2013

A national-scale picture of U.S. cattle movements obtainedfrom Interstate Certificate of Veterinary Inspection data

Daniel A. Grear
USGS, dgrear@usgs.gov

Ryan S. Miller
USDA APHIS Veterinary Services

Jason E. Lombard
USDA APHIS Veterinary Services

Colleen T. Webb
Colorado State University, Colleen.Webb@ColoState.EDU

Michael Buhnerkempe
Colorado State University - Fort Collins, michael.buhnerkempe@gmail.com

See next page for additional authors

Authors

Daniel A. Grear, Ryan S. Miller, Jason E. Lombard, Colleen T. Webb, Michael Buhnerkempe, and Katie Portacci

Preventive Veterinary Medicine 112 (2013) 318–329

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

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

A national-scale picture of U.S. cattle movements obtained from Interstate Certificate of Veterinary Inspection data

Michael G. Buhnerkempe[a,1], Daniel A. Grear[a,*], Katie Portacci[b], Ryan S. Miller[b], Jason E. Lombard[b], Colleen T. Webb[a]

[a] Department of Biology, Colorado State University, Fort Collins, CO, USA
[b] USDA APHIS Veterinary Services, Centers for Epidemiology and Animal Health, Fort Collins, CO, USA

ARTICLE INFO

Article history:
Received 10 December 2012
Received in revised form 5 August 2013
Accepted 7 August 2013

Keywords:
Cattle transport
U.S. cattle industry
Network
Interstate Certificate of Veterinary Inspection
ICVI
Livestock health

ABSTRACT

We present the first comprehensive description of how shipments of cattle connect the geographic extent and production diversity of the United States cattle industry. We built a network of cattle movement from a state-stratified 10% systematic sample of calendar year 2009 Interstate Certificates of Veterinary Inspection (ICVI) data. ICVIs are required to certify the apparent health of cattle moving across state borders and allow us to examine cattle movements at the county scale. The majority of the ICVI sample consisted of small shipments (<20 head) moved for feeding and beef production. Geographically, the central plains states had the most connections, correlated to feeding infrastructure. The entire nation was closely connected when interstate movements were summarized at the state level. At the county-level, the U.S. is still well connected geographically, but significant heterogeneities in the location and identity of counties central to the network emerge. Overall, the network of interstate movements is described by a hub structure, with a few counties sending or receiving extremely large numbers of shipments and many counties sending and receiving few shipments. The county-level network also has a very low proportion of reciprocal movements, indicating that high-order network properties may be better at describing a county's importance than simple summaries of the number of shipments or animals sent and received. We suggest that summarizing cattle movements at the state level homogenizes the network and a county level approach is most appropriate for examining processes influenced by cattle shipments, such as economic analyses and disease outbreaks.

Published by Elsevier B.V.

1. Introduction

The most widespread data available for tracing large-scale cattle movements in the United States is the Interstate Certificate of Veterinary Inspection (ICVI), although ICVIs are not designed for this purpose. Specifically, when cattle shipments cross state lines, destination states require that most shipments must be accompanied by an ICVI, certifying that an accredited veterinarian has inspected the animals prior to shipment and they are apparently healthy with no signs of communicable diseases and that testing requirements for the destination state are met; cattle going to slaughter are a notable exception in some states. ICVIs list the origin and destination address for the shipment providing a useful source of data on interstate cattle movements. In addition to verifying destination state health requirements are met, ICVIs are the most complete source of national-level movement data. Unfortunately, the use of the ICVI system is limited by storage as paper records, incomplete data fields, and inability to rapidly retrieve the data which limit the ability to use these records for large

* Corresponding author at: 2150 Centre Avenue, Building B, Fort Collins, CO 80526, USA. Tel.: +1 970 494 7001; fax: +1 970 491 7269.
E-mail address: Daniel.A.Grear@aphis.usda.gov (D.A. Grear).
[1] Present address: Department of Ecology and Evolutionary Biology, University of California – Los Angeles, Los Angeles, CA, USA.

0167-5877/$ – see front matter. Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.prevetmed.2013.08.002

scale purposes. Although electronic systems are available for entering data on ICVIs, these systems are limited in scope and highly biased (Portacci et al., 2013).

Previous attempts to capture a snapshot of nation-wide cattle movements have tried to circumvent the accessibility issues associated with paper copies of ICVIs and focused on summary information obtained from ICVIs at the state level (Shields and Mathews, 2003). Some states keep track of which states they send and receive cattle from enabling coarse grained models and predictions of cattle movement at the state scale (Shields and Mathews, 2003), although these models are inadequate to capture spatial heterogeneities in cattle shipment patterns that arise at finer resolutions. Unfortunately, this summary information is also incomplete, with less than half of states maintaining summaries about interstate shipping partners, and provides no information on characteristics of individual shipments that could enhance understanding and interpretation of U.S. cattle movement patterns (Forde et al., 1998).

Other attempts to characterize cattle movement in the U.S. have avoided ICVIs entirely. Particularly, data on shipments of animals for feeding have been compiled to identify geographic areas serviced by cattle markets, although these data are limited to shipments from a single video auction and makes no attempt to develop a comprehensive description of national cattle movements (Bailey et al., 1995). Additionally, a study focused on movements from a limited number of counties in California was done by asking farmers directly about their direct and indirect contacts with other cattle facilities, and while the data are illuminating at a local scale, the cost and effort of expanding to a national scale characterization are immense (Bates et al., 2001). Another effort to characterize U.S. cattle movements at finer resolutions (i.e., premises level) has relied almost exclusively on models developed from expert opinion on shipment patterns at highly-local scales (Liu et al., 2012). Without comprehensive movement data to base these models on, management decisions must rely on predictions based on large amounts of unquantifiable uncertainty and unverified assumptions. Based on the limitations of other attempts at characterizing U.S. cattle movements, ICVI data may provide a better understanding of the movement within the U.S. cattle industry as a whole.

The value of ICVIs in characterizing cattle movement through the development of a network approach has not been thoroughly explored. Network models have been successful in describing cattle and sheep movements in other countries (Webb, 2005; Bigras-Poulin et al., 2006; Kiss et al., 2006; Robinson and Christley, 2007; Brennan et al., 2008; Natale et al., 2009; Vernon and Keeling, 2009; Volkova et al., 2010; Bajardi et al., 2011) and are likely to be helpful in characterizing animal movement in the United States. Also, using networks for post hoc analysis has proven useful in identifying determinants of disease spread and the importance of animal movements in disease outbreaks, most notably the 2001 outbreak of Foot-and-Mouth disease in the United Kingdom (Green et al., 2006; Ortiz-Pelaez et al., 2006; Kao et al., 2006). However, integrating movement data into disease spread models should not be limited to after-the-fact analysis but should also be used to develop predictions of risk based on animal movements. Predictions derived from these types of animal movement and disease simulation models can then be used to efficiently structure limited surveillance resources as well as determine the economic consequences of disease control strategies during an outbreak (Kao et al., 2007; Velthuis and Mourits, 2007). In addition, economic analyses of cattle shipment patterns can increase efficiency of the U.S. cattle industry as a whole through an understanding of market interactions (Bailey et al., 1995) and livestock hauling behavior and limitations (Hoffman et al., 1975).

