

Réseaux en épidémiologie des maladies infectieuses et écologie

Elisabeta Vergu

INRA, Université Paris Saclay, MaIAGE UR 1404, Jouy-en-Josas
équipe Dynenvie (Modélisation dynamique et statistique pour les écosystèmes, l'épidémiologie et l'agronomie)

*AG du Département MIA / NUMM
Palaiseau et Jouy-en-Josas, 21-23 mai 2019*



Introduction

Réseau : ensemble de **noeuds / sommets** (entités de nature très diverse en interaction) reliés par des **liens / arêtes** (relations entre les entités).

"Les réseaux représentent une façon de voir le monde et de structurer les données complexes." (A.L. Barabasi)

Théorie des graphes : origine attribuée à Euler (18ème siècle) ; premier modèle de graphe aléatoire par Erdős-Rényi (1959).

Résurgence de l'intérêt pour les réseaux à la fin des années 1990, dans la communauté des physiciens (Barabási et al, Newman et al).

Champ de recherche foisonnant dans de nombreux domaines : statistique, probabilités, physique, informatique, biologie des systèmes, écologie, épidémiologie,...

Réseaux et CT2 à MIA/NUMM : toutes les unités MIA ; LISC, TETIS.

Réseaux en écologie, enjeux maths-info*

Principalement deux types de réseaux d'interactions entre espèces :

- bipartis (ex. proies-prédateurs)
- orientés ou non, chaque noeud est une espèce

Deux grands types de questions/objectifs :

- apprendre le réseau à partir de données (**Inférence**)
 - ▶ données incomplètes/rares issues d'échantillonnage : développer des méthodes d'inférence adéquates pour données éparses
 - ▶ données NGS : développer des algorithmes qui passent à l'échelle
 - ▶ prise en compte de la dépendance temporelle dans les données, de covariables dans la modélisation
- trouver une organisation/structure dans un réseau connu (**Analyse**)
 - ▶ par ex par modèles à blocs latents (LBM) ou stochastiques (SBM) : développer des algorithmes d'estimation qui passent à l'échelle et traitent des extensions

*(suite échanges avec Nathalie Peyrard et Stéphane Robin)

Réseaux en épidémiologie des maladies infectieuses, enjeux maths-info

Noeuds = individus ou groupes d'individus (foyers, villes, fermes).

Liens = liens sociaux, sexuels, trafic aériens, mouvements d'animaux

Réseaux **orientés, pondérés, temporellement dynamiques**, souvent de **grande taille** (e.g. $> 10^5$ noeuds)

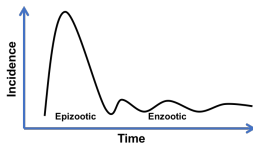
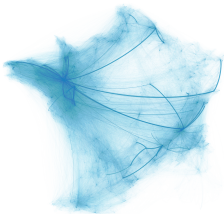
Questions et enjeux de modélisation en lien avec le risque de survenue et l'issue d'une épidémie se déployant sur réseaux :

- impact de la structure des réseaux, prise en compte des hétérogénéités
- modèles de réseaux pour données avec diverses structures de corrélation
- modèles de réseaux dynamiques et des processus épidémiques couplés, étude de leurs comportement
- modélisation des changements de comportement des individus (noeuds) modifiant dynamiquement le réseau.

Questions pour le cas des données de réseaux incomplètes...

Contexte : réseaux d'échanges d'animaux et maladies endémiques

- Données détaillées issues du traçage des animaux d'élevage :
nodes = fermes & opérateurs commerciaux ; **liens** = mouvements d'animaux
- Spécificités : dynamiques endémiques longues, populations d'animaux gérés, processus démographiques et infectieux à des échelles de temps comparables
- Questions d'intérêt :
 - ▶ vulnérabilité du réseau à la transmission de l'infection
 - ▶ centralité des noeuds (sentinelles ou cibles pour les mesures de maîtrise)
 - ▶ mécanismes à la base des échanges marchands, modèles de prédiction
 - ▶ prédiction de la propagation d'infections via les mouvements d'animaux, modèles de métapopulations sur réseaux



