Mapping U.S. cattle shipment networks: Spatial and temporal patterns of trade communities from 2009 to 2011

Article *in* Preventive Veterinary Medicine · November 2016

DOI: 10.1016/j.prevetmed.2016.09.023

7 authors, including:

Angela D Luis
University of Montana

Michael G Buhnerkempe
Southern Illinois University School of Medicine

Daniel A Grear
United States Geological Survey

Ryan Sanders Miller
United States Department of Agriculture

Preventive Veterinary Medicine 134 (2016) 82–91

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

Mapping U.S. cattle shipment networks: Spatial and temporal patterns of trade communities from 2009 to 2011

Erin E. Gorsich[a,*], Angela D. Luis[b], Michael G. Buhnerkempe[c,d], Daniel A. Grear[e], Katie Portacci[e], Ryan S. Miller[e], Colleen T. Webb[a]

[a] Department of Biology, Colorado State University, Fort Collins, CO, USA
[b] Department of Ecosystem and Conservation Sciences, University of Montana, Missoula, MT, USA
[c] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA
[d] Fogarty International Center, National Institutes of Health, Bethesda, MD, USA
[e] USDA APHIS Veterinary Services, Center for Epidemiology and Animal Health, Fort Collins, CO, USA

ARTICLE INFO

Article history:
Received 8 April 2016
Received in revised form
27 September 2016
Accepted 27 September 2016

Keywords:
Cattle shipment
Community detection
Dynamic network
Network analysis
Interstate Certificate of Veterinary Inspection
U.S. cattle industry

ABSTRACT

The application of network analysis to cattle shipments broadens our understanding of shipment patterns beyond pairwise interactions to the network as a whole. Such a quantitative description of cattle shipments in the U.S. can identify trade communities, describe temporal shipment patterns, and inform the design of disease surveillance and control strategies. Here, we analyze a longitudinal dataset of beef and dairy cattle shipments from 2009 to 2011 in the United States to characterize communities within the broader cattle shipment network, which are groups of counties that ship mostly to each other. Because shipments occur over time, we aggregate the data at various temporal scales to examine the consistency of network and community structure over time. Our results identified nine large (>50 counties) communities based on shipments of beef cattle in 2009 aggregated into an annual network and nine large communities based on shipments of dairy cattle. The size and connectance of the shipment network was highly dynamic; monthly networks were smaller than yearly networks and revealed seasonal shipment patterns consistent across years. Comparison of the shipment network over time showed largely consistent shipping patterns, such that communities identified on annual networks of beef and diary shipments from 2009 still represented 41–95% of shipments in monthly networks from 2009 and 41–66% of shipments from networks in 2010 and 2011. The temporal aspects of cattle shipments suggest that future applications of the U.S. cattle shipment network should consider seasonal shipment patterns. However, the consistent within-community shipping patterns indicate that yearly communities could provide a reasonable way to group regions for management.

1. Introduction

Network analysis provides a conceptual framework to investigate patterns of animal movement. When networks are used to describe livestock shipments, the production units of interest are represented as nodes, and the shipment of animals between them are represented as edges (Dubé et al., 2011). Network analysis can then be used to describe features of the livestock industry (Buhnerkempe et al., 2013), evaluate the animal welfare or economic consequences of shipment practices (Hakansson et al., 2016), and study disease spread (Fèvre et al., 2006).

For a given network, shipment patterns can be better understood by considering higher-order network phenomena such as communities, which are defined as sets of nodes in the network with high levels of connections among them and low levels of connections to other nodes (Newman, 2010). While many livestock shipment networks have communities that also represent geographic regions (Lentz et al., 2011; Grisi-Filho et al., 2013), communities are properties of the shipment network (Buhnerkempe et al., 2016). As a result, these communities describe the underlying structure of the industry based on how the commodity of interest flows without imposing arbitrary geographic or administrative boundaries. These communities are most useful if they consistently describe shipping patterns and capture variability among seasons or years (Green et al., 2011). However, most communities are identified on a static network, where the data are aggregated over a

* Corresponding author at: Department of Biomedical Sciences, Oregon State University, Corvallis, OR, USA.
E-mail address: eringorsich@gmail.com (E.E. Gorsich).

http://dx.doi.org/10.1016/j.prevetmed.2016.09.023

year or multiple years (Kao et al., 2006; Green et al., 2011; Lentz et al., 2011; Grisi-Filho et al., 2013). In reality, these shipments can be dynamic over time, and considering higher resolution temporal data may result in changes to the network structure (Noremark et al., 2011; Mweu et al., 2013; Dutta et al., 2014).

A national scale, data-driven description of the U.S. livestock shipment network has recently become possible based on movement data from Interstate Certificates of Veterinary Inspection (ICVI; Buhnerkempe et al., 2013). For non-slaughter shipments of cattle across state lines, ICVIs certify that an accredited veterinarian has inspected the animal's health and that the testing requirements of the destination state are met prior to shipment. Previous analyses of livestock movement patterns in the U.S. have been based on questionnaires (Bates et al., 2001; Marshall et al., 2009; McReynolds et al., 2014) or expert opinion (Liu et al., 2012) and were smaller in scale. Thus, although ICVIs were not designed for tracing cattle movements, they are an improvement over previous descriptions of livestock shipments because they are the most comprehensive and consistently collected shipment data for the U.S. ICVIs also include the origin and destination county for the shipment as well as temporal information for the shipment (Portacci et al., 2013). This allows shipment networks to be constructed, where each county is a node in the network and the directional shipments of cattle in the ICVI data are represented as edges between them, along with temporal information to inform our understanding of temporal variability in the network structure. A basic description of a static, annually aggregated cattle shipment network based on ICVI data from 2009 is presented by Buhnerkempe et al. (2013).

In this study, we use three years of ICVI data from 2009 to 2011 to consider two alternative hypotheses for the spatial and temporal patterns of cattle shipments in the U.S. First, network structure may vary in time and communities identified on a static network from one year will be unable to describe shipment patterns in future months or years. This hypothesis describes an industry where shipment patterns are dominated by the influence of grass and feed availability, such that shipment timing and locations (as represented by network communities) respond to the price of feed and cattle. Second, network structure may be time invariant if movement patterns are dominated by the influence of fixed infrastructure. In the U.S., the feedlot-slaughter system is a particularly concentrated, spatially fixed, infrastructure that may buffer drought or economic drivers of livestock shipments. To evaluate these hypotheses, we describe the underlying structure of trade communities and build network models of data aggregated at daily, weekly, monthly, and yearly time scales to examine features of the networks that remain stable or change through time.