In this study, we use ICVIs to characterize cattle shipments and to build network models that provide the first comprehensive quantitative characterization of interstate national cattle movement in the U.S. This demonstrates the utility of an ICVI beyond tracing individual animals and ensuring destination state health requirements are met. We note that ICVIs are limited to shipments that travel interstate and, thus, ICVI based inference about the cattle movement in the U.S. is also restricted to interstate cattle transport patterns. However, we evaluate the robustness of the ICVI-based networks using several independent data sources to identify any potential systematic bias that may be introduced by sampling interstate data. In addition, attributes derived from the cattle movement network can then be tied to the underlying infrastructure of the cattle industry to help guide national scale strategies on disease risk prevention, surveillance, and control, as well as economic analyses.

2. Methods

2.1. Data collection

Because ICVIs are maintained by both the destination and origin states, we requested that all states send a 10% sample (either photocopies or scans) of their calendar year 2009 cattle ICVIs for cattle shipments originating in the state by taking a systematic sample of every tenth cattle ICVI. This enabled us to obtain estimates of the total number of ICVIs for 2009 and avoid potential temporal and spatial biases that could arise by sampling the first 10% of records. We requested ICVI records from 2009 because it was the most recent complete year at the time of request. We specifically requested origin ICVIs to avoid duplication in the data due to copies of an ICVI being sent to both the destination and origin state veterinarian's office. Although destination ICVIs may have been more accurate due to a state's increased vigilance over cattle entering rather than leaving their state, certificates sent to destination states are often accompanied by numerous other forms (test results, etc.) increasing the potential burden on states to sort and ship 2009 cattle ICVIs.

We obtained ICVIs from 48 states, with the exceptions being New Jersey (did not participate) and Alaska (no ICVIs to report). In general, we obtained a 10% systematic sample of 2009 ICVIs from the shipment origin states, but modifications to this sampling design were implemented in three

Table 1
Number of Interstate Certificates of Veterinary Inspection shipments by farm type and production types determined by the decision tree analysis (Appendix C).

	Shipment purpose			
	Feeding	Breeding	Unknown	Proportion
Production type				
Beef	3970	1977	2681	0.435
Dairy	343	2308	41	0.136
Unknown	5401	666	983	0.356
Proportion	0.490	0.250	0.187	0.927[a]

[a] 1447 (7.3%) of shipments were classified as show/exhibition.

states to accommodate time and budget constraints (see Appendix A).

2.2. Data entry

A database was constructed to capture the data from the paper ICVIs, The database included: origin and destination address; dates the animals were inspected, shipped, and the ICVI received at the state veterinarian's office; the purpose of the shipment; whether the shipment was beef or dairy cattle; the number of animals; and the breeds, age, and gender distributions of the cattle in the shipment. All entered data was verified by individuals not responsible for original entry with random spot checks by a third person to ensure consistent accuracy of data entry and verification. All address data was converted to latitude/longitude coordinates using standard geocoding methodology. Data accuracy was confirmed during this process with only 0.9% of origin locations and 1.9% of destination locations failing to be assigned coordinates at least to the county-level. Subsequent data cleaning limited this number to 2 origin and 3 destination addresses that were removed. The final database contained 19,817 shipment records.

2.3. Production type classification

To see how well the sample matched standing cattle populations across the industry, each shipment was classified into a production type (e.g., show/exhibition, beef-feeding, beef-breeding, dairy-feeding, etc.). We developed a decision tree based on shipment details to classify shipments, which was subsequently confirmed by a classification tree analysis (Breiman et al., 1984; see Appendix C). Because beef and dairy shipments and feeding and breeding type shipments needed to be separated, the classification included available data on production type given on the ICVI directly as well as information on breed, purpose, age, and sex; for a fuller description of the shipment types see Table 1 and Appendix B.

2.4. Network models

Network models consist of a set of nodes that define the individual units of study and edges that describe the interactions between nodes (Dube et al., 2009). For our network models, we defined nodes at two scales. First, premises were aggregated to the county scale, a scale which is already used by the National Agricultural Statistics Service for reporting national agricultural census data (3109 potential nodes in the conterminous U.S.; USDA, 2007). Data on individual identity and characteristics of individual premises are not available in the U.S., and detailed data on the origin and destination premises (i.e., beef farm, market/abattoir, feedlot, etc.) are not required on the ICVIs. Therefore, it was not appropriate to perform any analyses or make any inferences about shipment characteristics to scales smaller than the county. Second, premises were aggregated to the state-level in the continental U.S. (48 nodes). Due to concerns over heterogeneities in state and county size across the US and their potential effect on network structure (Flowerdew et al., 2008), we also approximated a state and county network using equally-sized geographic areas as nodes. We did this by aggregating shipment locations to arbitrary regular 500 and 50 km grids generating 46 and 2350 nodes, respectively. Directed edges were defined by the shipments between nodes and assigned weights based on the number of shipments. Where indicated, the number of cattle over all the shipments between nodes was used as an alternative edge weight.

For the observed network, we calculated six metrics to describe the overall structure of the sample of interstate movements including:

(1) Diameter – a measure of network size that describes the maximum number of steps in the shortest path between any two nodes (i.e., the longest, shortest path length).
(2) Reciprocity – the proportion of edges for which there is another edge in the opposite direction (i.e., node i to j and node j to i).
(3) Transitivity – the probability that any two neighbors of a node (i.e., connected by an edge) are connected themselves (also known as the clustering coefficient).
(4) Degree assortativity – the correlation between shipment activity and the network degree of the nodes at the ends of each edge (i.e., the sum of a node's in and out shipments). This characterizes whether high activity nodes in the network generally interact with other high activity nodes (i.e., values close to one) or low activity nodes (i.e., values close to negative one).
(5) Giant strongly connected component (GSCC) – the GSCC describes the largest set of nodes for which there are bi-directional paths between any two nodes in the set.
(6) Giant weakly connected component (GWCC) – the GWCC describes the largest set of nodes for which all nodes are accessible to each other regardless of the direction of the edges between them.

We also calculated four metrics specific to a node to measure its centrality to the observed network. These included:

(1) In-degree – the number of shipments into a node.
(2) Out-degree – the number of shipments from a node.
(3) Betweenness – the number of shortest paths between any two nodes that go through the node in question.

(4) Closeness – the inverse of the number of steps required to reach every other node from a given node.

We calculated the network statistics using the igraph package (Csardi and Nepusz, 2006) for R statistical software (R Development Core Team, 2012).

2.5. Comparison with NASS data

We assessed the relationships between observed interstate shipment network patterns and the underlying infrastructure of the US cattle industry by using the 2007 NASS agricultural census as the quantitative benchmark of the U.S. cattle industry. To compare the network metrics with underlying infrastructure in the U.S. cattle industry, we calculated Pearson's correlations between our network node (county) metrics and county-level summaries of U.S. farm distributions. We use "farm" as a general term for any type of premises where cattle are traded as a commodity according to the USDA National Agricultural Statistics Service (NASS) definition: any establishment from which $1000 or more of agricultural products were sold or would normally be sold during the year (USDA, 2007).

We evaluated how well an approximation of the number of individual farms per county, identified by unique addresses, as destination and origin, in the ICVI sample matched the distribution of cattle farms per county according to the 2007 NASS agricultural census. To identify a proxy for the number of individual farms in the ICVI data, we considered each unique latitude–longitude location according to the addresses listed on the ICVIs to be a unique farm. We recognize several sources of error that lead to under or over counting farms using the address information including missing, illegible, or unusable (i.e., post office box) address information on the ICVI forms; errors in transcription of addresses into the database; and multiplicities in names and addresses for farms (Portacci et al., 2013). Therefore, we consider the number of unique addresses per county from the ICVI data to be a proxy for the true number of farms identified through the sampling of shipments and limit the inference of our analyses to evaluating the relative frequency of farms per county.

