

## **ANALYSE DU RESEAU D'ECHANGES BOVINS FRANÇAIS. FOCUS SUR DES SOUS-GROUPES D'ETABLISSEMENTS A STRUCTURE VULNERABLE FACE A LA DIFFUSION DE MALADIE ; LES « GIANT STRONG COMPONENTS »\***

**Séverine Rautureau<sup>1</sup>, Barbara Dufour<sup>2</sup> et Benoît Durand<sup>1</sup>**

### **RESUME**

Les mouvements d'animaux forment un réseau complexe reliant les exploitations entre elles et représentent un risque important de dissémination d'agents infectieux entre établissements d'élevage et ce, sur de longues distances. Le réseau de mouvements de bovins à travers la France a été étudié, avec les méthodes d'analyse des réseaux sociaux, dans le but de déterminer des caractéristiques qui pourraient avoir des implications pour la diffusion d'épizooties.

Dans cet objectif, des sous-groupes d'établissements de grande taille fortement connectés entre eux, les « giant strong components » (GSC) ont été analysés. Ces GSCs représentent des vulnérabilités structurelles du réseau d'élevages français vis-à-vis d'un risque épizootique, de façon permanente et sur tout le territoire. Les conditions permettant d'éviter leur formation afin de diminuer la vulnérabilité du réseau ont été alors étudiées. L'analyse des flux passant par les marchands (centres de rassemblement et marchés) a montré qu'une action de maîtrise des flux ciblant à la fois les deux types d'établissements était nécessaire. Mais, un ciblage sur des établissements selon leur « centralité » (positionnement) dans le réseau était plus pertinent et efficace en termes de nombre d'établissements à surveiller.

**Mots-clés** : analyse de réseaux sociaux, mouvements de bovins, diffusion de maladie.

### **SUMMARY**

Movements of animals form complex networks linking holdings and are associated with a major risk of dissemination of infectious agents between farms. Using the social network analysis method, the network of cattle movements throughout France has been studied to suggest features that might have important implications for disease spread.

For that purpose, large cohesive sub-groups, so-called « giant strong components » (GSC), have been identified; GSCs represent structural vulnerability of the French cattle trade facing epidemic risks. They are consistently present and widely spread all over the country.

.../..

\* Texte de la communication orale présentée au cours des Journées scientifiques AEEMA, 21 mai 2010

<sup>1</sup> Anses, Laboratoire de santé animale, Unité d'épidémiologie (EPI), 23, avenue du Général de Gaulle, 94701 Maisons-Alfort Cedex, France

<sup>2</sup> ENVA : Unité de recherche d'épidémiologie des maladies animales infectieuses (EPIMAI), 7, avenue du Général De Gaulle, 94704 Maisons-Alfort, France

.../..

The way to avoid their emergence in order to decrease network vulnerability was studied. The analysis of commercial flows has shown that control measures affecting both types of merchants (dealers and markets) were necessary. On the other hand, targeting holdings based on their central position in the network were more relevant and more efficient in terms of number of holdings calling for surveillance.

**Keywords:** Social network analysis, Cattle movements, Disease spread.




---

## I - INTRODUCTION

---

Les mouvements d'animaux représentent la principale voie de transmission de maladies contagieuses d'élevage à élevage. L'épizootie de fièvre aphteuse du Royaume-Uni en 2001 qui a été extrêmement sévère, notamment lors de la phase silencieuse [Gibbens *et al.*, 2001], confirme non seulement le rôle des mouvements d'animaux dans la diffusion de maladies mais attire également l'attention sur une diffusion à bas bruit lors d'échanges de routine, sans mesure de prévention ciblée ou systématique sur ces mouvements. Ces transferts d'animaux reliant les exploitations entre elles forment en fait un réseau complexe

à travers lequel les agents pathogènes se propagent empruntant les mouvements comme voies entre les entités. Il est alors intéressant d'étudier plus précisément l'organisation de ces échanges ; la filière est-elle toujours vulnérable ? Ce réseau favorise-t-il la diffusion de maladies ? L'étude de cette structure par l'analyse de réseau permet de caractériser ces relations. Les indicateurs calculés renseignent sur la place et le rôle structural des types d'exploitations présentes et mettent en évidence des groupes d'établissement fortement connectés.