2. Methods

2.1. Data collection

To explore temporal variability in the U.S. cattle shipment network, we compiled ICVI data from 2009, 2010, and 2011. The 2009 ICVI data consist of a 10% systematic sample of cattle ICVI records for shipments leaving a state (Buhnerkempe et al., 2013). ICVI records are maintained and stored by the state veterinarian's office in both the state where the shipment originated and the state of destination. We requested origin ICVIs to avoid duplicated records. This dataset included all states in the U.S. with the exception of New Jersey (no response), Alaska (zero origin ICVIs), and Hawaii (zero origin ICVIs), resulting in 19,817, non-slaughter shipment records from 2433 counties in 47 states from 2009. Because our analyses address questions about the timing and consistency of interstate shipping patterns, we further excluded 713 additional shipments if the ICVIs were issued in 2009 but not sent until 2010 or if a shipment was both sent and received by the same state. For this study, we further compiled a 10% systematic sample of cattle export ICVI records using similar methods to 2009 from the following states in 2010 and 2011: California (CA), Iowa (IA), Minnesota (MN), New York (NY), North Carolina (NC), Tennessee (TN), Texas (TX), and Wisconsin (WI). These eight states were chosen to compare U.S. cattle shipment networks among years based on multiple criteria. The primary criterion for inclusion of a state in the 2010 and 2011 sampling was that states were identified as influential to the flow of cattle in 2009 based on high values for a number of network statistics such as out-degree, in-degree, and betweenness (Buhnerkempe et al., 2013). Secondary criteria stipulated that the state generated large potential outbreaks in a disease spread model (Buhnerkempe et al., 2014), allowed representation from both diverse geographic locations and locations traditionally representing a beef or dairy focus, and met additional expert opinion provided by USDA regarding the relevance of the states chosen to the U.S. beef and dairy industries. This subset of states includes 35% of operations and 36% of U.S. cattle based on summaries from the USDA National Agricultural Statistics Service (NASS; USDA, 2012).

We constructed networks by aggregating the ICVI data to the county level, such that each county represents a node in the network and each edge defines the directional shipments between nodes. Edges in the network are either unweighted or weighted by the number of shipments moving between the counties. Previous analyses have compared data aggregated at additional scales (state-level, and 50 and 500 km grid sizes). The 50 and 500 km grids generate 2350 and 46 equally sized nodes, which is roughly similar to the number of counties (3108) and states (48) in the contiguous U.S., respectively. This work suggests that a county level aggregation is the most appropriate scale because it captures heterogeneity in shipments better than coarser scales and is an administrative unit (Buhnerkempe et al., 2013). By analyzing the temporal patterns of cattle shipments, we extend the analyses in Buhnerkempe et al. (2013), where the ICVI data were aggregated across an entire year in 2009. Because the patterns of live animal transport in the beef and dairy industry are different (Bates et al., 2001), we constructed separate networks for beef and dairy shipments. However, production type was only specified in the ICVI's for 64% of shipments in 2009, 54% in 2010, and 48% in 2011. To estimate whether the remaining shipments contained beef or dairy cattle, we used a classification tree analysis to calculate the probability of unknown shipments being either beef or dairy and assigned them according to the higher probability (detailed methods and evaluation in Buhnerkempe et al., 2013).

2.2. Community detection

To identify communities within the U.S. cattle shipment network, we aggregated the 2009, 10% ICVI data into an annual network at the county scale and identified groups of highly connected counties by applying a community detection algorithm to both the unweighted network and the network weighted by the number of cattle shipments. We do not consider communities based on data from 2010 or 2011 because data from only eight states were collected for those years. Communities are formed by maximizing the modularity, Q, where

$$Q = \frac{1}{2m}\Sigma_{i,j}\left[A_{i,j} - \frac{k_i k_j}{2m}\right]\delta\left(c_i, c_j\right).$$

In this equation, $A_{i,j}$ is the weight of the edge between i and j. For the unweighted networks used in this analysis, $A_{i,j}$ is 1 if an edge exists, $k_i = \Sigma_j A_{i,j}$ is the sum of the edges attached to i, and c_i is the community to which i is assigned. The delta function is 1 if $c_i = c_j$ and

0 otherwise $m = \frac{1}{2}\Sigma_{i,j}A_{i,j}$. Modularity ranges between −1 and 1. Higher values occur when the density of edges within communities is larger than the density of edges between communities.

We applied the community detection algorithm described in Blondel et al. (2008) that detects communities in two stages, which are repeated iteratively until a maximum modularity is obtained. In the first stage, each node is assigned to its own community. The algorithm sequentially evaluates the reassignment of each node to its neighboring communities and places the node in the community with the largest gain in modularity. This process is repeated for all nodes and until no reassignments improve the modularity. In the second stage, a new network is constructed with nodes representing the communities identified in the first stage and edges representing the between-community links. The first stage of the algorithm is applied to this new, community-level network, and community memberships of all nodes are updated. The two-stage process is then repeated through several iterations until no increases in modularity are achieved. Many alternative algorithms exist to determine communities on networks (Newman, 2010). We use the algorithm described in Blondel et al. (2008) because it performs well on large networks compared to other algorithms (Lancichinetti and Fortunato, 2009). One criticism of modularity maximization algorithms is the possibility of multiple distinct solutions with high modularity scores if a clear global maximum does not exist (Good et al., 2010). To address this, we visually assessed the communities identified in each stage of the algorithm to ensure our final set of communities was robust to small changes in modularity (i.e. communities were visually similar over the later iterations of the algorithm and did not oscillate between different community structures across different iterations). We also visually compared the regions identified by the community detection algorithm with census regions defined in the USDA 2012 Census of Agriculture.

We tested for spatial clustering of the communities with join count statistics (Fortin and Drake, 2005) in the spdep package (Bivand and Piras, 2015) in R statistical software version 3.1.2 (R Development Core Team, 2014). Join count statistics test for positive or negative spatial associations of categorical data against the null hypothesis of spatial randomness. We tested for non-random association patterns for each large community (>50 counties) and used a Bonferroni correction to account for multiple comparisons. For the Bonferroni correction, we adjusted the individual confidence level upward from at least 95% confidence to at least 100(1–0.05/k)% confidence, where k is set equal to 9 for the number of communities with greater than 50 counties that were identified in the beef and in the dairy networks. Thus, significant clustering occurs when p-values are less than p = 0.0055.

2.3. Network models

We compared networks based on different temporal aggregations of the 2009 ICVI data to determine the consequences of different timescales or units on network structure. To compare a range of timescales, the data were aggregated into one yearly network; 13, 28-day (roughly monthly) networks; 52 weekly networks; and 365 daily networks (Table 1). We refer to the 28-day networks as monthly networks throughout the text. We chose to investigate these particular timescales as they are natural units of general interest. For each network, we calculated five node-level network measures and five measures of network structure (Dubé et al., 2011). The node-level network measures were:

(1) in-degree—number of unique counties that sent at least one shipment to the county in question,
(2) out-degree—number of unique counties that received at least one shipment from the county in question,
(3) weighted in-degree—number of shipments received by the county in question,
(4) weighted out-degree—number of shipments sent from the county in question,
(5) betweenness—the number of shortest paths between any two counties that go through the county in question.