We evaluated how the distribution of farms per county, obtained from the 10% sample of ICVIs represented the 2007 NASS farm census using a generalized linear regression model with quasipoisson errors and a log-link function for the response. We used the ICVI farm proxy from each county i as the response variable (μ_i^{cvi}) with the NASS census farm count as an explanatory variable, ($NASS_i$). However, given the interstate nature of the data, we were concerned with the effect of bias in sampling along the state borders (i.e., these farms are more likely to ship across state lines due to their proximity). Thus, we also included an indicator variable of whether or not a county was adjacent to an interstate border ($border_i$) and a factor for the effect of the state, j, the county is in ($state_j$). Because we considered these main effects and their interactions, we used the model:

$$\begin{aligned}\ln(\mu_i^{cvi}) &= \beta_0 + \beta_1 NASS_i + \beta_{2,j} state_j + \beta_3 border_i \\ &+ \beta_{(1,2),j} NASS_i * state_j + \beta_{(1,3)} NASS_i * border_i \\ &+ \beta_{(2,3),j} state_j * border_i + \beta_{(1,2,3),j} NASS_i * state_j * border_i \\ &+ \varepsilon_i\end{aligned}$$

where ε_i is the quasipoisson distributed random error associated with county *i*. We selected a best model with drop-in-deviance χ^2 tests starting with the full interaction model and sequentially removing the interaction terms and then the main effect terms if the drop-in-deviance test was not significant ($p > 0.05$). From the model with minimum deviance, we evaluated β_1 to determine if and what was the relationship between the ICVI farm count proxy per county in the NASS census; $\beta_{2,j}$ to evaluate how the average ICVI farm count proxy varied by state, β_3 to evaluate if the count of farms in the ICVIs were greater in counties bordering a state line; $\beta_{(1,2),j}$ to determine how the relationship between the ICVI farm count proxy per county in the NASS census varied across states; $\beta_{(1,3)}$ to determine if there was a general state border bias in the sampling; $\beta_{(2,3),j}$ to determine if a border sampling effect was constant across states; and $\beta_{(1,2,3),j}$ to determine if a border sampling bias varied by state.

We also examined the relationship between the node-level (county) network metrics, in-degree, out-degree, betweenness, and closeness with the number of farms, cattle inventory, density of farms per km^2, and density of animals per km^2 obtained from the 2007 NASS census. For each cattle infrastructure metric, we calculated correlations for the total per county, the operations with cattle on feed, and the operations with cows that calved. Some county-level NASS cattle inventory data are withheld to protect privacy and any counties with withheld data were omitted for this analysis.

3. Results

3.1. Shipment characteristics

Interstate shipments of cattle not going to slaughter in the U.S. were primarily small shipments consisting of less than 10 animals (Fig. 1). These shipments were dominated by animals going to feed (45.1%), although breeding (17.1%), sale (11.4%), and show (6.8%) movements were also represented (Appendix B, Fig. B1). The ICVI sample contained both local and long distance cattle movements, including rare extreme long-distance movements (>3000 km; Appendix B, Fig. B2). There were no apparent differences in the distributions of movement distances by shipment type (Fig. B3). As expected with the large number of feeding shipments, most animals being shipped in the U.S. are young castrated males (<2 years old; Fig. 2). Information on production type was limited on the ICVIs with only 23.9% and 2.7% of shipments identified as beef and dairy respectively. Incorporating information on breed, purpose, age, and sex in the decision tree framework (Appendix C) increased our ability to categorize shipments by production type identifying 43.5% and 13.64% of movements as beef and dairy, respectively (Table 1), with

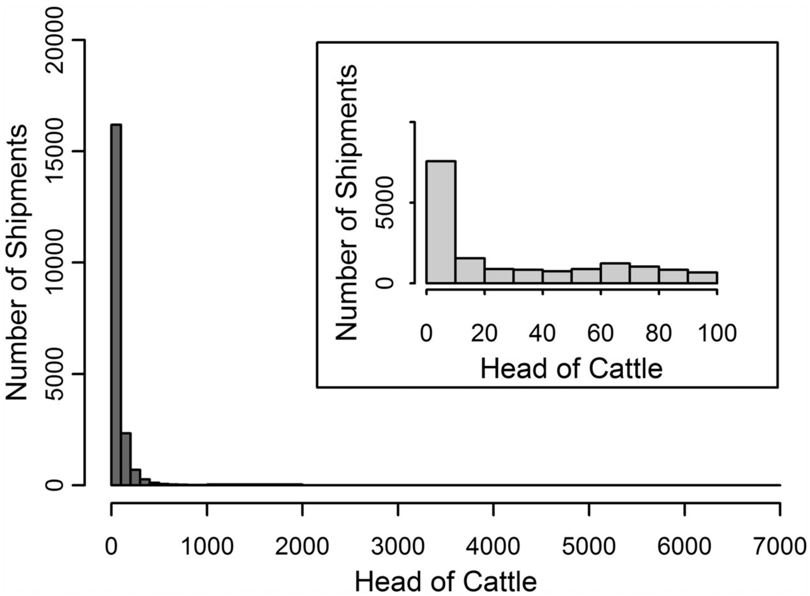

Fig. 1. Histogram of US cattle ICVI shipment sizes. The tail of the shipment size distribution drops off sharply after shipment sizes of 300, although the largest shipment was 6200 cattle. The shipments in the 0–100 size class are dominated by shipments of size 0–10 (inset).

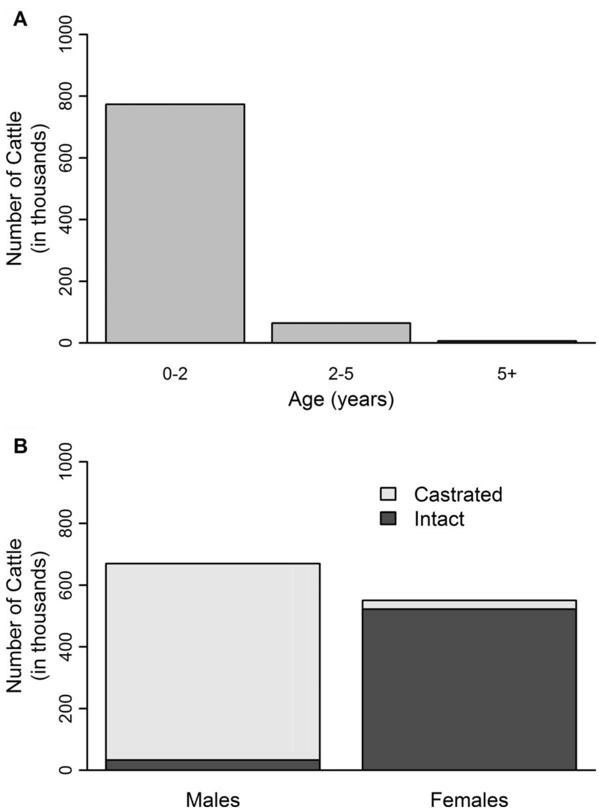

Fig. 2. Number of cattle being shipped in various (A) age classes and (B) gender classes. Gender is further broken in castrated male and females (light gray) and intact male and females (dark gray).

show/exhibition/rodeo shipments accounting for 7.3% of the shipments.

