthodes d'analyse des réseaux temporels et de l'étude de leur mise en pratique devrait permettre, à terme, de les utiliser pour aider le décideur public à réduire les risques de propagation d'épizooties, par exemple en mettant au point des stratégies de lutte ou de contrôle ciblées.

Aurore Payen

Centre d'études et de prospective

12. Büttner K., Krieter J., Traulsen A. et Traulsen I., 2016, « Epidemic spreading in an animal trade network - Comparison of distance-based and network-based control measures », *Transboundary and emerging diseases*, 63(1), e122-e134.

13. Payen A., Tabourier L. et Latapy M., 2019, « Spreading dynamics in a cattle trade network: Size, speed, typical profile and consequences on epidemic control strategies », *PLoS one*, 14(6).

Ministère de l'Agriculture et de l'Alimentation
Secrétariat Général

Service de la statistique et de la prospective
Centre d'études et de prospective
3 rue Barbet de Jouy
75349 PARIS 07 SP
Sites Internet : www.agreste.agriculture.gouv.fr
www.agriculture.gouv.fr

Directrice de la publication : Corinne Prost

Rédacteur en chef : Bruno Héroult
Mel : bruno.herault@agriculture.gouv.fr
Tél. : 01 49 55 85 75

Dépôt légal : À parution © 2020