Table 1
Network and node-level properties for different temporal aggregations of the beef cattle movement network created with the 2009 ICVIs. Mean values are displayed for daily, weekly and monthly networks. Total values reflect the sum of incoming and outgoing values (e.g. total shipments is the total of incoming and outgoing shipments).

	daily	weekly	monthly	yearly
number (no.) of nodes	72.0	376.3	890.0	2283
no. of edges	43.1	297.6	1115.1	10934
diameter	1.8	4.5	14.0	14
GSCC	1.0	1.5	16.0	1247
GWCC	6.8	137	711.9	2255
mean out-degree	0.6	0.8	1.2	4.8
maximum (max.) out-degree	3.9	9.4	20.1	100
mean in-degree	0.6	0.8	1.2	4.8
max. in-degree	2.4	9.6	28.3	154
mean no. total shipments	0.6	0.8	1.4	7.0
max. no. total shipments	4.0	10.0	27.4	262
mean no. total cows	45.0	60.7	103.0	524
max. no. total cows	629.5	1610.8	3626.2	29172

The five measures of network structure included:

(1) Number of nodes—the number of observed counties in the network. We use the number of nodes as a measurement of network size
(2) Number of edges—the number of unique origin and destination pairs that were involved in at least one cattle shipment,
(3) Diameter— the maximum number of steps in the set of shortest paths between all county pairs,
(4) Giant strongly connected component (GSCC)—the largest set of counties that have bi-directional shipments between any counties in the set, and
(5) Giant weakly connected component (GWCC)—the largest set of counties that are accessible to each other regardless of the direction of the edges between them.

We calculated the mean and maximum value of each node-level network measure and each measure of network structure for each temporal aggregation. We calculated network measures using the igraph package (Csardi and Nepusz, 2006) for R statistical software version 3.1.2 (R Development Core Team, 2014).

2.4. Temporal patterns in network size, network degree, and community structure

We explored temporal patterns in network size and in node-level network metrics such as in-degree, out-degree, and betweenness by identifying monthly networks from the 2009 ICVI data with small or large measures. Then, we tested for statistically significant monthly variation in each measure based on monthly networks created as described above with data from the subset of eight states available in all three years. We tested for statistical differences in network size among months using a Friedman's test with year as a repeated measure. We also tested for statistical differences in node-level metrics using a Friedman's test but with county as the unit of repeated measure because out-degree, in-degree, and betweenness are all calculated at the county level. The Friedman's test is a repeated measures nonparametric test that compares the ranks of observations to test for significant differences among groups (Conover, 1999). We also tested for monthly variation in network size and node-level metrics in each state individually, using the Bonferroni correction for multiple comparisons. For the Bonferroni correction, we have adjusted the individual confidence level upward to at least 100(1–0.05/k)% confidence. Because 8 states were compared (k = 8), significant differences among months occur when p-values are less than p = 0.00625. Statistical analyses associated with annual variation in network size and node-level metrics are described in Appendix C.

To examine how consistent communities were through time, we evaluated how well the communities identified on the annual network from 2009 described monthly shipment patterns from 2009. For each month and each community, we calculated the proportion of within community shipments, calculated as the number of shipments that were both sent and received by counties in a community divided by the total number of shipments sent by counties in a community. To evaluate how well communities identified on the annual network from 2009 describe monthly shipments from the same year, we calculated the proportion of within community shipments for each month based on shipments from the 47 states available in 2009. To evaluate how well communities identified on the annual network from 2009 describe shipments in 2010 and 2011, we calculated the proportion of within community shipments in 2010 and 2011 based on the subset of eight states available in all three years. We also applied the community detection algorithm to each monthly network in 2009. All analyses were conducted separately for beef and dairy networks to determine if the two production types have different movement patterns.

3. Results

3.1. Community detection and spatial patterns of the 2009 cattle shipment network

We identified communities in the U.S. cattle shipment network based on 16,054 beef shipments and 3050 dairy shipments that occurred in 2009. The community detection algorithm identified 26 different communities within the unweighted beef cattle network (Fig. 1a). There were seven large communities of 572, 387, 336, 269, 229, 192 and 120 counties; two additional communities of 63 and 55 counties; and 17 smaller communities of two to 24 counties in size. The circular visualization in Fig. 1a summarizes the shipment network. The outer circle represents the nine communities colored in the map and the colored flows within the circle represent the volume of shipments moving between each community. The colors within the circle indicate the community from which a shipment originated. Thus, because the flows are primarily within the same color, beef shipments occurred primarily within each community. Sixty-seven percent of shipments were sent and received within the same community while 33% of shipments were sent to other communities (Appendix A; Table A1a).

We identified 41 different communities within the unweighted dairy cattle network, nine of which consisted of more than 50 counties (Fig. 1b). Seventy-three percent of dairy shipments were sent and received within the same community while 27% percent of dairy shipments were sent to other communities (Appendix A; Table A1b). For both the beef and dairy networks, similar communities were identified on weighted (Fig. 1a and b) and unweighted networks (Appendix A; Fig. A1). One exception to their similarity is that the light blue beef community identified on the unweighted network consists of two, smaller communities on the weighted beef network. The geographic regions identified by the community detection algorithm do not follow the census regions defined in the USDA 2012 Census of Agriculture (Fig. 1c). In the circular visualization in Fig. 1c, the outer circle represents the USDA 2012 Census of Agriculture regions. The flows between census regions show that many shipments moved between the census regions. Only 39% of beef shipments were sent and received within the same census region while 48% of dairy shipments were sent and received within the same census region.

The communities identified on the beef and dairy cattle shipment networks do not visually follow state lines but generally appear spatially contiguous as determined by join count statistics. Geographic clustering of the communities identified on the beef cattle shipment network (Fig. 1a) occurs in the western (purple, $Z = 28.66$, $p < 0.0001$), central-north (light blue, $Z = 47.78$, $p < 0.0001$), central-south (light green, $Z = 27.36$, $p < 0.0001$), upper mid-west (dark blue, $Z = 13.50$, $p < 0.0001$), northeastern (orange, $Z = 12.53$, $p < 0.0001$), and in the two mid-western/central communities (dark pink, $Z = 16.62$, $p < 0.0001$; light pink, $Z = 28.66$, $p < 0.0001$). The dark green community was not geographically clustered after accounting for multiple comparisons ($Z = 2.78$, $p = 0.007$, Bonferroni corrected cutoff $p = 0.0055$). The yellow community was geographically clustered ($Z = 5.82$, $p < 0.0001$), but it is spread throughout the southeast and mid-west portions of the country. As a result, the west (purple), central-north (light blue), central-south (light green), northeast (orange) and mid-west/central (light pink, dark pink, dark blue) portions of the United States are largely distinct in terms of cattle shipments from each other while the states in the southeast have counties belonging to many different communities. Geographic clustering of the communities identified on dairy cattle shipment networks occurred in the orange, light pink, light green, light blue, purple, dark green, and dark pink communities (Fig. 1b: orange, $Z = 15.98$, $p < 0.0001$; light pink, $Z = 13.53$, $p < 0.0001$; light green, $Z = 9.26$ $p < 0.0001$; light blue, $Z = 15.20$, $p < 0.0001$; purple, $Z = 6.57$, $p < 0.0001$; dark green = 8.72, $p < 0.0001$; dark pink, $Z = 4.41$, $p = 0.0001$). The yellow and dark blue communities were not clustered (yellow, $Z = 1.35$, $p = 1$; dark blue, $Z = 2.51$, $p = 0.120$).

