# Assessment of epidemiological risk related to pathogen spread on the network of French cattle movements

Elisabeta Vergu[1]

with Bhagat L. Dutta[1,2], Mathieu Moslonka-Lefebvre[1,3], Gaël Beaunée[1,2] and Pauline Ezanno[2]

other collaborators: Hervé Monod[1], João Filipe[3] and Chris Gilligan[3]

[1]INRA - MIA UR341, Jouy-en-Josas, France; [2]INRA, Oniris, LUNAM Université - UMR1300 BioEpAR, Nantes, France; [3]University of Cambridge - Department of Plant Sciences, United Kingdom

ModStatSAP 2014, Paris 11 December

# Context: cattle trade as a pathway for pathogen spread

- Important to understand the contact structure among herds underlying pathogen transmission
  - In particular for infectious diseases endemic in livestock: animal health and welfare; economic losses in animal productions; zoonoses
  - Animal trade networks characteristics impact infection spread
- Network representation: farms as nodes, trade relationships as links

# Animal trade as networks in the literature

- Aims of published studies
  - Demographic structure and pathogen dynamics
  - Dynamical patterns in longitudinal data, surveillance optimization
- Methods
  - Static networks of aggregated data
  - Continuous increments in steps of small time windows
  - Time-stamped networks
- Different countries and several animal markets
  - Cattle: Canada, Denmark, France, Italy, Sweden, UK, US
  - Pigs: Canada, France, Germany, Sweden

[Keeling and Eames, 2005; Bigras-Poulin *et al.*, 2006; Kao *et al.*, 2006; Kiss *et al.*, 2006, Natale *et al.*, 2009; Dube *et al.*, 2010; Danon *et al.*, 2011; Bajardi *et al.*, 2011, 2012; Rautureau *et al.*, 2011, 2012; Noremark *et al.*, 2011; Vernon and Keeling 2009; Vernon 2011; Buttner *et al.*, 2011; Dorjee *et al.*, 2013; Mweu *et al.*, 2011; Buttner *et al.*, 2013; Buhnerkempe *et al.*, 2013; Konschake *et al.*, 2013; Noremark and Widgren, 2014]

# Objectives

- Study of the French animal trade network over several years[1]
  - To identify if and which structural patterns underlying pathogen spread are stable in time
  - To explore control measures based on network-related characteristics of nodes (rely on previous data if real-time information not available)
- Joint economic and epidemiologic perspective[2]

[1] B.L. Dutta, P. Ezanno, E. Vergu. (2014) Characteristics of the spatio-temporal network of cattle movements in France over a 5-year period. *Prev. Vet. Med. 117(1):79-94.*

[2] M. Moslonka-Lefebvre, C. Gilligan, H. Monod, C. Belloc, P. Ezanno, J.A.N. Filipe, E. Vergu. (2014) Market analyses of livestock trade networks to inform the prevention of joint economic and epidemiological risk. *Submitted.*

# Some basic descriptors for the animal movement data in France (2005 - 2009)

- The data include all individual information for cattle (animal ID, date of birth, date, origin and destination holdings - farms (F), markets (M) and assembly centers (AC) - for each movement, etc).

| | **2005** | **2006** | **2007** | **2008** | **2009** |
|---|---|---|---|---|---|
| **Total holdings (V)** | 243,324 | -3.7 | -7.5 | -11.7 | -14.7 |
| **(M; AC)** | (84; 1547) | | | | |
| **No of movements (W)** | 8,636,018 | -0.1 | -9.2 | -11.7 | -12.2 |
| **No of links (A)** | 1,279,576 | -9.3 | -13.1 | -19.8 | -22.8 |
| **No of batches** | 2,791,261 | -4.7 | -12.8 | -17.7 | -20.4 |
| **No of cattle** | 5,533,854 | -0.8 | -6.1 | -7.6 | -9.6 |
| **No of holdings as origin** | 228,400 | -3.6 | -7.8 | -12.3 | -15.2 |
| **No of holdings as destination** | 141,249 | -3.4 | -7.7 | -12.3 | -16.1 |

The yearly values are shown as % in change with respect to observations in 2005.