3.2. Network metrics

The network metrics for all cattle for both the county and 50 km grid scales were similar (Table 2). In particular, it took a maximum 12 movements to get across either network (network diameter). The low reciprocity in these networks indicated that return shipments among counties are rare. Also, the US cattle industry showed low clustering in general with low transitivity in both networks indicating few connections between any two neighbors. The networks tended to be mildly assortative with high degree nodes (whether in- or out-degree) shipping more often to other high degree nodes. However, when edges were weighted by the number of cattle being shipped, the assortativity in both networks became slightly negative indicating no real preferential interaction between nodes that ship a high volume of cattle. Despite this, the networks in general were relatively well connected with 64.0% and 66.5% of nodes in the county and 50 km grid networks, respectively, belonging to the GSCC (for county scale see Fig. 3). When direction is not considered, both networks were essentially one component with the GWCC encompassing over 99% of the network at both scales (for county scale see Fig. 3).

The state and 500 km grid networks were also similar (Table 2). In contrast to the county and 50 km grid scales, however, the state and 500 km grid networks were much smaller requiring only 3 steps to traverse the network. These coarse-grained networks also show a large amount of association between neighbors with high transitivity indicating clustering and over one-third of edges in the network being bi-directional. Neighbors at this scale also tend to behave similarly in terms of shipment activity as evidenced by the high assortativity, although this relationship

Table 2
Network metric describing the connectivity of the U.S. cattle industry at two spatial scales, county and state, as well as lattice-based approximations to these network nodes (i.e., 50 km and 500 km grids).

	County	50 km Grid	State	500 km Grid
Number of nodes	2436	2354	48	46
Diameter	12	12	3	3
Reciprocity	0.0282	0.0249	0.3455	0.3173
Transitivity	0.0530	0.0550	0.7687	0.7369
Assortativity	0.2021	0.1534	0.6417	0.6666
Weighted Assortativity	−0.0792	−0.0576	−0.1701	0.1222
Size of GSCC[a]	1559	1566	47	46
Size of GWCC[b]	2422	2346	48	46

[a] Giant strongly connected component: the largest set of nodes for which there are bi-directional paths between any two nodes in the set.
[b] Giant weakly connected component: the largest set of nodes for which all nodes are accessible to each other, regardless of the path direction.

again breaks down when cattle weights are considered. The strength of connection in these networks eliminates any isolated nodes in the network with all of the nodes belonging to the GWCC and only one node, New Jersey, was not included in the GSCC at the state scale due to a lack of data.

3.3. Node metrics

Owing to the similarity in our networks at the county and 50 km grid scales as well as at the state and 500 km grid scales, we will only present node metrics for the more familiar county and state scales. Results for the 50 km and 500 km grid networks can be found in Appendix D.

At the county scale, most shipments were destined for the central plains states with these regions having the highest in-degree nodes in the network (Fig. 4A). These same regions were still involved with cattle exports as well, but the nodes with the highest out-degree were shifted to the northwest (Fig. 4B). The shift in spatial distribution between in-degree and out-degree was mirrored by a relatively weak correlation between the two measures (Table 3). When movements were weighted by the number of cattle, the network showed similar patterns due to the high predictability of weighted in- and out-degree (i.e., number of cattle imported and exported) from the in- and out-degree weighted by number of shipments (Fig. 5). Betweenness centrality was similar to that of both in-degree and out-degree (Fig. 4C). Counties with high shipping activity, whether it was imports or exports, had high betweenness verified by the high correlation between both in- and out-degree and betweenness (Table 3). Conversely, closeness centrality showed no apparent spatial pattern (Fig. 4D). Across the U.S., a county's closeness was relatively constant and had a weaker relationship with in- and out-degree than betweenness centrality (Table 3).

At the state scale, most shipments were destined for the central plains states, with these states having the highest in-degree nodes in the network (Fig. 6A). In contrast to the county scale, in-shipments at the state scale are much more homogeneous across the U.S. Similar to the county network, we found a general shift in the sending centers to more northern states, but outgoing shipment behavior is still fairly homogeneous at the state scale (Fig. 6B). Although sending and receiving hubs were not identical, a state's in-degree and out-degree were highly correlated (Table 3). State-level betweenness centrality exhibited a pattern that was significantly different than in- and out-degree with New York, Texas, and Pennsylvania having the largest betweenness (Fig. 6C). At the state scale, betweenness pointed to network structure that is not well described by in- or out-degree and exhibited smaller correlation with these measures than at the county scale (Table 3). Closeness centrality, again, showed no spatial pattern (Fig. 6D).

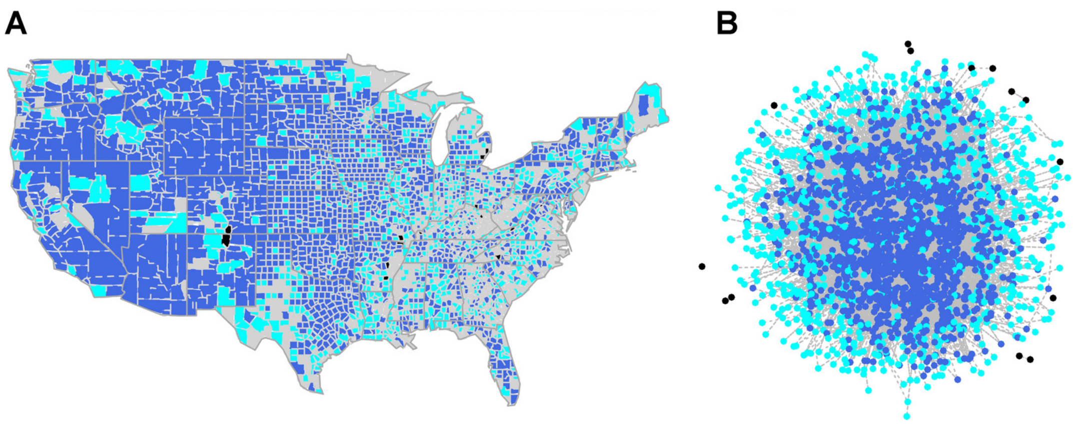

Fig. 3. County membership in both the GSCC and the GWCC (dark blue), GWCC only (light blue), or neither (black; isolated from the GWCC) shown both (A) geographically and (B) in network space. Counties for which there are no data are given in gray.