3.2. Consequences of different temporal aggregations of the data on networks

Different temporal aggregations resulted in large changes in both beef and dairy cattle shipment networks constructed from the 2009 ICVI data (beef: Table 1; dairy: Appendix B; Table B1). Network size dropped as we aggregated the networks over smaller time scales (Fig. 2a). For beef shipments, an average of 39% of counties in the 2009 data sent or received shipments in a given month, and only 16% of counties sent or received shipments in a given week. Similarly, for dairy shipments, only 24% of counties in the yearly data sent or received shipments in a given month and only 8% of counties sent or received shipments in a given week. The number of connections (edges) between counties and the size of the GSCC and GWCC also decreased with smaller temporal aggregations (beef: Table 1; dairy: Appendix B; Table B1).

3.3. Seasonal and annual variation in network size

There were two seasonal peaks in the network size (number of counties sending or receiving shipments) of both beef and dairy networks in 2009. The largest monthly networks were in April and October with 1003 and 1103 active beef counties and 304 and 324 active dairy counties, respectively. The smallest monthly networks occurred in July and December with 719 and 675 active beef counties and 221 and 140 active dairy counties, respectively. Based on the subset of states whose shipments were observed in all years, this seasonal variation in network size was consistent in 2009–2011; network size taken across years varied significantly by month (Fig. 2b, beef, Friedman $\chi^2 = 28.25$, $p = 0.005$; Appendix

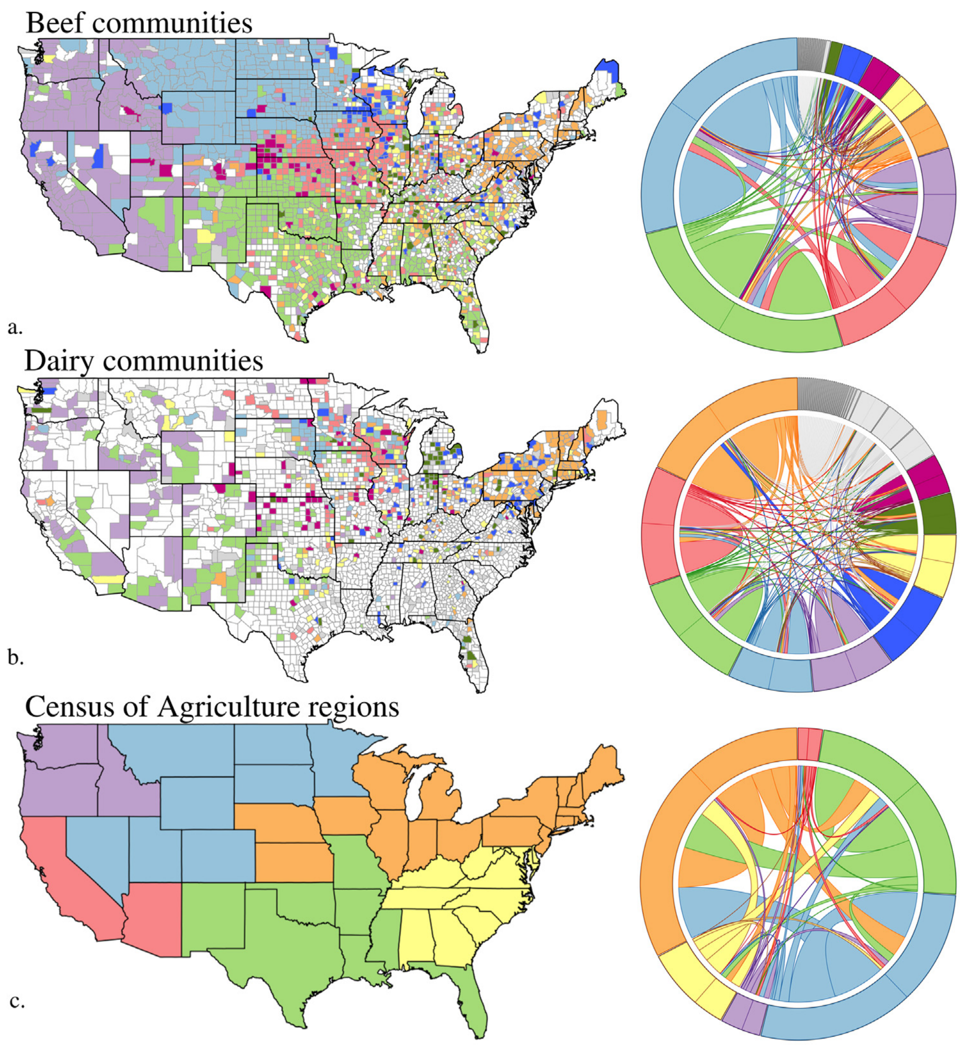

Fig. 1. Spatial patterns of the beef and dairy movement networks. Counties are colored by their community identified by the community detection algorithm on the (a) 2009 annual unweighted beef network and (b) 2009 annual unweighted dairy network. The nine largest communities (>50 counties) are colored in each map. Counties colored in gray belong to smaller communities while counties colored in white were not present in the network. In the circular visualization, the outer circle represents the communities in the map; the size of the each community is scaled in association with the number of shipments it sends or receives. Moving clockwise around the circle, the thin band of color on the outer circle divides the shipments entering and leaving a community. Connections within the circular visualization represent shipments between communities and are also colored by the community they are leaving. (c) States in the map and communities in the circular visualization are colored by their 2012 USDA Census of Agriculture region. Connections within the circular visualization represent beef shipments.

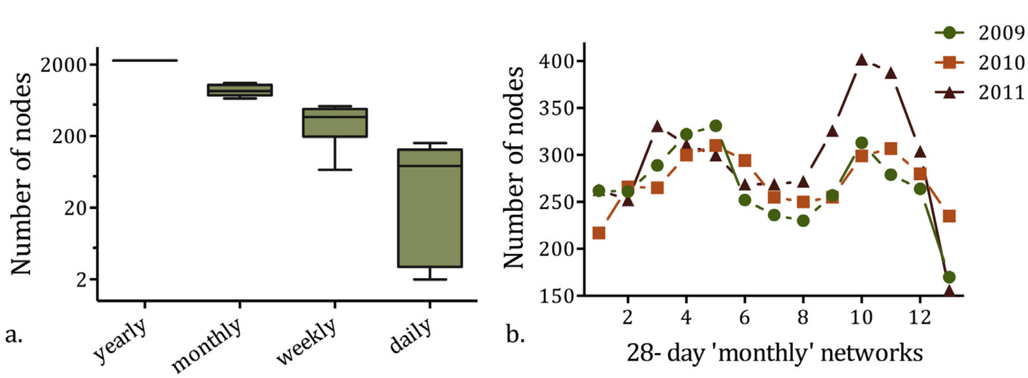

Fig. 2. (a) The number of nodes present in beef networks constructed at different temporal aggregations shows that network size decreases at smaller timescales. The boxplots display the median and interquartile range; each box is bounded by the first quartile and third quartiles of each distribution. The boxplot's whiskers extend to 1.58 times the ratio of the interquartile range divided by the square root of the sample size. (b) Network size displayed seasonal patterns that were consistent among years. For comparison among years, the beef networks displayed consist of the eight states represented in each year (California, Iowa, Minnesota, New York, North Carolina, Tennessee, Texas, and Wisconsin).