- Main flows: F $\longrightarrow$ AC (40%); F $\longrightarrow$ F (18%); AC $\longrightarrow$ F (15%),

Dutta *et al.* (2014) *Prev. Vet. Med.*

# Sub-network of dairy herds in Finistère as an example

Subnetwork (2005-2009) of dairy herds trading non beef animals $> 6$ months

- 1945 herds
- 89,281 movements

# Some elements for network analysis (1)

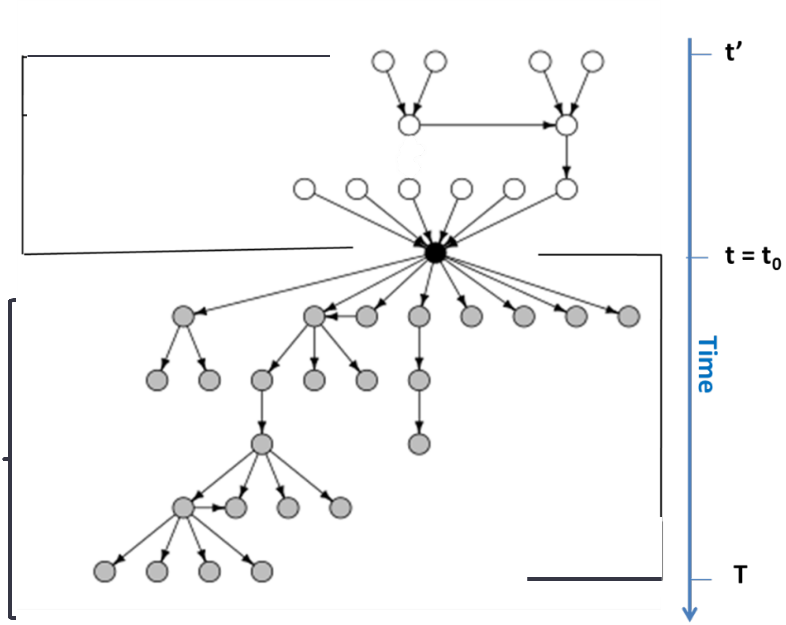
- The explored data are represented by directed, weighted and time-varying networks.

- Basics for directed weighted networks $W = \{w_{ij}\}$, $w_{ij} \geq 0$ the no. of mvts. $i \to j$ (unweighted $A = \{a_{ij}\}$, $a_{ij} = 1$ if $w_{ij} > 0$ and 0 otherwise)
  - *degree* ($k$): $k_i^{out} = \sum_j a_{ij}$, $k_i^{in} = \sum_j a_{ji}$
  - *strength* ($s$): $s_i^{out} = \sum_j w_{ij}$, $s_i^{in} = \sum_j w_{ji}$
  - *betweenness centrality*: $C_B(i) = \sum_{j \neq i \neq l} \sigma_{jl}(i)/\sigma_{jl}$, where $\sigma_{jl}(i)$ is the no. of shortest paths between $j$ and $l$ passing through $i$, and $\sigma_{jl}$ is their total no.

- Economic-like descriptors
  - *activity (flow-share)* ($x$): $x_i = (s_i^{in} + s_i^{out})/\sum_j(s_j^{in} + s_j^{out})$
    by definition, $0 \leq x_i \leq 1$ and $\sum_i x_i = 1$; for given $\delta_1, \delta_2 > 0$ agents $i$ for which $x_i < \delta_1$, $x_i \in [\delta_1, \delta_2]$ and $x_i > \delta_2$ are denoted *nichers*, *followers* and *leaders*
  - *flow polarity* ($\delta$): $\delta_i = (k_i^{in} - k_i^{out})/(k_i^{in} + k_i^{out})$
    by construction $-1 \leq \delta_i \leq 1$; for a given $\epsilon > 0$ agents $i$ for which $\delta_i < -\epsilon$, $\delta_i \in [-\epsilon, \epsilon]$ and $\delta_i > \epsilon$ correspond to *suppliers*, *wholesalers* and *demanders*