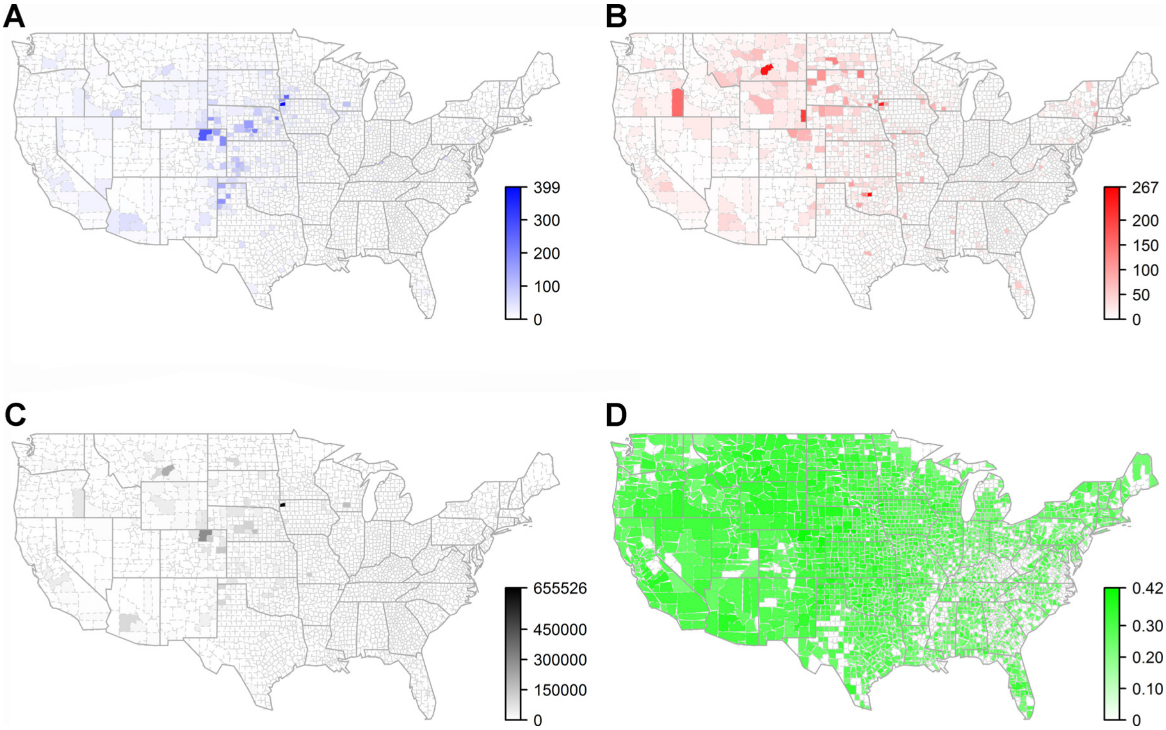

Fig. 4. Maps of (A) in-degree, (B) out-degree, (C) betweenness, and (D) closeness for the county scale network.

3.4. NASS comparison

The regression model that best explained the relationship between the proxy of farm numbers sampled from the ICVI data was the full model including the explanatory variables: NASS farm count, state, state border factor, and all interactions ($\chi^2 = 1699.6$, df = 91, $P < 0.01$). Nine states had greater than the mean ICVI farm count proxy (β_2; Table E1): Iowa, Kansas, Minnesota, Montana, Nebraska, New Mexico, Oklahoma, Oregon, and Wyoming. Five states had lower than mean ICVI farm count proxy (β_2; Table E1): Kentucky, Michigan, North Carolina, Virginia, and West Virginia. There was not an overall relationship between ICVI farm count proxy and NASS census farm count ($\beta_1 = 0.001$, 95% Confidence Interval (CI) = [−0.002, 0.003]), but a positive interaction between NASS farm count was identified in 19 states ($\beta_{(1,2)}$; Table E1): Arizona, California, Colorado, Georgia, Iowa, Idaho, Kentucky, Massachusetts, Michigan, Montana, North Carolina, North Dakota, Nebraska, New York, South Dakota, Tennessee, Utah, Virginia, and Washington. There was not an overall significant effect of a county bordering a state boundary ($\beta_3 = 0.04$, 95% CI [−0.43, 1.23]) and a single state showed a negative interaction of bordering a state boundary and ICVI farm count proxy, North Carolina ($\beta_{(2,3),NC} = -1.65$, 95% CI [−3.29, −0.15]). Overall, there was no effect of state boundary on the slope of the relationship between NASS census farm count and ICVI farm proxy count ($\beta_{(1,3)} = -1.41$, 95% CI [−1.74, 0.001]). There was a weak but significantly positive three-way interaction between the effect of a county bordering a state boundary, the slope of the relationship between ICVI farm proxy and NASS farm census in 3 states: Illinois ($\beta_{(1,2,3),IL} = 3.6 \times 10^{-3}$, 95% CI = [$4.9 \times 10^{-4}$, 6.7×10^{-3}]), North Carolina ($\beta_{(1,2,3),NC} = 3.7 \times 10^{-3}$, 95% CI = [$2.7 \times 10^{-4}$, 7.4×10^{-3}]), and Oregon ($\beta_{(1,2,3),OR} = 2.5 \times 10^{-3}$, 95% CI [$4.9 \times 10^{-4}$, 4.4×10^{-3}]).

3.5. Network correlations with NASS census of cattle infrastructure

There were few strong correlations between node (county) network metrics and county-level summaries of cattle infrastructure and inventory for the 2007 NASS census (Table 4). The strongest correlation was between in-degree and the number of cattle on feed (Table 4), and this correlation appeared to be restricted to incoming cattle

Table 3
Correlations between node metrics at the county and state scales.

	County				State			
	In-degree	Out-degree	Between	Close	In-degree	Out-degree	Between	Close
In-degree	1	0.342	0.712	0.486	1	0.666	0.440	0.585
Out-degree		1	0.612	0.547		1	0.480	0.732
Betweenness			1	0.430			1	0.667
Closeness				1				1

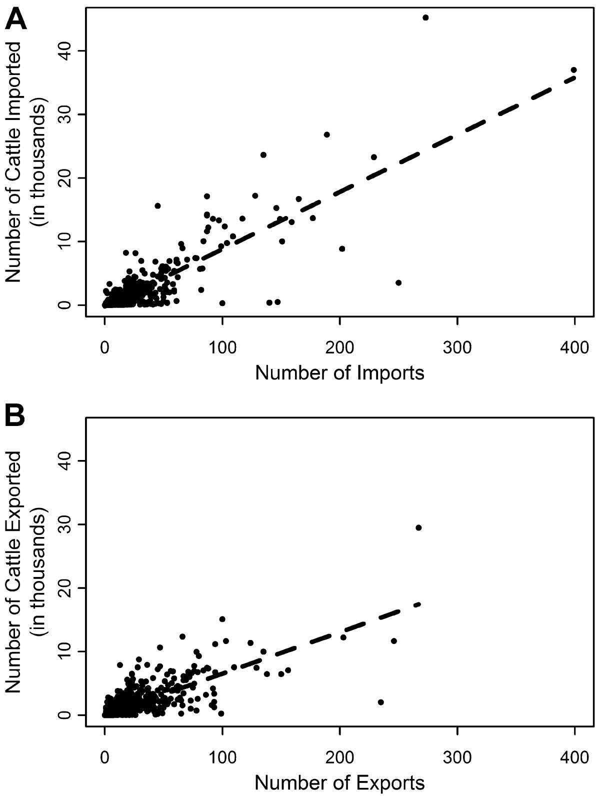

Fig. 5. County scale relationship between (A) the number of cattle imported (cattle-weighted in-degree) and the number of imports (shipment-weighted in-degree) and (B) the number of cattle exported (cattle-weighted out-degree) and the number of exports (shipment-weighted out-degree).

shipments, as there was weak correlation between out-degree and number of cattle on feed (Table 4 and Fig. 7). Although correlations were generally low, count metrics predicted network structure better than density metrics, and cattle inventory was always a better predictor of network structure than farm number (Table 4). Additionally, in most cases, cattle infrastructure was a better predictor of in-degree, but cow–calf inventory and operations appeared to be more associated with a node's out-degree (Table 4).

4. Discussion

ICVIs, to this point, have primarily been used for to ensure state animal health requirements are met. Here, we co-opted the utility of ICVIs for tracing individual shipments to obtain a sample of the interstate cattle movement network in the U.S. This work represents one of the first truly comprehensive studies of national-level cattle shipment contents and spatial patterns in the U.S. by describing shipment patterns at a county resolution; enhancing previous efforts at summarizing interstate U.S. cattle movement (i.e., Shields and Mathews, 2003). The interstate nature of the data restricts the direct inferences we can make about cattle transports within state boundaries; however, such a data-driven and nation-wide characterization of interstate cattle movement has been notably lacking for disease risk prevention, surveillance, and control (NRC, 2012).