---

## II - MATERIELS ET METHODES

---

### 1. DONNEES

#### 1.1. DONNEES DE MOUVEMENTS

Depuis 2001, la France a renforcé considérablement son processus de maîtrise de la traçabilité des animaux : les mouvements de bovins sont désormais enregistrés : chaque détenteur doit notifier chaque événement de la vie d'un bovin. Ainsi nous avons collecté trois jeux de données « détenteurs » avec l'ensemble de leurs notifications, issus de la base de donnée nationale d'identification (BDNI) gérée par le ministère de l'Agriculture ; les éleveurs (263 907), les centres de rassemblement (1 315) et les marchés (76) pour l'année 2005 (cette année a été choisie

comme plus « représentative » des mouvements habituels d'animaux car étant située avant les restrictions de mouvements liées à l'épizootie de FCO intervenue en France à partir de 2006).

Les enregistrements ont été rassemblés, recréant la chronologie des mouvements entre chacun des détenteurs (ou établissements). Ainsi, sur les 27 millions de bovins différents qui ont transité pendant l'année 2005, six millions se sont déplacés au moins une fois d'un établissement à un autre. Ont été ainsi enregistrés 18 millions de mouvements individuels et six millions de mouvements de lots d'animaux.

## 1.2. CREATION DE RESEAUX

En termes de réseau, chaque établissement est assimilé à un « nœud » et les mouvements à des liens ou plus spécifiquement des arcs en tenant compte du sens des mouvements (figure 1.a). Ici, un lien correspondait à au moins un échange d'animaux entre deux exploitations réalisé au cours de la période étudiée. Ainsi, pour l'année, le réseau étudié comportait 244 097 nœuds et 1 416 208 arcs. Des réseaux mensuels et hebdomadaires ont été compilés. En moyenne, un réseau mensuel

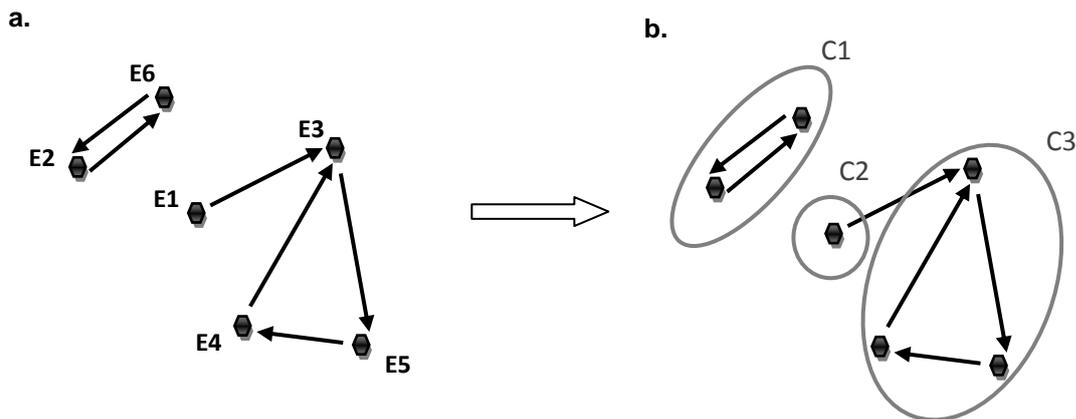
(hebdomadaire) était composé de 126 613 (49 246) nœuds et 212 497 (59 933) arcs (tableau 1).

De plus, des réseaux transformés ont été construits ; des liens directs ont été recréés entre établissements, court-circuitant soit tous les centres de rassemblement (réseau « sans centres ») soit tous les marchés (réseau « sans marchés ») soit les deux (réseau « sans marchands »). Des réseaux mensuels et hebdomadaires ont également été construits.