Questions abordées / en cours

- Explorer les caractéristiques du réseaux de commerce des bovins en France (2005-2009) en lien avec le risque épidémique (Dutta et al 2014 ; Moslonka-Lefebvre et al 2016)
 - ▶ clustering, topologie, vulnérabilité
- Comprendre les mécanismes à la base de l'évolution de ces réseaux et construire des modèles de prédiction (Hoscheit et al 2016 ; Dalmau & Salvi, soumis)
 - ▶ données issues d'enquêtes de terrain
 - ▶ modèles mécanistes ; modèles statistiques
- Etudier les dynamiques couplées épidémiques et des réseaux sous-jacents à l'aide de modèles génériques (PhD P. Montagnon, co-encadrement V. Bansaye ; post-doc M. Salvi, collab. V. Bansaye, V.C. Tran)
 - ▶ construction de modèle et étude de comportement (différentes asymptotiques)
- Evaluer des stratégies de contrôle pour diverses maladies (BVD, paraTB ; FMD, bTB) avec des modèles spécifiques (Beaunée et al, 2015, 2017 ; Qi et al, 2019)
 - ▶ modèles de simulation, exploration numérique

*(projet ANR CADENCE)

Questions abordées / en cours

- Explorer les caractéristiques du réseaux de commerce des bovins en France (2005-2009) en lien avec le risque épidémique (Dutta et al 2014 ; Moslonka-Lefebvre et al 2016)
 - ▶ clustering, topologie, vulnérabilité
- Comprendre les mécanismes à la base de l'évolution de ces réseaux et construire des modèles de prédiction (Hoscheit et al 2016 ; Dalmau & Salvi, soumis)
 - ▶ données issues d'enquêtes de terrain
 - ▶ modèles mécanistes ; modèles statistiques
- Etudier les dynamiques couplées épidémiques et des réseaux sous-jacents à l'aide de modèles génériques (PhD P. Montagnon, co-encadrement V. Bansaye ; post-doc M. Salvi, collab. V. Bansaye, V.C. Tran)
 - ▶ construction de modèle et étude de comportement (différentes asymptotiques)
- Evaluer des stratégies de contrôle pour diverses maladies (BVD, paraTB ; FMD, bTB) avec des modèles spécifiques (Beaunée et al, 2015, 2017 ; Qi et al, 2019)
 - ▶ modèles de simulation, exploration numérique

*(projet ANR CADENCE)

Questions abordées / en cours

- Explorer les caractéristiques du réseaux de commerce des bovins en France (2005-2009) en lien avec le risque épidémique (Dutta et al 2014 ; Moslonka-Lefebvre et al 2016)
 - ▶ clustering, topologie, vulnérabilité
- Comprendre les mécanismes à la base de l'évolution de ces réseaux et construire des modèles de prédiction (Hoscheit et al 2016 ; Dalmau & Salvi, soumis)
 - ▶ données issues d'enquêtes de terrain
 - ▶ modèles mécanistes ; modèles statistiques
- Etudier les dynamiques couplées épidémiques et des réseaux sous-jacents à l'aide de modèles génériques (PhD P. Montagnon, co-encadrement V. Bansaye ; post-doc M. Salvi, collab. V. Bansaye, V.C. Tran)
 - ▶ construction de modèle et étude de comportement (différentes asymptotiques)
- Evaluer des stratégies de contrôle pour diverses maladies (BVD, paraTB ; FMD, bTB) avec des modèles spécifiques (Beaunée et al, 2015, 2017 ; Qi et al, 2019)
 - ▶ modèles de simulation, exploration numérique

*(projet ANR CADENCE)

Questions abordées / en cours

- Explorer les caractéristiques du réseaux de commerce des bovins en France (2005-2009) en lien avec le risque épidémique (Dutta et al 2014 ; Moslonka-Lefebvre et al 2016)
 - ▶ clustering, topologie, vulnérabilité
- Comprendre les mécanismes à la base de l'évolution de ces réseaux et construire des modèles de prédiction (Hoscheit et al 2016 ; Dalmau & Salvi, soumis)
 - ▶ données issues d'enquêtes de terrain
 - ▶ modèles mécanistes ; modèles statistiques
- Etudier les dynamiques couplées épidémiques et des réseaux sous-jacents à l'aide de modèles génériques (PhD P. Montagnon, co-encadrement V. Bansaye ; post-doc M. Salvi, collab. V. Bansaye, V.C. Tran)
 - ▶ construction de modèle et étude de comportement (différentes asymptotiques)
- Evaluer des stratégies de contrôle pour diverses maladies (BVD, paraTB ; FMD, bTB) avec des modèles spécifiques (Beaunée et al, 2015, 2017 ; Qi et al, 2019)
 - ▶ modèles de simulation, exploration numérique

*(projet ANR CADENCE)