Shipment characteristics derived from the ICVIs qualitatively matched with the conventional wisdom about the U.S. cattle industry. The frequency of relatively small shipments we observed matched the large proportion of cattle holdings in the U.S. with fewer than 100 head of cattle (90.4% of beef farms, USDA, 2011; 76.7% of dairy farms,

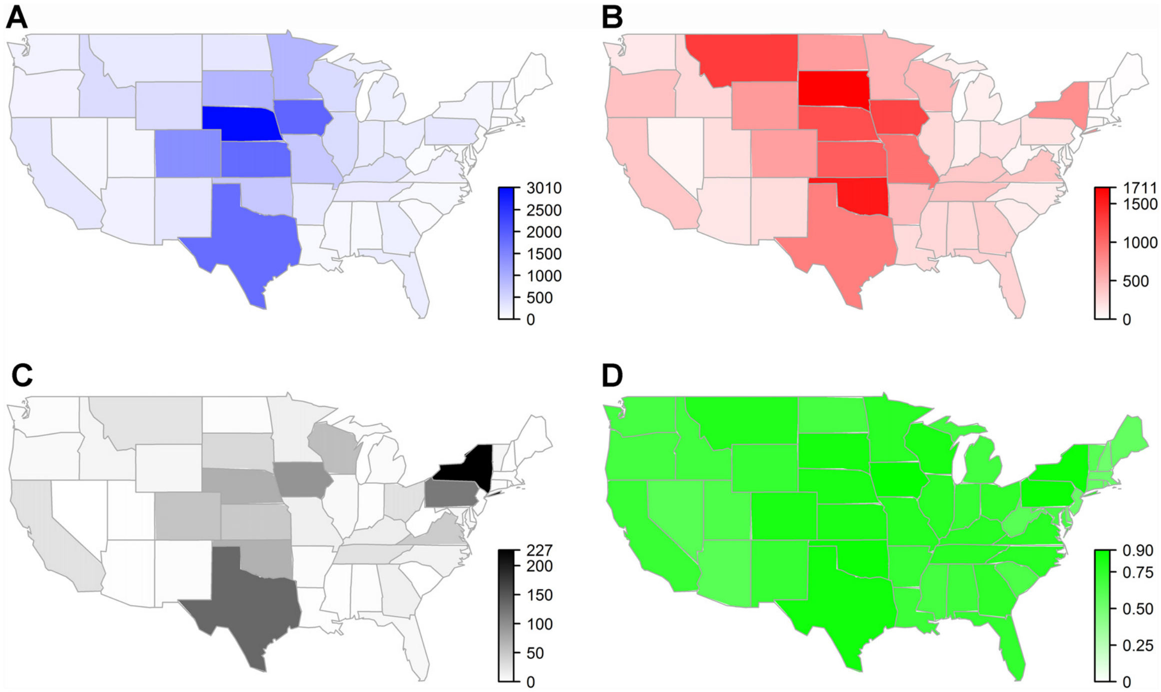

Fig. 6. Maps of (A) in-degree, (B) out-degree, (C) betweenness, and (D) closeness for the state scale network.

Table 4
Pearson's correlation coefficients between county-scale network metrics and county-level cattle industry characteristics from the 2007 NASS farm census. Numbers in brackets give the 95% confidence interval based on 1000 bootstrap replicates.

NASS data category	Node (county) network statistic			
	In-degree[a]	Out-degree[a]	Betweenness	Closeness
Farm count				
Total	0.05 [0.01, 0.10]	0.15 [0.11, 0.19]	0.15 [0.11, 0.22]	0.22 [0.19, 0.26]
Operations with animals on feed	0.27 [0.19, 0.37]	0.13 [0.05, 0.21]	0.25 [0.13, 0.34]	0.22 [0.19, 0.25]
Operations with cows that calved	0.02 [−0.02, 0.06]	0.13 [0.10, 0.18]	0.12 [0.07, 0.19]	0.21 [0.17, 0.24]
Cattle inventory				
Total	0.52 [0.43, 0.63]	0.32 [0.27, 0.39]	0.42 [0.34, 0.55]	0.48 [0.42, 0.54]
Animals on feed	0.64 [0.52, 0.74]	0.15 [0.08, 0.25]	0.41 [0.30, 0.53]	0.30 [0.26, 0.34]
Cows that calved	0.21 [0.15, 0.28]	0.29 [0.25, 0.36]	0.24 [0.17, 0.37]	0.38 [0.32, 0.46]
Farm density per km^2				
Total	−0.05 [−0.08, −0.02]	0.01 [−0.02, 0.04]	0.03 [0.00, 0.06]	0.03 [0.00, 0.07]
Operations with animals on feed	0.20 [0.12, 0.28]	0.05 [−0.02, 0.13]	0.15 [0.05, 0.25]	0.14 [0.10, 0.18]
Operations with cows that calved	−0.08 [−0.10, −0.05]	0.00 [−0.03, 0.03]	0.01 [−0.02, 0.05]	0.02 [−0.02, 0.06]
Animals density per km^2				
Total	0.48 [0.39, 0.55]	0.20 [0.14, 0.28]	0.33 [0.26, 0.39]	0.41 [0.38, 0.44]
Animals on feed	0.59 [0.49, 0.68]	0.11 [0.05, 0.22]	0.33 [0.21, 0.44]	0.29 [0.26, 0.33]
Cows that calved	0.10 [0.07, 0.13]	0.15 [0.12, 0.20]	0.14 [0.11, 0.20]	0.26 [0.23, 0.29]

[a] Weighted by number of shipments.

USDA, 2008). Not only did the shipment size distribution reflect the farm size distribution, but shipment volumes by production type mirrored standing cattle populations. We found a roughly 3-to-1 ratio of beef to dairy shipments amongst the shipments that could be identified to this level.

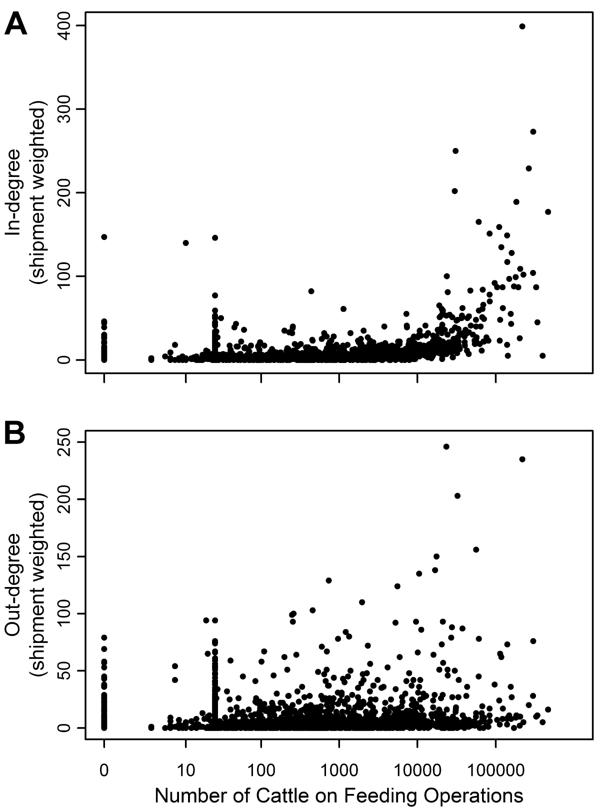

Fig. 7. Correlation between the county-level ICVI shipment weighted network in-degree (A) and out-degree (B) and the inventory of cattle on farms with feeding operations from the 2007 NASS farm census.