Figure 1

### Réseau d'établissements d'élevage et mise en évidence de « strong components » (sous-groupe d'établissements fortement connectés où chaque entité est accessible par chacun)

- a. Exemple avec un réseau de six établissements  
b. Trois « strong components » mis en évidence



## 2. ANALYSE DE RESEAUX

L'analyse de réseaux sociaux a été utilisée. Cette méthode caractérise les réseaux en calculant des indicateurs pour les nœuds et leurs relations indiquant leur place et leur rôle dans le réseau [Wasserman, 1994]. Des groupes de nœuds proches et très connectés peuvent être mis également en évidence. Ainsi, l'étude a été dirigée vers des sous-structures cohésives représentant des établissements fortement reliés par des mouvements et donc des établissements dont l'organisation peut favoriser la diffusion de maladies entre eux.

### 2.1. MISE EN EVIDENCE DES « GIANT STRONG COMPONENTS » (GSCs)

Les sous-groupes étudiés sont appelés « strong components » en termes de réseaux sociaux. A l'intérieur de ces groupes, chaque entité peut être accessible par chacun des autres établissements (figure 1.b). Pour chaque réseau construit, la recherche de ces sous-groupes a été faite ainsi qu'une distribution de leur taille. Dès qu'un sous-groupe de grande taille s'est distingué des autres de façon notable, il a été considéré comme un « giant » (GSC) et a constitué notre unité d'étude (figure 2).

Ainsi, nous nous sommes intéressés à des groupes d'établissements en nombre important tous connectés entre eux, ce qui a défini une structure à risque face à la diffusion de maladie.

## 2.2. ETUDE DES GSCs

Tout d'abord, l'émergence et la répartition géographique des GSCs ont été étudiées pour les deux types de réseaux (entiers et transformés). Ces observations doivent permettre de comparer deux situations ; la première où les mouvements ne sont pas entravés, la deuxième où des établissements (marchés et/ou centres) sont fermés et les flux se réorganisent (ou la situation où une « biosécurité parfaite » est artificiellement mise en place ; les flux passant par ces établissements ne croisent pas).

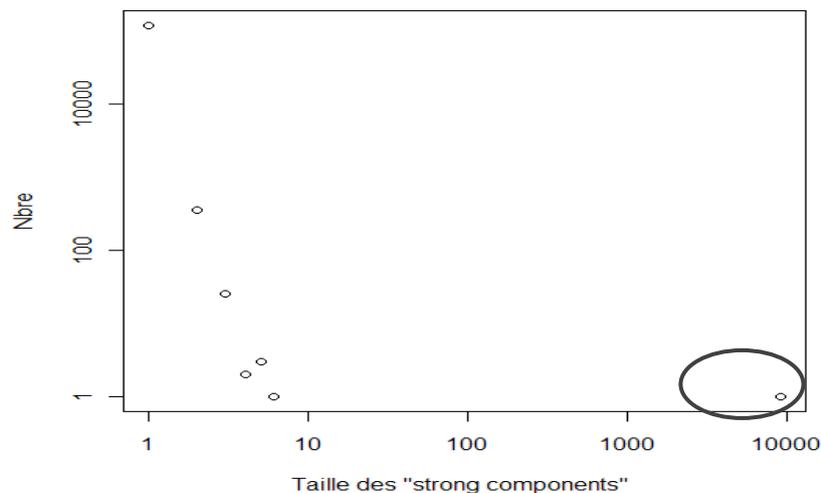
Ensuite, l'étude structurale doit préciser les caractéristiques des GSCs issus des réseaux

entiers. Leur composition par type d'établissements a été observée et les mesures de centralité des nœuds (établissements) ont été calculées selon la méthode d'analyse de réseaux sociaux [Freeman, 1978/1979 ; Wasserman, 1994]. Cette approche peut permettre de trouver des fragilités dans ces structures afin d'orienter des moyens d'action pour les déstructurer [Albert *et al.*, 2000]. Ces « attaques » du réseau par éviction de nœuds pourraient être alors assimilées à des mesures de prévention ou de lutte (fermeture ou interdiction de mouvements pour un établissement).