# Some elements for network analysis (2)

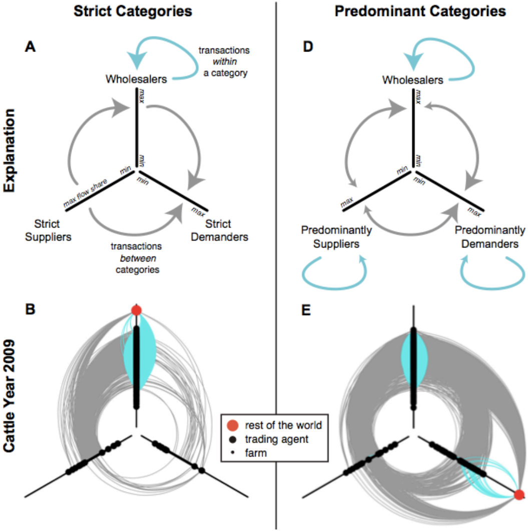
- Measures for assessing time-varying characteristics of networks
  - *short range similarity (SRS)*: $SRS_V = |V_t \cap V_{t+\Delta t}|/|V_t \cup V_{t+\Delta t}|$, $SRS_E = |E_t \cap E_{t+\Delta t}|/E_t \cup E_{t+\Delta t}|$ (where network $G_t = (V_t, E_t)$, with $V$, $E$ sets of vertices and edges; $G_t$, $G_{t+\Delta t}$ are snapshots of the same network at two consecutive time intervals)
  - *reachability ratio (RR)*: fractional size of *set of influence* (for node $i$, it includes nodes reachable from $i$ by *time respecting paths*) averaged over all nodes with out-movements
- Proxies for pathogen spread and its control
  - *giant strongly connected component (GSCC)*
  - *percolation: fractional change in the GSCC or RR size with respect to removal of nodes*


t'
t = t0
Time
T

# Distribution of flows by main categories based on flow share ($x$) and polarity ($\delta$)

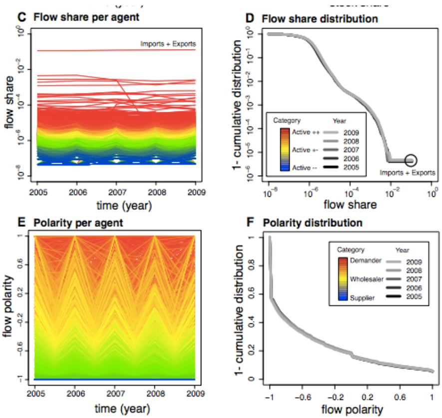
- each axis corresponds to a strict (A,B) or predominant (D,E) flow polarity ($\delta_i$)
- distribution along axes according to flow share ($x_i$)
- only links with $\geq 12$ transactions/yr are represented

Mid-wholesalers (intermediary levels of $\delta_i$ and $x_i$) play the central role in terms of flows $\rightarrow$ potentially in terms of pathogen transmission also.



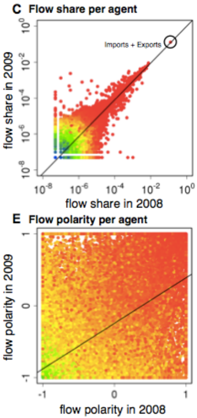

Moslonka-Lefebvre *et al.* (2014) *Submitted.*

# Stability over time: flow share ($x$) and polarity ($\delta$)




Individual trajectories of herds (all years - left, comparison 2008 vs 2009 - right)

- quite stable over time for $x$ (C, left and right)
- more variable for $\delta$, especially for wholesalers (E, left and right)

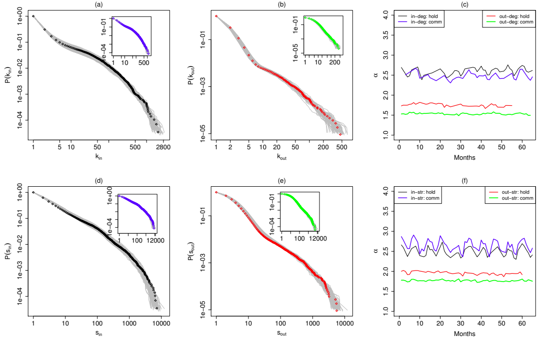



Moslonka-Lefebvre *et al.* (2014) *Submitted.*

# Topological stability over time: monthly degree ($k$) and strength ($s$) distributions (2005-2009)

No statistically significant difference (K-S test) between monthly observed distrib.