This ratio is close to the 25 million to 9 million beef/dairy cattle ratio reported in the 2007 NASS farm census. Additionally, many ICVIs were issued for feeding shipments and contained a large number of young animals and castrated males underscoring the central role the feedlot system plays in the U.S. cattle industry, similar to the patterns presented in the previous coarse-grained description of U.S. cattle transports presented by Shields and Mathews (2003). Thus, we conclude that we have a representative sample that reflects true patterns of interstate cattle transport in the U.S. cattle industry.

Spatial representation of the interstate cattle movement network highlighted the feedlot system. Cattle destinations (in-degrees) were greatest in states with traditionally large feeding infrastructure in the central plains states, and there was a positive correlation between network in-degree and the inventory of feeding cattle per county (Fig. 7A). Exports, however, were shifted to the northern plains and mountain states and were generally more dispersed with cattle moving to the central plains states from across the country (Fig. 5A and B). However, these general patterns at the state scale overlook heterogeneities that become apparent at the county scale. Notably, state-level importance is often due to the influence of a small number of hub counties within the state, and many counties that have high node centrality (e.g., in- and out-degree), as well as network centrality (e.g., betweenness), are in states with low centrality (Figs. 4 and 6). We conclude that quantifying interstate movements on the county scale captures heterogeneities in the interstate cattle movement network that are homogenized when the network is viewed at the state scale. This is an important distinction for disease modeling because models that consider movement at a state scale are unlikely to identify the heterogeneities, represented by the few high-centrality counties, among state-level nodes that are critical for surveillance or control actions.

The central hub structure of the U.S. cattle industry at the county-scale is apparent in the ICVI sample with nearly two-thirds of the counties present in the sample belonging to the GSCC and with almost all of the counties observed on the ICVIs linked through the GWCC (Table 2, Fig. 3). Thus, we observed a core of nodes (counties) that are capable of both sending and receiving to one another, as well as satellite nodes that send cattle into the network core (Fig. 3B). The connectance within the core, however, is not determined by simple out and back movements. The county network's low reciprocity indicates that these types of movements were rare, and it took a maximum of 12 links to reach the most distant counties (in network space). Thus, despite the central geography of the feedlot system, oftentimes relationships between nodes are more circuitous than might be expected within the unique structure of the U.S. cattle industry.

Despite the recognition of the hub structure of the county network, the determinants of node importance were unclear with several potential factors that could influence network structure. First, immense heterogeneities in county and state size in the U.S. created the potential for large counties and states to play disproportionate roles in the network because they contain more sources and destinations of shipments (Flowerdew et al., 2008). For example, in the conterminous U.S., the largest county by area (San Bernadino, California) is larger than each of the 8 smallest states by area and has more cattle than each of the 12 states with fewest cattle (NASS 2007). To address this issue, we created approximations to equally-sized county and state networks, the 50 and 500 km grid networks. We found no evidence to support the hypothesis that the heterogeneity in geographic size of the nodes influenced node importance. Our networks with nodes defined by regular grids of equal area (no heterogeneity) exhibited similar patterns and structure to the actual county and state networks. However, we did find that large differences in geographic size of the nodes across spatial scales, represented by the county vs. state scales, influenced the network structure (Table 2). Based on the available data for interstate cattle movements, we found no evidence for the influence of node size heterogeneity on network structure, and we view the county and state scales as appropriate boundaries for national management, recognizing the effect of absolute node size (i.e., the state scale homogenizes node heterogeneity vs. the county scale) on network structure.

In addition to size heterogeneities, the number of farms within a county may impact node importance in the network. We tested the naïve assumption that the number of origin and destinations of interstate movements was positively associated with the number of farms in a county (per NASS census) and found no overall association across the U.S. However, we did find a positive association between the number of farms in a county per the NASS census and the proxy for the number of farms as origins or destinations of interstate shipments within 19 states, although the magnitude of these associations was consistently smaller than the association expected under a 10% sample (i.e., $\beta_{(1,2),j} = 0.1$; Table E1). Thus, farm census data is not a good predictor of shipment behavior, and consequently, more details about the cattle industry are necessary to explain how individual counties are involved in interstate cattle movements. Attempts to explain the variation using production types (i.e., beef-feeding, beef-breeding, dairy-feeding, and dairy-breeding) were inconclusive due to overlapping network structures for some production types and the lack of shipments in each network relative to the number of nodes (Appendix F). We also calculated correlations between county-level network metrics and county-level animal numbers broken down by those animals that are feeding and those that are on breeding operations (i.e., cow-calf; Table 4). We found that the number of premises with feeding animals and the number of animals on feed were the only meaningful predictors of node in-degree suggesting an aggregating effect of feedlots. These same measures were less strongly associated with node betweenness and out-degree. We speculate that more specific information on the types of infrastructure found at both the origin and destination counties, such as markets or abattoirs, will explain much of the variation we did not capture (Bigras-Poulin et al., 2006; Natale et al., 2009); although this level of detailed premises information is not systematically available on ICVIs (Portacci et al., 2013).

Even with the unique feeding structure in the U.S., the interstate U.S. county network is similar in many ways to farm-level livestock networks in other countries. In particular, the relatively large GSCC and low transitivity observed in interstate U.S. cattle movements at the county scale is similar to that observed in sheep movement networks United Kingdom (Kao et al., 2006; Kiss et al., 2006). Given the importance of animal transports in the 2001 Foot-and-Mouth disease epidemic in the U.K. (Kao et al., 2006), large-scale epidemics in the U.S. cattle industry are a justifiable concern. However, identifying counties important for disease spread is not as straightforward as in the U.K. sheep movement network, where a high correlation between a node's in- and out-degree allows the use of first order network structures to assign potential importance in disease spread. Instead, the interstate U.S. cattle movement network exhibits a correlation between in- and out-degree that is more similar to cattle movement networks in Italy, where control policies based on a node's betweenness were found to be reasonably effective at reducing simulated disease spread (Natale et al., 2009). Similarly, betweenness has also been found to be important in predicting disease outcomes based on U.K. cattle movement networks (Ortiz-Pelaez et al., 2006) and identifying market hubs where in- and out-degree are more strongly correlated (Robinson and Christley, 2007). Betweenness is useful as a higher order metric in that it quantifies a node's role in the flow of cattle through the network as a whole, thus capturing more global structure than degree centrality. This difference was observed at the state scale where betweenness identified hubs that were not seen when considering imports or exports alone (Fig. 5). ICVIs allow complex networks of movements to be built and patterns of animal transports to be discovered beyond primary state-to-state trading partners. Thus, the benefits of understanding the higher order properties of the network, especially betweenness, point to the untapped potential of ICVI data to understand and inform policy decisions regarding interstate U.S. cattle movement.