Figure 2

### Distribution de la taille des « strong components » et mise en évidence d'un « giant strong component »

Réseau d'établissements français de bovins, Janvier 2005



Ainsi, une analyse avec retraits de nœuds selon leurs mesures de centralité a été menée. Les nœuds à plus forte centralité sont éliminés un à un et l'émergence de GSC est examinée jusqu'à l'obtention d'un « strong component » de taille inférieure à 51. Cette taille correspond, en effet, au plus grand « strong component secondaire » observé dans les réseaux entiers hebdomadaires. Les

indicateurs ont été recalculés et classés entre chaque éviction.

Les analyses ont été réalisées avec les logiciels suivants : Pajek, programme pour les analyses de réseaux de grande taille (v1.25, <http://pajek.imfm.si/doku.php?id=pajek>) [Batagelj et Mrvar, 1998] et le package Igraph du logiciel R (v 0.5.2, <http://igraph.sourceforge.net>) [Collectif, 2009].

### III - RESULTATS

#### 1. EMERGENCE DES GSCS

La distribution des tailles des « strong components » réalisée pour les réseaux entiers (sur des pas de temps annuel, mensuels et hebdomadaires) a révélé la présence de larges « strong components » que l'on a considérés comme des « giants ». Les « giants » mensuels avaient une taille

moyenne de 10 277 établissements, les hebdomadaires de 2 115 (tableau 1 et figure 3).

On retrouve dans les tailles des GSCs mensuels les fluctuations saisonnières des flux avec un pic au printemps et en automne (figure 3).

Tableau 1

Caractéristiques des réseaux d'élevages et des « giant strong components » émergents à l'échelle annuelle, mensuelle et hebdomadaire, France, 2005

	Réseau annuel	Réseau mensuel moyen	Réseau hebdomadaire moyen
Taille (nombre établissements)	244 097	126 613	49 246
Liens entre établissements (nombre)	1 416 208	212 497	59 933
Nombre SC <sup>(1)</sup>	1 259	362	125
Taille SC1 <sup>(2)</sup>	<b>108 904 (44,6%)</b>	<b>10 277 (8,1%)</b>	<b>2 115 (4,3%)</b>
Taille SC2 <sup>(3)</sup>	9 (3,7.10 <sup>-3</sup> )	7 (5,5.10 <sup>-3</sup> %)	9 (1,8.10 <sup>-2</sup> %)
<b>Détails GSC :</b>			
- liens (nombre)	804 331	39 592	7 417
- longueur moyenne du chemin le plus court <sup>(4)</sup>	4,44	4,74	5,07
- diamètre <sup>(5)</sup>	20	16	14

(1) SC = « strong component »

(2) SC1 = le « strong component » le plus grand ⇒ possible « giant » (GSC) et le pourcentage d'établissements concernés

(3) SC2 = le second « strong component » plus grand

(4) Indicateur SNA : average shortest path length = nombre de liens entre deux établissements pour le chemin le plus court

(5) Indicateur SNA : diamètre = le plus long des « shortest path », nombre de liens

La même étude sur les réseaux transformés, soit sans centres, soit sans marchés, a permis de mettre en évidence l'émergence de GSCs mensuels et hebdomadaires avec des caractéristiques similaires à ceux issus des réseaux entiers. Par contre, quand à la fois les centres et les marchés sont court-circuités, on observe la disparition de tout GSC.

#### 2. LOCALISATION DES GSCS

Les GSCs étaient largement répartis sur le territoire français et couvraient les principales zones d'élevage (figure 4.a). La dispersion géographique des GSCs issus des réseaux « sans centres » ou « sans marchés » demeuraient semblables. Par contre, pour les GSCs issus des réseaux « sans marchands », les « strong components » observés étaient de petite taille (inférieure en moyenne à 10) et localement répartis (figure 4.b.).