C, Fig. C1, dairy, $\chi^2 = 23.06$, $p = 0.027$). When each state was considered independently, there was evidence for variation among months in the number of counties sending or receiving beef shipments in California and Iowa (Appendix C; Fig. C2). However, monthly variation in the size of the beef network was not driven by California or Iowa alone. Significant differences in network size among months remain after excluding beef shipments from California ($\chi^2 = 27.88$, $p = 0.006$), from Iowa ($\chi^2 = 29.44$, $p = 0.004$), and from both California and Iowa ($\chi^2 = 28.74$, $p = 0.004$), suggesting monthly variation in network size is driven by small increases and decreases in the number of active counties in each state. There was no significant variation among months in the number of counties sending or receiving dairy shipments when each state was considered independently (Appendix C; Fig. C3). When comparing the overall number of active counties among years, there was no significant difference for beef networks while dairy networks had more active counties in 2009 compared to later years (Appendix C; Fig. C1). Appendix C summarizes the overall and state specific analyses of annual variation on both beef (Table C1) and dairy networks (Table C2).

3.4. Seasonal and annual variation in node-level metrics

The temporal and spatial patterns of in-degree (Fig. 3a–c), out-degree (Fig. 3d–f), the number of shipments, and the number of cattle shipped remained relatively constant within 2009. There were slight decreases in out-degree during the summer months that are driven by changes in network size and changes in the distribution of out-degree among counties (Appendix C; Fig. C4). Seasonal changes in betweenness were more pronounced (Fig. 3g–i). Betweenness for some counties was high in months 1, 2, 4, and 11 and low in the remaining months (See Appendix C for spatial patterns of in-degree, out-degree, and betweenness for monthly data aggregations for beef and dairy networks in Fig. C5a –d, and Fig. C6a ––b). There were also minimal differences in the temporal patterns of shipments among years (Appendix C; Fig. C7, Table C1, Table C2).

3.5. Stability of community structure among months in 2009–2011

In any given month in 2009, 62–77% of beef shipments leaving counties in the light blue community were also received by that community; 69–81% of beef shipments leaving counties in the light green community were also received by that community. In a given month, 52–65, 48–71, and 41–67% of shipments leaving the light pink, purple, and orange communities, respectively, were also received by that community (Fig. 4a). The other large, annual communities also consistently captured monthly movements in 2009. Based on data from 2010 and 2011, the communities identified on the annual network from 2009 also described beef shipments from 2010 and 2011 (50–54, 65–66, 49–50, 60–61, 43–45% of shipments that were sent and received in 2010 and 2011 were within the light blue, light green, light pink, purple, and orange communities, respectively, where community identity was based on the 2009 annual network).

The communities identified on the annual dairy network also capture a large proportion of the monthly dairy shipments in 2009, 2010, and 2011. In any given month in 2009, 58–95% of dairy shipments leaving counties in the light blue community were also received by that community; 64–92% of dairy shipments leaving counties in the light green community were also received by that community. In a given month, 50–89, 65–87, and 59–84% of shipments leaving the light pink, purple, and orange communities, respectively, were also received by that community (Fig. 4b). The communities determined from the 2009 annual network also described dairy shipments from 2010 and 2011 (41–48, 56–63, 44–45, 57–59, and 44–47%, of shipments that were sent and received in 2010 and 2011 were within the light blue, light green, light pink, purple, and orange communities, respectively).

In Fig. 5, the outer circles represent the communities identified on the annual beef network, while the flows between them represent the number of shipments in each month. The consistency across months in the relative sizes and outflows of the communities shows that the flow of shipments among counties identified in the annual communities was relatively stable over time, although some tradeoffs between the blue and green communities are observed (Fig. 5). This indicates that there is temporal variation in which counties are shipping in a given month. When the community detection algorithm was applied to monthly networks, this additional heterogeneity within the annual communities becomes apparent. Because not all counties are active in a given month, the communities identified in the monthly networks were smaller than the communities identified on the yearly network. The average size of a monthly community was 10 counties, and the maximum size of a monthly community was 60 counties. Furthermore, each monthly community represents a smaller geographic area compared to the annual communities and, therefore, captures which subsets of the annual communities are shipping in a given month (Appendix C; Fig. C8).

4. Discussion

The U.S. cattle shipment network consists of trade communities that consistently describe monthly and annual shipping patterns. The consistent community structure suggests that shipment locations are driven more by spatially fixed elements, such as infrastructure, compared to local economic or environmental conditions. We also identified seasonal variation in the overall number of cattle shipments. Seasonal shipping patterns were largely consistent between years and independent of the fundamental structure of the network. Taken together, these results support a national-scale description of the U.S. cattle industry where seasonal prices of feed and cattle influence the timing of shipments within a spatially fixed backbone of the feedlot and slaughter system.

4.1. Temporal patterns in the U.S. cattle shipment network

We compared the properties of beef and dairy networks created with data from 2009 aggregated over one year, one month, one week, or one day. Our analyses show that the temporal scales at which data were aggregated influenced the size and connectance of both beef and dairy cattle shipment networks (beef: Table 1; dairy: Appendix B; Table B1). On average, only 39% of counties in the yearly beef data and only 24% of counties in the yearly dairy data send or receive shipments in a given month. Previous studies of livestock shipment networks in the United States and Europe have noted similar consequences of reduced temporal aggregations (Noremark et al., 2011; Rautureau et al., 2011; Mweu et al., 2013; Dutta et al., 2014; Grear et al., 2014). Determining the appropriate time scale over which to aggregate contact data is an important consideration for the study of movement in general, but especially for the study of disease spread, which depends on both the transmission dynamics of the pathogen and the question of interest (Craft, 2015). For example, the data required to create a relevant contact network for the dynamics of pathogens with short infectious periods, such as foot and mouth disease (FMD) should be aggregated over days or weeks. Conversely, the data required for pathogens with a longer infectious period, such as bovine tuberculosis should