Questions abordées / en cours

- Explorer les caractéristiques du réseaux de commerce des bovins en France (2005-2009) en lien avec le risque épidémique (Dutta et al 2014 ; Moslonka-Lefebvre et al 2016)
 - ▶ clustering, topologie, vulnérabilité
- Comprendre les mécanismes à la base de l'évolution de ces réseaux et construire des modèles de prédiction (Hoscheit et al 2016 ; Dalmau & Salvi, soumis)
 - ▶ données issues d'enquêtes de terrain
 - ▶ modèles mécanistes ; modèles statistiques
- Etudier les dynamiques couplées épidémiques et des réseaux sous-jacents à l'aide de modèles génériques (PhD P. Montagnon, co-encadrement V. Bansaye ; post-doc M. Salvi, collab. V. Bansaye, V.C. Tran)
 - ▶ construction de modèle et étude de comportement (différentes asymptotiques)
- Evaluer des stratégies de contrôle pour diverses maladies (BVD, paraTB ; FMD, bTB) avec des modèles spécifiques (Beaunée et al, 2015, 2017 ; Qi et al, 2019)
 - ▶ modèles de simulation, exploration numérique

*(projet ANR CADENCE)

Modèles de propagation d'épidémies SIR sur graphe explicite

Montagnon P, 2019 *Stoch. Proc. Appli*;

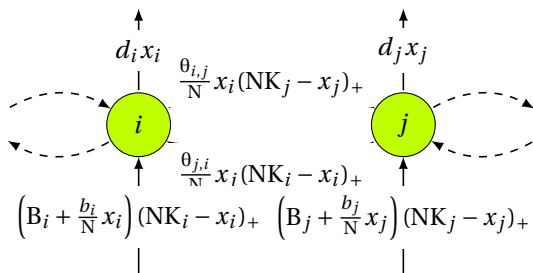
Montagnon P, 2019 *J Math. Biol.*;

Montagnon P, Bansaye V, Deslandes F, Vergu E, 2019 (en préparation)

Objectifs : construction et analyse de comportement d'un modèle générique ;
construction d'indicateurs de vulnérabilité / importance des noeuds du réseaux

1- Le modèle de population

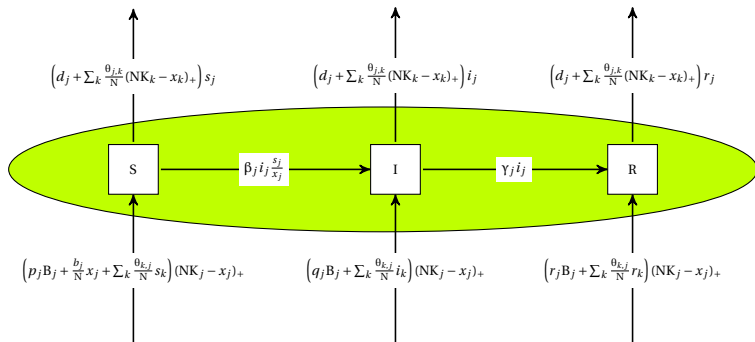
- n nœuds représentant les opérateurs et fermes
- Modèle markovien à espaces d'états discrets (processus de sauts $X^N = (X^N(t))_{t \geq 0}$ à valeurs dans \mathbb{N}^n , N paramètre d'échelle)
- Le nœud i a une capacité K_i , population $x = (x_1, \dots, x_n)$



Modèle de métapopulations avec immigration et transferts logistiques.

Le processus déterministe limite admet un équilibre $z^* = (z_1^*, \dots, z_n^*)$ (pas d'information sur l'unicité ou la stabilité en général).

2- Le processus épidémique couplé à la dynamique de population



Au sein de chaque nœud j , on distingue trois compartiments (S, I et R), $S^N(t) + I^N(t) + R^N(t) = X^N(t)$ et on considère un processus markovien de sauts SIR.

2- Limite d'échelle du processus épidémique

On suppose que $X^N(0) = \lfloor Nz^* \rfloor$, que $I^N(0) = I(0) \in \mathbb{R}_+^n$ et que $R^N(0) = 0$. On considère (s, i) tel que $(s, i)(0) = (z^*, 0)$ et

$$\begin{cases} s'_j = (p_j B_j + b_j z_j^*) (K_i - z_j^*)_+ - d_j s_j \\ \quad + \sum_{k=1}^n \left(\theta_{j,k} s_j (K_k - z_k^*)_+ - \theta_{k,j} s_k (K_j - z_j^*)_+ \right) - \beta_j \frac{i_j s_j}{z_j^*} \\ i'_j = q_j B_j + \beta_j \frac{i_j s_j}{z_j^*} - d_j i_j - \gamma_j i_j \\ \quad + \sum_{k=1}^n \left(\theta_{j,k} i_j (K_k - z_k^*)_+ - \theta_{k,j} i_k (K_j - z_j^*)_+ \right). \end{cases}$$

Limite d'échelle du processus épidémique

Pour tout $T \geq 0$, presque sûrement

$$\lim_{N \rightarrow +\infty} \sup_{t \in [0, T]} \left\| \frac{(S^N(t), I^N(t), R^N(t))}{N} - (s(t), i(t), z^* - s(t) - i(t)) \right\|_{\infty} = 0.$$

3- Comportement du processus épidémique

$W_{i,j}$ = progéniture dans noeud j d'un individu né dans i pour un processus de **branchement (BP)** à taux de naissance β_i , taux de mort $d_i + \gamma_i$ et mêmes mouvements que le processus de population.