Figure 3

## Répartition du nombre d'établissements participant à des échanges dans les « giant strong components » mensuels et hebdomadaires sur l'année 2005

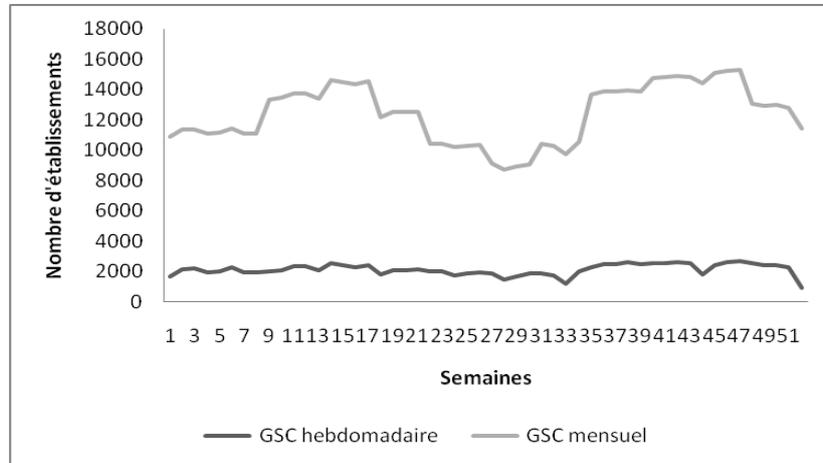
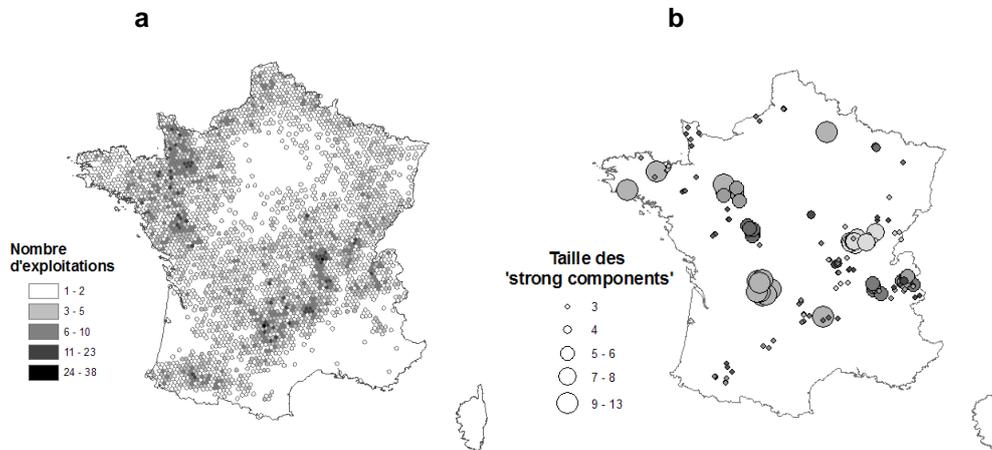


Figure 4

## Répartition géographique des établissements participant à des échanges dans les « strong components » - Janvier 2005

- (a) le « giant strong component » du réseau entier - : 9 137 établissements.  
 (b) les « strong components » du réseau « sans marchands » : 46 strong components, 167 établissements