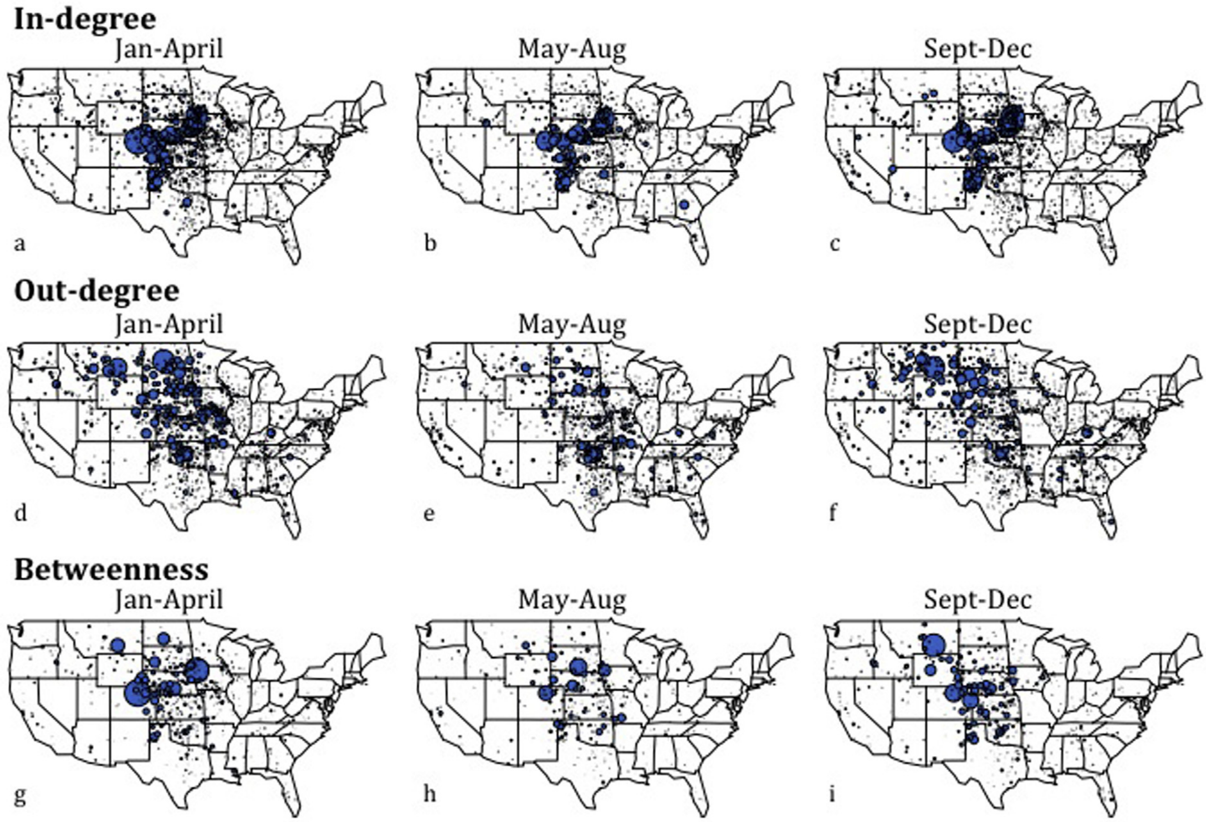

Fig. 3. Seasonal and spatial patterns for beef networks for 2009. The data were lumped into 4-month time periods for visualization. Circle sizes reflect the (a) in-degree in Jan–April, (b) in-degree in May–August, (c) in-degree in September–December, (d) out-degree in Jan–April, (e) out-degree in May–August, (f) out-degree in September–December, (g) betweenness in Jan–April, (h) betweenness in May–Aug, (i) betweenness in September–December.

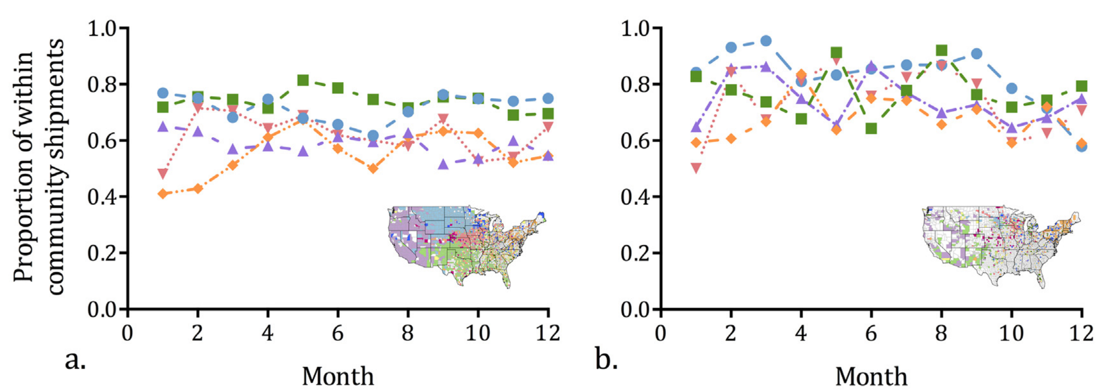

Fig. 4. Stability of communities in the (a) beef and (b) dairy network. For each month, the communities identified on the annual network explain a large proportion of shipments that occur monthly. The light blue, light green, purple, pink and orange annual communities are displayed according to their color in Fig. 1. The y-axis is the proportion of monthly shipments leaving counties within the annual community that are received by counties within the same annual community.

be aggregated over a longer timescale to represent the effective network each pathogen faces.

Our analyses also identified two seasonal peaks in network size that are consistent with conventional wisdom about the national-scale patterns of livestock shipments in the beef industry (Shields and Mathews, 2003). Beef calves are thought to move from calf/cow areas to cool-season pastures in the fall and then to feedlots or summer pastures (Shields and Mathews, 2003). This shipment pattern is consistent with the increases in network size observed in April and October for both beef and dairy networks. We also observed minimal variation among months and years in the mean number of outgoing shipments per active county, reflecting that most counties send and receive few shipments. Months with high network size had both higher numbers of well-connected counties and higher numbers of counties with few connections. This pattern suggests that both small operations with few connections and large operations with many connections are seasonally active and contribute to the seasonal patterns in network size. By highlighting the importance of seasonality, this result can provide information for future disease risk assessments and economic models that explicitly incorporate network structure.