Not "far" from power-law distributions



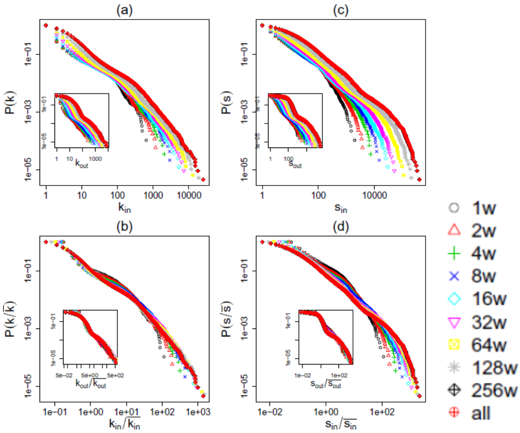

$k^{in}$ (a), $k^{out}$ (b), $s^{in}$ (d), $s^{out}$ (e); holdings as nodes (except for insets, commune as nodes); estimated exponents for $k$ (c), for $s$ (f)

Dutta *et al.* (2014) *Prev. Vet. Med.*

# Topological stability over time: degree ($k$) and strength ($s$) distributions over increased time windows (2005-2009)

For small time windows the distributions differ (K-S test, mult. testing correction).

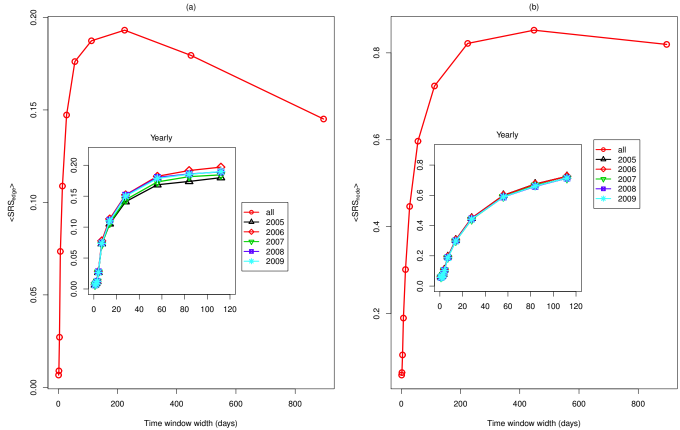
For increasing time windows distributions approach stability.




Dutta *et al.* (2014) *Prev. Vet. Med.*

# Stability over time: common bakbone of consecutive periods assessed through *SRS*

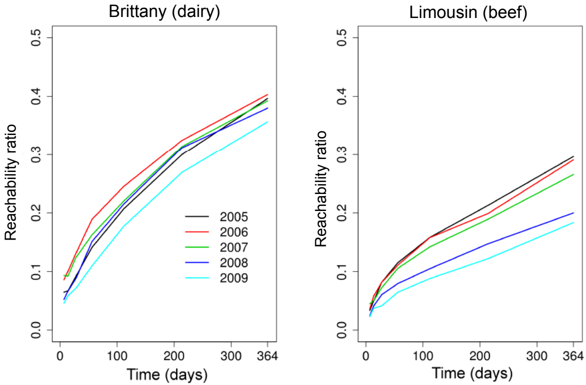
Nodes remain active over two consecutive years ($\sim$ 80%), whereas links are much more prone to variation (less then 20% common to two consecutive years).