While we advocate the utility of the ICVI data, it does have limitations in its scope of inference for use at the national scale to help manage diseases. Although we did sample some intrastate movement, ICVIs are designed to capture interstate movement. Thus, our network's strength lies in capturing the long-distance tail of U.S. cattle shipments. These global processes are often missing from other network studies, and although we lack the data to describe local movement processes for a more complete understanding of U.S. cattle movement, methods have been developed to bridge this gap (Lindström et al., 2013). We also acknowledge that our network is limited to a static picture of one year of movement from a sample of ICVI records. This limitation underscores the need for a wider availability of ICVIs (or other comparable systematic electronic cattle transport data) to enable study of the changes in the structure of the cattle network across years and the effect of sampling on network structure. We also hypothesized that proximity to a state boundary could bias our characterization of cattle shipments because the proximity to a state boundary may make moving cattle interstate (thus requiring an ICVI) more likely. However, we found little evidence to support this hypothesis and consider our characterization based on the ICVI sample to be representative of the cattle operations that move cattle interstate, regardless of the uncertainty as to what characteristics underlay the shipment patterns (Table 4).

5. Conclusions

Understanding cattle movements at the national scale can have far reaching implications for disease response planners and is one of the most important applications of the network analysis of this novel data set. Identification of nodes that are highly connected can be used to develop targeted surveillance programs, and these nodes may also indicate locations where movement controls can be implemented in the event of a disease outbreak. Viewing the U.S. interstate cattle movement network at the state scale has the potential to homogenize the network but some disease outbreaks may require coarse grained surveillance and control making the state scale a necessary component of long-term disease management strategies. Disease simulation models can be used in conjunction with this network model to evaluate emergency mitigation options for highly infectious diseases in cattle. In addition, a network modeling approach can also be used to extend inference to unobserved cattle movements and quantify uncertainty for applications of the ICVI data sample (Lindström et al., 2013), as well as explain and understand disease spread for endemic diseases by evaluating the risk factors associated with movements from and to geographic regions and cohorts of cattle.

Conflicts of interest

None declared.

Acknowledgements

Funding provided by the Research and Policy for Infectious Disease Dynamics (RAPIDD) Program, Science and Technology Directorate, U.S. Department of Homeland Security, and Fogarty International Center, National Institutes of Health; Foreign Animal Disease Modeling Program, Science and Technology Directorate, U.S. Department of Homeland Security (Grant ST-108-000017); and USDA Cooperative Agreement 11-9208-0269-CA 11-1 and 09-9208-0235-CA. Data was 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 acknowledge the National Institute for Mathematical and Biological Synthesis for supporting and hosting the Modeling Bovine Tuberculosis working group, where the initial ideas for collecting and using ICVI data were developed. We also thank to the state veterinarians and staff, whose cooperation and effort made the data collection possible.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.prevetmed.2013.08.002.

References

Bailey, D., Brorsen, B.W., Thomsen, M.R., 1995. Identifying buyer market areas and the impact of buyer concentration in feeder cattle markets using mapping and spatial statistics. Am. J. Agric. Econ. 77, 309–318.

Bajardi, P., Barrat, A., Natale, F., Savini, L., Colizza, V., 2011. Dynamical patterns of cattle trade movements. PLoS ONE 6, e19869.

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.

Bigras-Poulin, M., Thompson, R.A., Chriel, M., Mortensen, S., Greiner, M., 2006. Network analysis of Danish cattle industry trade patterns as an evaluation of risk potential for disease spread. Prev. Vet. Med. 76, 11–39.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth, Belmont, CA.

Brennan, M.L., Kemp, R., Christley, R.M., 2008. Direct and indirect contacts between cattle farms in north-west England. Prev. Vet. Med. 84, 242–260.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJ. Complex Syst., 1695, http://igraph.sf.net

Dube, C., Ribble, C., Kelton, D., McNab, B., 2009. A review of network analysis terminology and its application to foot-and-mouth disease modeling and policy development. Transbound. Emerg. Dis. 56, 73–85.

Flowerdew, R., Manley, D.J., Sabel, C.E., 2008. Neighbourhood effects on health: does it matter where you draw the boundaries? Soc. Sci. Med. 66, 1241–1255.

Forde, K., Hillberg-Seitzinger, A., Dargatz, D., Wineland, N., 1998. The availability of state-level data on interstate cattle movements in the United States. Prev. Vet. Med. 37, 209–217.

Green, D.M., Kiss, I.Z., Kao, R.R., 2006. Modelling the initial spread of foot-and-mouth disease through animal movements. Proc. R. Soc. B 273, 2729–2735.

Hoffman, L.A., Boles, P.P., Hutchinson, T.Q., 1975. Livestock trucking services: quality, adequacy, and shipment patterns. Economic Research Service, USDA. Agricultural Economic Report No. 312.

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. Proc. R. Soc. B. 273, 1999–2007.

Kao, R.R., Green, D.M., Johnson, J., Kiss, I.Z., 2007. Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the UK. J. R. Soc. Lond. Interface 4, 907–916.

Kiss, I.Z., Green, D.M., Kao, R.R., 2006. The network of sheep movements within Great Britain: network properties and their implications for infectious disease spread. J. R. Soc. Lond. Interface 3, 669–677.
Lindström, T., Grear, D.A., Buhnerkemp, 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, e5432, 10.371/journal.pone.0053432.
Liu, H., Schumm, P., Lyubinin, A., Scoglio, C., 2012. Epirur_Cattle: a spatially explicit agent-based simulator of beef cattle movements. Procedia Comp. Sci. 9, 857–865.
Natale, F., Giovannini, A., Savini, L., Palma, D., Possenti, L., Fiore, G., Calistri, P., 2009. Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread. Prev. Vet. Med. 92, 341–350.
NRC (National Research Council), 2012. Evaluation of the updated site-specific risk assessment for the Department of Homeland Security's planned national bio- and agro-defense facility in Manhattan, Kansas. National Academies Press, Washington, DC.
Ortiz-Pelaez, A., Pfeiffer, D.U., Soares-Magalhães, R.J., Guitian, F.J., 2006. Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. Prev. Vet. Med. 76, 40–55.
Portacci, K., Miller, R.S., Riggs, P., Buhnerkempe, M., Abrahamsen, L., 2013. Assessment of paper Interstate Certificates of Veterinary Inspection to support animal disease tracing. J. Am. Vet. Med. Assoc. 243, 555–560.
R Development Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/
Robinson, S.E., Christley, R.M., 2007. Exploring the role of auction markets in cattle movements within Great Britain. Prev. Vet. Med. 81, 21–37.
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 pp.
USDA, 2007. Census of Agriculture. USDA National Agricultural Statistics Service. http://www.nass.usda.gov/census
USDA, 2008. Dairy 2007. Part II: Changes in the U.S. Dairy Cattle Industry, 1991–2007. USDA-APHIS-VS, CEAH. Fort Collins, CO, #N481.0308.
USDA, 2011. Small-scale U.S. Cow–Calf Operations. USDA-APHIS-VS, CEAH, Fort Collins, CO, #596.0411.
Velthuis, A.G.J., Mourits, M.C.M., 2007. Effectiveness of movement-prevention regulations to reduce the spread of foot-and-mouth disease in The Netherlands. Prev. Vet. Med. 82, 262–281.
Vernon, M.C., Keeling, M.J., 2009. Representing the UK's cattle herd as static and dynamic networks. Proc. R. Soc. B 276, 469–476.
Volkova, V.V., Howey, R., Savill, N.J., Woolhouse, M.E.J., 2010. Sheep movement networks and the transmission of infectious diseases. PLoS ONE 5, e11185.
Webb, C.R., 2005. Farm animal networks: unraveling the contact structure of the British sheep population. Prev. Vet. Med. 68, 3–17.