4.2. Spatial patterns in the U.S. cattle shipment network

We identified nine beef communities and nine dairy communities consisting of over 50 highly connected counties based on cattle

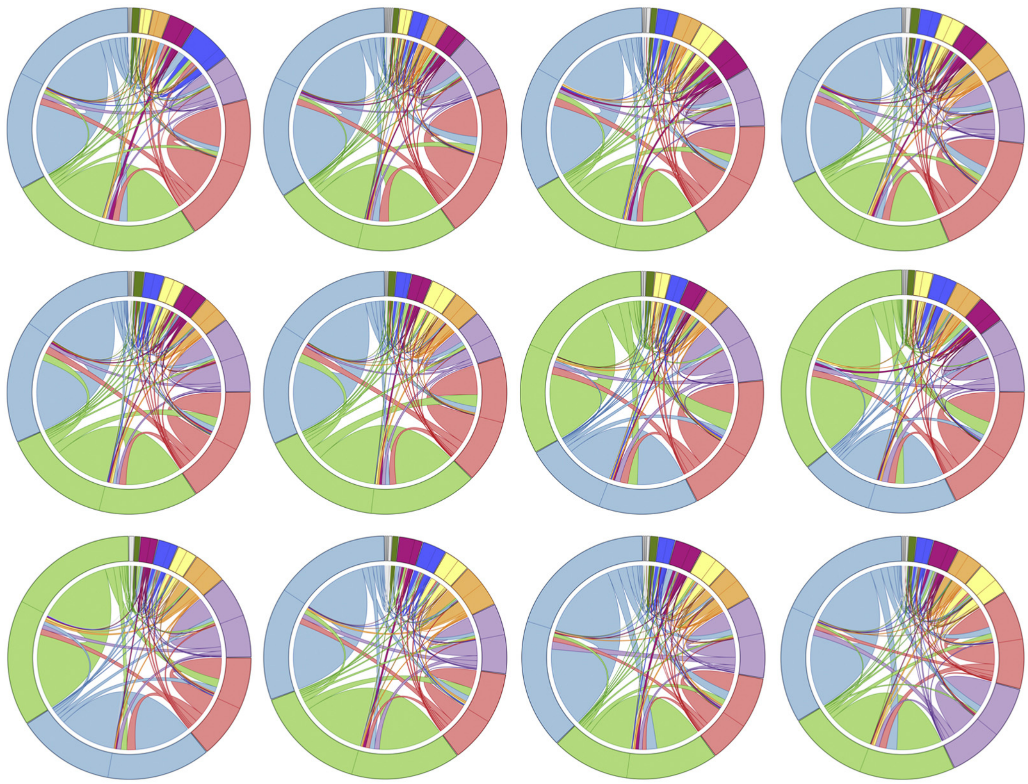

Fig. 5. Circular representations of the 12 monthly beef networks colored by the communities identified on the annual beef network (Fig. 1a). Months are ordered from left to right. The size of each community is scaled in association with the number of shipments it sends or receives in each month. The circular visualization is interpreted as in Fig. 1.

shipments. Eight of these large beef communities and seven dairy communities also represented geographically clustered groups of counties that did not follow the regions defined in the 2012 USDA Census of Agriculture. Spatially defined communities with respect to livestock shipments have also been identified in the U.K. (Kao et al., 2006), Germany (Lentz et al., 2011), and Brazil (Grisi-Filho et al., 2013). However, the U.S. communities appear less clustered than the distribution of communities elsewhere, perhaps due to the central role the feedlot system plays in the U.S. cattle industry (Shields and Mathews, 2003).

The stability of shipment patterns across different monthly aggregations in 2009, 2010, and 2011 is also encouraging for potential surveillance and emergency preparedness purposes. We note, however, that not all counties are active in a given month. Monthly networks had on average 10% of the connections in the yearly network (Fig. 2a). Previous work has shown that community detection algorithms are sensitive to the average number of contacts in the network (Lancichinetti and Fortunato, 2009). Because there are fewer connections in the monthly networks, we expect the strongly connected components to be smaller and identify additional heterogeneity within the annual communities. Despite this heterogeneity, the communities identified on the annual network in 2009 consistently explained 41–81% of shipments in a given month of the same year and 43–66% of shipments in future years. This indicates that although there is some restructuring of which sections of the annual communities are active in a given month (Fig. 3), the flow of shipments among the counties identified in the annual communities was relatively stable (Fig. 5). Additional, local-scale analyses examining the timing and frequency of existing trade connections within a community could also be conducted with this dataset.

Notably, our dataset covers the 2009 and 2011 drought periods in Texas, allowing us to evaluate features of the network that remain stable or change in response to drought. Shipment volume was more variable among years than shipment locations. More beef cattle shipments left Texas in 2011 compared to other years (Appendix C), but shipments remained within the light green community, which includes Texas (the light green community in 2009 explained 72% of shipment patterns compared to 65% in 2010, and 66% in 2011). Because we did not observe changes in community structure in response to the 2011 drought, we conclude that national-scale shipment locations are driven less by the consequences of drought and economic conditions. This consistent community structure can have applications for risk-based disease surveillance strategies by describing sites that regularly ship to each other and identifying areas that are connected through livestock shipments. Moving potentially infectious animals between herds through the shipment of animals is a primary risk factor for long-distance transmission and the rapid dissemination of a potential infection (Fritzemeier et al., 2000; Greger, 2007). Therefore, communities can provide information for disease control zones that balance the benefits of disease control with the expected economic costs from blocked trade (Scott et al., 2006), with the caveat that communities may change over longer time periods than that investigated in this study.

4.3. Caveats and interpretation

The data used to construct networks can be considered partial in two ways. First, observation of shipments, or edges, was based on a 10% systematic sample of cattle export ICVI records. Missing data may underestimate edges between counties that rarely ship to each other. However, unpublished data thinning studies indicate that a 10% sample is sufficient to recreate major network characteristics, and node-level summary statistics from networks based on the 10% sample of ICVIs aggregated across 2009 were correlated with key features of the industry such as the total cattle inventory and number of feeder cattle (Buhnerkempe et al., 2013). Second, ICVI data only capture shipments between states and do not describe shipment patterns below the state-level. These observation biases are critical for epidemiological modeling because the lack of short distance or rare shipments will underestimate disease spread (Lindstrom et al., 2013; Buhnerkempe et al., 2014). It is encouraging, however, that ICVIs capture information on short-distance movements (Buhnerkempe et al., 2013) that can be used to predict within-state movements (Lindstrom et al., 2013). Additionally, we argue that these limitations do not prohibit comparisons among seasons and years.

5. Conclusions

A quantitative and national-scale description of cattle shipments is a first step towards unraveling the infrastructure and economic forces shaping cattle shipments in the U.S. Our analyses describe a relatively stable network with minimal variation in community structure but seasonal changes in the overall number of cattle shipments. These results support descriptions of a cattle industry dominated by the influence of a fixed feedlot-slaughter system with seasonal changes in the volume of shipments moving between largely consistent locations. These results suggest that decisions based on communities identified on the annual network can be used to guide economic or disease surveillance suggestions applied to additional timescales. In addition, the temporal patterns uncovered here could be included in future network modeling approaches to understand the economic benefits or disease risks associated with shipments from different geographic areas and seasons.

Acknowledgements

Funding was provided by the Foreign Animal Disease Modeling Program, Science and Technology Directorate, U.S. Department of Homeland Security (Grant ST-108-000017) and a USDA Cooperative Agreement 13-9208-0345-CA. MGB and CTW were supported by the Research and Policy for Infectious Disease Dynamics (RAPIDD) program of the Science and Technology Directorate, Department of Homeland Security, and Fogarty International Center, National Institutes of Health; through interagency agreement #HSHQDC-09-X-00135. Data were provided by the U.S. Department of Agriculture, Animal and Plant Health Inspection Service, Veterinary Services. However, the analyses, views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the regulatory opinions, official policies, either expressed or implied, of the USDA-APHIS-Veterinary Services or the U.S. Department of Homeland Security. We also thank the state veterinarians and staff whose cooperation and effort made the data collection possible.