Dutta *et al.* (2014) *Prev. Vet. Med.*

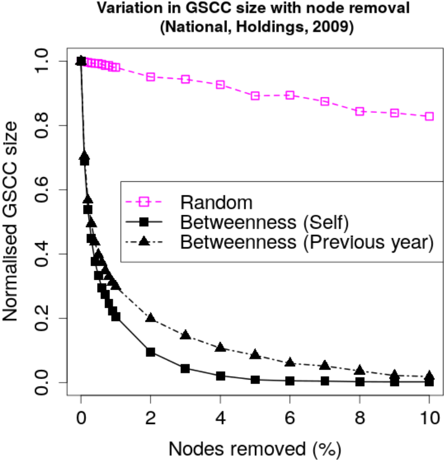
# Reachability ratio (*RR*)

Dairy network related to a potentially larger epidemic size (expressed through the proxy RR) than beef netork in the temporal formalism
(contrary to time-aggregated analysis: GSCC greater for beef than for dairy netw.)




Dutta *et al.* (2014) *Prev. Vet. Med.*

# Percolation wrt GSCC: based on degree ($k$), strength ($s$) and betweeness centrality $C_B$

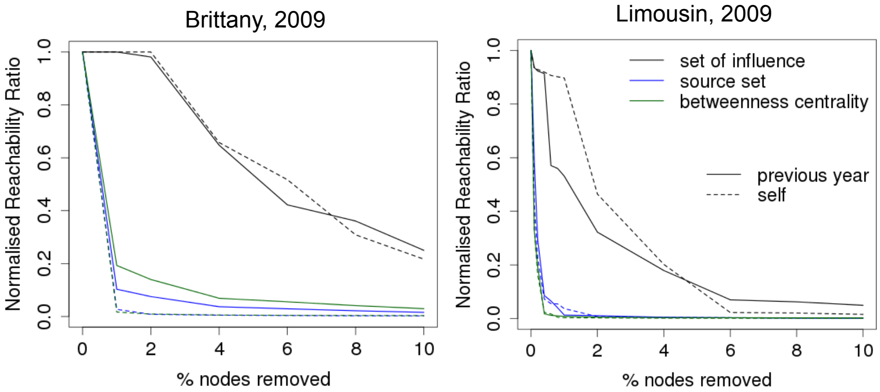



Site percolation based on nodes $C_B$ is the most efficient strategy.

High resilience to random removal o nodes (scale-free)

Dutta *et al.* (2014) *Prev. Vet. Med.*

# Percolation wrt RR: based on betweeness centrality $C_B$, set of influence, source set

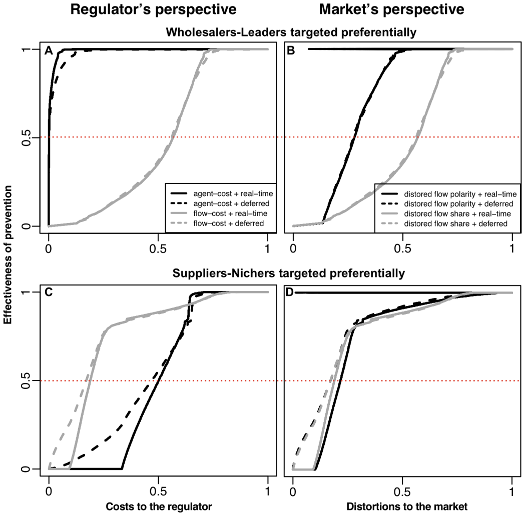



Node removal according to betweenness and source set are better strategies.

Beef region seems easier to control than dairy region.

Dutta *et al.* (2014) *Prev. Vet. Med.*

# Percolation wrt GSCC: based on flow share ($x$) and polarity ($\delta$)




Targenting the WL is more efficient than SN if the cost is proportional to the number of agents to target (A-B, black curves).

The opposite occurrs if the cost is per flow unit (A-B grey curves).

SN strategies can induce less distortions than WL ones for most levels of prevention-effectiveness.

Moslonka *et al.* (2014) *Submitted*

# Main implications of network analysis for infection spread

- Topological stability (of main indicators distributions) over time.
- Links related characteristics globally less stable over time than nodes related ones $\longrightarrow$ important variation over time of the network backbone.
- Important to consider temporal formalism when assessing epidemiological risk and control.
- Maximizing the efficacy of interventions could lead to different targeting strategies based on criteria used (e.g. cost per node - holding, vs cost per flows).
- Need of other appropriate indicators and computationally efficient programmes.