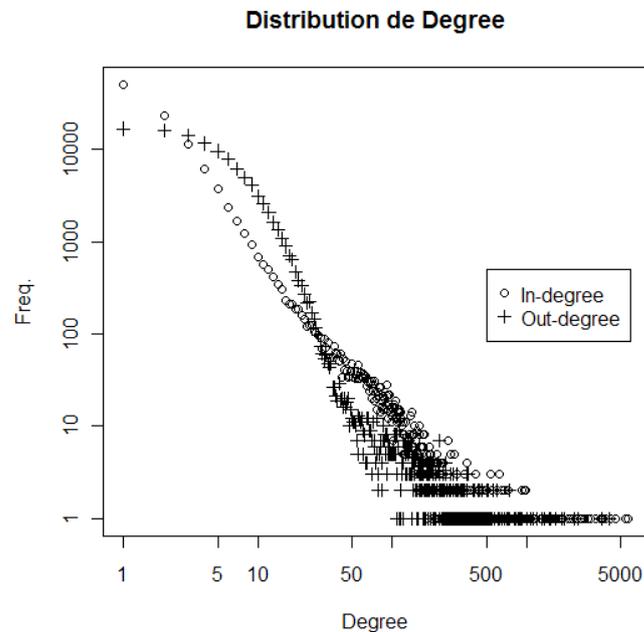
## 3. STRUCTURE DES GSCS

L'étude des nombres de liens de sortie ou d'entrée (degrés) a mis en évidence une distribution caractéristique (figure 5) ; beaucoup d'établissements ont peu de liens avec les autres et un faible nombre assure la majorité des liens. Cette distribution dont la fin de la courbe tend à suivre une loi de puissance permet de qualifier la structure comme un réseau « scale-free ». Ce type de réseau a

comme propriétés de favoriser la diffusion de maladie et d'être sensible à des attaques ciblées plutôt qu'aléatoires [Albert *et al.*, 2000 ; Barabasi, 2009]. Cette robustesse signifie que le réseau conserve sa structure avec la majorité des liens si une attaque consiste à retirer des nœuds de façon aléatoire. Par contre, une attaque ciblée notamment parmi des nœuds fortement connectés (se trouvant en fin de courbe), déstructure plus facilement le réseau.

**Figure 5**  
**Distribution de degrés ; liens d'entrée et sortie d'établissements**  
**détenant des bovins pour l'année 2005 - France**

In Degree: Mouvements d'entrée, Out Degree: Mouvements de sortie



Par ailleurs, l'« average shortest path » (nombre de liens permettant d'accéder d'un établissement à un autre) était en moyenne de cinq pour les différents pas de temps. Chaque établissement était donc relié en moyenne par cinq mouvements, avec quatre établissements intermédiaires (tableau 1). Cet indicateur est en relation avec la vitesse de diffusion d'une maladie.

#### 4. DESTRUCTURATION DES GSCS

L'analyse avec retraits de nœud selon leurs mesures de centralité a permis de limiter

l'apparition de GSC en évinçant, pour la procédure la plus satisfaisante, en moyenne 299 établissements pour les réseaux mensuels et 82 pour les réseaux hebdomadaires. Cela représentait moins de 1% de la population impliquée dans la période d'étude. L'« attaque » selon les mesures de centralité a réduit le nombre d'établissements à cibler. Les centres et les marchés étaient largement représentés. Ainsi, au lieu de cibler la totalité des marchands, cette méthode proposait de ne dédier des mesures de lutte que sur 35 à 56% des marchés et sur 6 à 23% des centres.

---

## IV - DISCUSSION

---

### 1. LES DONNEES

Cette étude ciblait seulement les échanges entre les établissements français, permettant une estimation du risque de diffusion de maladies pour le territoire national. De plus, seuls les établissements hébergeant les animaux ont été retenus. Les abattoirs et équarrissages ont été exclus, car étant

considérés comme ne participant pas activement à la diffusion de maladies.

Les données provenaient des notifications réglementées et obligatoires qui concernent les mouvements avec changement de détenteurs. Les mises en pâture en commun ou les transhumances notamment, ou autres commerces ne sont pas enregistrées dans ce

système qui comprend tout de même la majorité des mouvements de bovins.

## 2. L'ANALYSE DE RESEAU

Les réseaux obtenus sont des réseaux de très grande taille. C'est pourquoi, cette première approche a été faite sur des réseaux simples (un lien = au moins un mouvement entre deux exploitations) pour faciliter les calculs d'indicateurs. Ce choix semble cependant avoir peu d'incidence car sur l'année, 82% de ces liens sont des mouvements uniques entre établissements.