Appendices A–C. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.prevetmed.2016.09.023.

References

Bates, T.W., Thurmond, M.C., Carpenter, T.E., 2001. Direct and indirect contact rates among beef dairy, goat, sheep, and swine herds in three California counties, with reference to control of potential foot-and-mouth disease transmission. Am. J. Vet. Res. 62, 1121–1129.

Bivand, R., Piras, G., 2015. Comparing implementations of estimation methods for spatial econometrics. J. Stat. Softw. 63, 1–36.

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theory E, P10008.

Buhnerkempe, M.G., Grear, D.A., Portacci, K., Miller, R.S., Lombard, J.E., Webb, C.T., 2013. A national-scale picture of U.S. cattle movements obtained from Interstate Certificate of Veterinary Inspection data. Prev. Vet. Med. 112, 318–329.

Buhnerkempe, M.G., Tildesley, M.J., Lindstrom, T., Grear, D.A., Portacci, K., Miller, R.S., Lombard, J.E., Werkman, M., Keeling, M.J., Wennergren, U., Webb, C.T., 2014. The impact of movements and animal density on continental scale cattle disease outbreaks in the United States. PLoS One 9, 10.

Buhnerkempe, M.G., Webb, C.T., Merton, A.A., Buhnerkempe, J.E., Givens, G.H., Miller, R.S., Hoeting, J.A., 2016. Identification of migratory bird flyways in North America using community detection on biological networks. Ecol. Appl. 26, 740–751.

Conover, W.J., 1999. Practical Nonparametric Statistics. Wiley, New York.

Craft, M.E., 2015. Infectious disease transmission and contact networks in wildlife and livestock. Phil. Trans. R. Soc. B 370, 1669.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. Int. J. Complex Syst., 1695.

Dubé, C., Ribble, C., Kelton, D., McNab, B., 2011. Introduction to network analysis and its implications for animal disease modelling. Rev. Sci. Tech. OIE 30, 425–436.

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

Fèvre, E.M., Hamilton, K.A., Cleaveland, S., 2006. Animal movements and the spread of infectious diseases. Trends Microbiol. 14, 125–131.

Fortin, M., Drake, M., 2005. Spatial Analysis: A Guide for Ecologists. Cambridge University Press, Cambridge.

Fritzemeier, J., Teuffert, J., Greiser-Wilke, I., Staubach, C., Schlüter, H., Moennig, V., 2000. Epidemiology of classical swine fever in Germany in the 1990. Vet. Microbiol. 77, 29–41.

Good, B.H., de Montjoye, Y.-A., Clauset, A., 2010. Performance of modularity maximization in practical contexts. Phys. Rev. E 81, 046106-.

Grear, D.A., Kaneene, J.B., Averill, J.J., Webb, C.T., 2014. Local cattle movements in response to ongoing bovine tuberculosis zonation and regulations in Michigan, USA. Prev. Vet. Med. 114, 201–212.

Green, D.M., Werkman, M., Munro, L.A., Kao, R.R., Kiss, I.Z., Danon, L., 2011. Tools to study trends in community structure of live fish movements in Scotland. Prev. Vet. Med. 99, 225–228.

Greger, M., 2007. The long haul: risks associated with livestock transport. Biosecur. Bioterror. 5, 301–312.

Grisi-Filho, J.H.H., Amaku, M., Ferreira, F., Dias, R.A., Neto, J.S.F., Negreiros, R.L., Ossada, R., 2013. Detecting livestock production zones. Prev. Vet. Med. 110, 304–311.

Hakansson, N., Flisberg, P., Algers, B., Jonsson, A., Ronnqvist, M., Wennergren, U., 2016. Improvement of animal welfare by strategic analysis and logistic optimization of animal slaughter transportation. Anim. Welf. 25, 255–263.

Kao, R.R., Danon, L., Green, D.M., Kiss, I.Z., 2006. Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. J. R. Soc. Interface 273, 1999–2007.

Lancichinetti, A., Fortunato, S., 2009. Community detection algorithms: a comparative analysis. Phys. Rev. E 80, 056117.

Lentz, H.H.K., Konschake, M., Teske, K., Kasper, M., Rother, B., Carmanns, R., Petersen, B., Conraths, F.J., Selhorst, T., 2011. Trade communities and their spatial patterns in the German pork production network. Prev. Vet. Med. 98, 176–181.

Lindstrom, T., Grear, D.A., Buhnerkempe, M., Webb, C.T., Miller, R.S., Portacci, K., Wennergren, U., 2013. A bayesian approach for modeling cattle movements in the United States: scaling up a partially observed network. PLoS One 8, e53432.

Liu, H., Schumm, P., Lyubinin, A., Scoglio, C., 2012. Epirur_Cattle: a spatially explicit agent- based simulator of beef cattle movements. ICCS In: Proceedings of the International Conference on Computational Science, 9, pp. 857–865.

Marshall, E.S., Carpenter, T.E., Thunes, C., 2009. Results of a survey to estimate cattle movements and contact rates among beef herds in California, with reference to the potential spread and control of foot-and-mouth disease. J. Am. Vet. Med. Assoc. 235, 573–579.

McReynolds, S.W., Sanderson, M.W., Reeves, A., Sinclair, M., Hill, A.E., Salman, M.D., 2014. Direct and indirect contact rates among livestock operations in Colorado and Kansas. J. Am. Vet. Med. Assoc. 244, 1066–1074.

Mweu, M.M., Fournié, G., Halasa, T., Toft, N., Nielsen, S.S., 2013. Temporal characterization of the network of Danish cattle movements and its implication for disease control: 2000–2009. Prev. Vet. Med. 110, 379–387.

Newman, M.E.J., 2010. Networks: An Introduction. Oxford University Press, New York.

Noremark, M., Hakansson, N., Lewerin, S.S., Lindberg, A., Jonsson, A., 2011. Network analysis of cattle and pig movements in Sweden: measures relevant for disease control and risk based surveillance. Prev. Vet. Med. 99, 78–90.

Portacci, K., Miller, R.S., Riggs, P.D., Buhnerkempe, M.G., Abrahamsen, L.M., 2013. Assessment of paper interstate certificates of veterinary inspection used to support disease tracing in cattle. J. Am. Vet. Med. Assoc. 243, 555–560.

R Development Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org/.

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.
Scott, A., Zepeda, C., Garber, L., Smith, J., Swayne, D., Rhorer, A., Kellar, J., Shimshony, A., Batho, H., Caporale, V., Giovannini, A., 2006. The concept of compartmentalization. Rev. Sci. Tech. Off. Int. Epiz. 25, 873–879.
Shields, D.A., Mathews, K.H., 2003. Interstate Livestock Movements. United States Department of Agriculture Economic Research Service Outlook Report LDP-M-108-01.21.
USDA, 2012. Census of Agriculture. USDA National Agricultural Statistics Service. http://www.nass.usda.gov/census.