Fonction génératrice des $W_{i,j}$:

$$G(s) = (\text{diag}(\lambda_1(s), \dots, \lambda_n(s)) - \Theta)^{-1} \omega$$

où $\Theta = \left(\frac{\theta_{i,j}}{N} (NK_j - z_j^*) \right)_{i,j}$, $\lambda_i(s) = (1 - s_i)\beta_i + \gamma_i + d_i + \sum_{j \neq i} \frac{\theta_{i,j}}{N} (NK_j - z_j^*)$ et $\omega_i = d_i + \gamma_i$.

3- Identification des nœuds rendant le graphe vulnérable

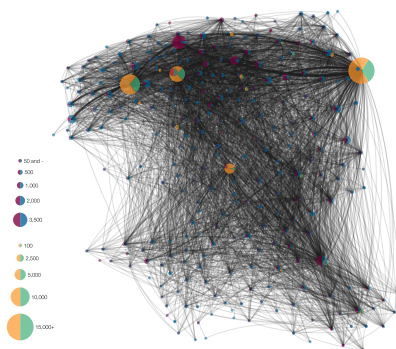
Cas sans immigration d'infectieux ($q_i = 0, \forall i$).

Plusieurs indicateurs :

- π_i , **probabilité d'épidémie majeure** au départ du nœud i (explosion du BP), $1 - s$ avec s point fixe de G .
- $\pi'_i := B_i(K_i - z_i^*)\pi_i$, probabilité de major outbreak pondérée par le flux d'immigration (qui conditionne le nœud dans lequel l'épidémie apparaît).
- $m_i := \mathbb{E}\left(\sum_{j \neq i} 1_{W_{i,j} > 0}\right)$, nombre moyen de nœuds atteints à la première étape par une épidémie au départ du nœud i , donné par $G((1, \dots, 1) - e_i)$ (« R_0 » d'un modèle centré sur les nœuds).
- $v_i := \mathbb{E}\left(\sum_{j=1}^n (1 - (1 - \pi_j)^{W_{i,j}})\right)$, indicateur « à deux pas » ($1 - (1 - \pi_j)^{W_{i,j}} =$ probabilité pour que les infectieux de seconde génération provoquent un épisode épidémique majeur), approché par $v'_i := -\sum_{j=1}^n \mathbb{E}(W_{i,j}) \log(\zeta_j)$.

4- Application numérique (données)

Réseau d'échanges de bovins du Finistère en 2015.



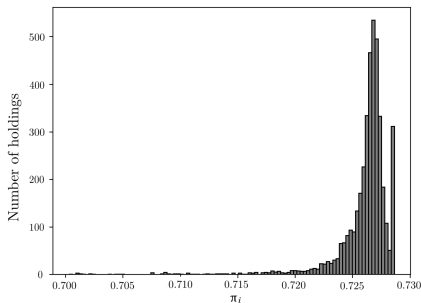
- 4.163 fermes, 3 marchés, 17 centres de rassemblement
- 118.311 animaux échangés au sein du département
- 55.325 animaux reçus de l'extérieur
- 241.747 animaux envoyés à l'abattoir ou hors du dép.
- Population moyenne au cours de l'année : 424.385 animaux

Paramètres démographiques calibrés sur ces données.

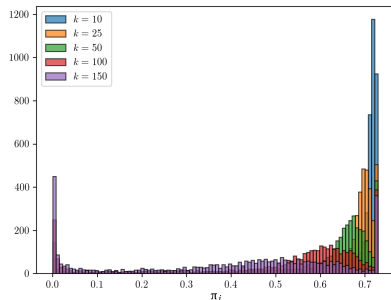
Paramètres épidémiologiques choisis compatibles avec ceux de la fièvre aphteuse :
 $\beta = 0.67j^{-1}$, $\gamma = 1/5.5j^{-1}$