L'analyse de réseau offre un large panel d'outils pour étudier ce réseau d'élevages. Notre choix s'est porté avant tout sur l'étude de structures à risque simples, des sous-groupes d'établissements fortement connectés, les GSCs. La cohésion des établissements dans ces sous-groupes et la prise en compte du sens des mouvements permettent d'approcher un risque de vulnérabilité de diffusion de maladies. La chronologie n'est cependant pas respectée mais seulement approchée par un découpage en mois et en semaine. Cette étude pourrait alors être couplée à d'autres outils comme la détection d'agrégats spatio-temporels ou par des simulations.

Il existe par ailleurs, d'autres entités structurales intéressantes à étudier, par exemple les groupes d'établissements qui viendraient alimenter les GSCs (ou inversement) induisant un risque d'extension plus large de maladies.

## 3. GSCS ET DIFFUSION DE MALADIES

Les GSCs par leurs caractéristiques sont des structures favorisant la diffusion de maladies : propriétés des réseaux « scale free », grande cohésion et large nombre d'entités réparties sur la quasi totalité du territoire français. Ces propriétés indiquent aussi comment agir pour limiter cette cohésion ; le réseau est sensible aux « attaques » ciblées sur des nœuds ayant une place centrale.

Naturellement, nous nous sommes intéressés aux centres de rassemblements et aux marchés. L'analyse avec les flux transformés (remplacement de flux indirects par des flux directs d'élevage à élevage) a permis de démontrer qu'une action sur l'ensemble de ces acteurs était nécessaire. Cette application systématique de mesures de lutte ou de prévention apparaît toutefois difficilement réalisable et sur du long terme entrave lourdement le commerce.

Par contre, la suite de l'étude a permis de préciser les actions en triant les nœuds selon leurs mesures de centralité au sein du réseau. Une pression de surveillance ou une limitation de mouvements temporaire sur un plus faible nombre d'établissements par rapport à un ciblage par catégorie d'activité peut suffire à déconnecter le réseau et donc à limiter un risque de large diffusion de maladies.

Bien sûr, pour cela il faut une action rapide puisque ces indicateurs se placent dans un réseau donné et que tout événement notable survenant par la suite (comme des mesures de prévention ou de lutte à la suite d'une épizootie imminente ou déclarée) recrée un nouveau réseau ; des établissements sont isolés, fermés, les réseaux commerciaux cessent leur activité ou se réorganisent.

---

## V - CONCLUSION

---

Cette étude d'une partie de la structure du réseau d'échange de bovins français nous révèle la vulnérabilité structurelle de la filière. Ces sous-groupes, les GSC peuvent être considérés comme des structures à risque facilitant la diffusion de maladies contagieuses. Le fait de limiter leur importance permet alors de diminuer un risque de large diffusion par le biais de mouvements d'animaux. Pour cela, l'étude a suggéré de cibler les mesures de

lutte et de prévention sur un faible nombre d'établissements grâce à des indicateurs empruntés à la méthode d'analyse de réseaux sociaux. Une partie des marchands (centres et marchés) par leur fonctionnement de rassemblement et de redistribution d'animaux est naturellement visée. Ce ciblage permet de proposer une stratégie optimale pour le déploiement des ressources sans entraver trop lourdement les flux commerciaux.

---

## BIBLIOGRAPHIE

---

Albert R., Jeong H., Barabasi A.L. - Error and attack tolerance of complex networks. *Nature*, 2000, **406**, 378-382.

Barabasi A.L. - Scale-free networks: a decade and beyond. *Science*, 2009, **325**, 412-413.

Batagelj V., Mrvar A. - Pajek - Program for Large Network Analysis. *Connections*, 1998, **21**, 47-57.

Collectif R. - A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria, 2009.

Freeman L.C. - Centrality in social networks, conceptual clarification. *Soc. Networks*, 1978/1979, **1**, 215-239.

Gibbens J.C., Wilesmith J.W., Sharpe C.E., Mansley L.M., Michalopoulou E., Ryan J.B.M., Hudson M. - Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months. *Vet. Rec.*, 2001, **149**, 729-743.

Wasserman S.F.K. - Social Network Analysis : Methods and Applications. University Press Cambridge. Cambridge, 1994.