4- Distribution des indicateurs : π_i

Histogram of π_i for $k = 1$ for the 4118 nodes
such that $\pi_i > 0.7$



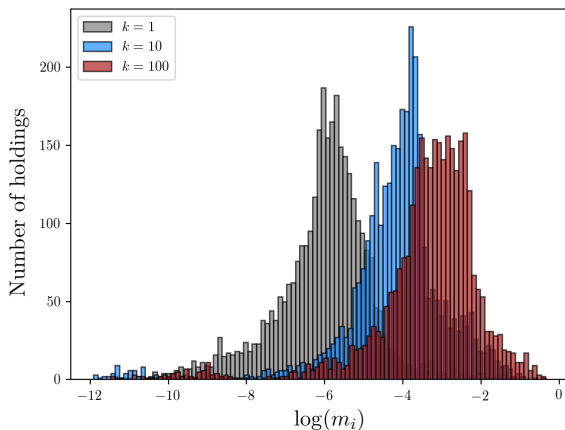
Histograms of π_i for various values of k



Probabilités d'un épisode épidémique majeur au départ des différents nœuds, en fonction d'un paramètre de ralentissement k (paramètres épidémiques remplacés par β/k et γ/k).

- $k = 1$ distribution très piquée, similaire à une pop.
- $k > 1$ la dispersion augmente en ralentissant l'épidémie

4- Distribution des indicateurs : m_i



Histogramme des $\log(m_i)$ pour différentes valeurs de k , sur les nœuds tels que $m_i > e^{-12}$ (3202 pour $k=1$, 3185 pour $k=10$ et 3199 pour $k=100$).

- distributions de « R_0 » se décalant vers la droite avec le ralentissement des épidémies (i.e. lorsque $k > 1$)

5- Simulations et stratégies de maîtrise (en cours)

Evaluation de la précision des indicateurs théoriques (basés sur une approximation par BP du modèle initial) par comparaison avec simulations de processus exacts.

Trois types de mesures de contrôle envisagées :

- Dépistage à l'importation (flux entrants bloqués ; $q_i \rightarrow 0$)
- Isolation des nœuds (flux sortants bloqués ; $\theta_{i,j} \rightarrow 0$)
- Vaccination d'animaux susceptibles (transfert d'animaux entre les compartiments S et R)

Chacune de ces mesures peut être ciblée sur les nœuds définis comme critiques au regard de l'un des indicateurs proposés (π_i , π'_i , m_i et v'_i).

Modélisation de la propagation de la BVD entre les fermes de bovins laitiers en Bretagne

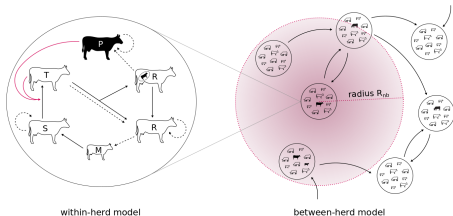
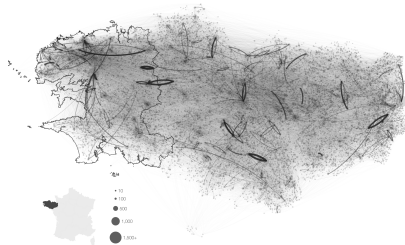
Qi et al, 2019 *Vet Res*

Objectifs : caractériser les rôles respectifs des différentes voies de transmission

Modèle stochastique de métapopulation sur réseau, propagation de la BVD

12750 troupeaux de bovins laitiers en Bretagne, 2005-2013.

Dynamique démographique pour chaque troupeau, dynamique épidémique intra et inter troupeaux, via le commerce d'animaux et les contacts au pâturage.



Paramètres calibrés sur données démographiques et d'échanges marchands et données épidémiologiques Finistère. Analyse des données simulées :

- rôle prépondérant de la contamination au pâturage (72%), mais épidémies plus durables si contamination suite à l'achat d'animaux
- caractérisation de la probabilité d'infection en fonction des caractéristiques réseau (degré pondéré) des troupeaux et de leurs voisins.

Conclusion et perspectives

- Descripteurs appropriés et proxy pour le risque et l'ampleur d'une épidémie, pour réseaux dynamiques, calculables en un temps raisonnable
- Modèles de réseaux dynamiques : phénoménologiques/statistiques ; mécanistes ; modèles de graphes aléatoires
 - ▶ dans le cas de mouvements commerciaux d'animaux : modèles différents en fonction du système d'élevage, des types d'animaux, probablement pas de modèle pour tout le réseau
- Modélisation multi-échelle des dynamiques épidémiques à une échelle régionale d'infection se propageant par le commerce d'animaux, pour
 - ▶ prédire la propagation d'épidémies
 - ▶ quantifier le rôle respectif des différentes voies de transmission
 - ▶ évaluer l'efficacité des stratégies de maîtrise