Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Epidemics 24 (2018) 67–75

Contents lists available at ScienceDirect

Epidemics

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

Role of animal movement and indirect contact among farms in transmission of porcine epidemic diarrhea virus

Kimberly VanderWaal*, Andres Perez, Montse Torremorrell, Robert M. Morrison, Meggan Craft

Department of Veterinary Population Medicine, University of Minnesota, Twin Cities, 1365 Gortner Avenue, St. Paul, MN 55113, USA

ARTICLE INFO

Keywords:
Animal movement
Network analysis
Computational modeling
Swine pathogens
Livestock
Epidemiology

ABSTRACT

Epidemiological models of the spread of pathogens in livestock populations primarily focus on direct contact between farms based on animal movement data, and in some cases, local spatial spread based on proximity between premises. The roles of other types of indirect contact among farms is rarely accounted for. In addition, data on animal movements is seldom available in the United States. However, the spread of porcine epidemic diarrhea virus (PEDv) in U.S. swine represents one of the best documented emergences of a highly infectious pathogen in the U.S. livestock industry, providing an opportunity to parameterize models of pathogen spread via direct and indirect transmission mechanisms in swine. Using observed data on pig movements during the initial phase of the PEDv epidemic, we developed a network-based and spatially explicit epidemiological model that simulates the spread of PEDv via both indirect and direct movement-related contact in order to answer unresolved questions concerning factors facilitating between-farm transmission. By modifying the likelihood of each transmission mechanism and fitting this model to observed epidemiological dynamics, our results suggest that between-farm transmission was primarily driven by direct mechanisms related to animal movement and indirect mechanisms related to local spatial spread based on geographic proximity. However, other forms of indirect transmission among farms, including contact via contaminated vehicles and feed, were responsible for high consequence transmission events resulting in the introduction of the virus into new geographic areas. This research is among the first reports of farm-level animal movements in the U.S. swine industry and, to our knowledge, represents the first epidemiological model of commercial U.S. swine using actual data on farm-level animal movement.

1. Introduction

Mathematical and computational modeling of infectious diseases is a common approach to simulating the spread of disease in a population, exploring key epidemiological parameters that drive transmission, and evaluating alternative control strategies (Brooks-Pollock et al., 2015; Craft, 2015; VanderWaal et al., 2017). In livestock populations, network-based models based on data on animal movements between farms have been a key area of research (Bajardi et al., 2012; Craft, 2015; Green et al., 2006; Kao, 2002; Kao et al., 2007; Rossi et al., 2015). However, animal movement data is rarely available for livestock industries in the United States due to the lack of a comprehensive national livestock traceability program. This limits capabilities to predict the dynamics of infectious diseases at the landscape, regional, and national levels and hinders development of risk-based surveillance and control measures based on movement data. Animal movement data may be particularly important for the swine industry, where production is highly vertically integrated in that pigs are moved between multiple premises between birth and slaughter, with each premise potentially located in different states (Valdes-Donoso et al., 2017). Such frequent and long distance movement makes the U.S. swine industry vulnerable to infectious disease epidemics.

In addition to direct contact among farms via animal movements, indirect contact may occur between farms due to windborne propagation of aerosols and dissemination of fomites by personnel, contaminated vehicles, and feed (Alonso et al., 2014; Alvarez et al., 2016; Beam et al., 2015; Dee et al., 2014; Kim et al., 2017; Lowe et al., 2014; O'Dea et al., 2015; Pasick et al., 2014). Although the potential importance of such mechanisms in creating transmission opportunities between swine premises has been shown in experimental studies and outbreak investigations (Alonso et al., 2014; Bowman et al., 2015; Lowe et al., 2014; Pasick et al., 2014), indirect contact is less often accounted for in epidemiological models (Arruda et al., 2016; Martinez-Lopez et al., 2011; Thakur et al., 2015; Yadav et al., 2016). Models of pathogen spread in livestock populations focus primarily on animal movement and, in some cases, local spatial spread based on proximity

* Corresponding author.
E-mail addresses: kvw@umn.edu (K. VanderWaal), aperez@umn.edu (A. Perez), torr0033@umn.edu (M. Torremorrell), craft@umn.edu (M. Craft).

https://doi.org/10.1016/j.epidem.2018.04.001
Received 28 November 2017; Received in revised form 3 March 2018; Accepted 9 April 2018
Available online 12 April 2018

between premises (Brooks-Pollock et al., 2015). Attempts to account for indirect contact are hindered due to lack of data on which farms are connected via indirect contact. However, the structured nature of U.S. swine companies provides an ideal opportunity to infer patterns of indirect contact among farms and explore the joint impact of direct and indirect transmission on the spread of pathogens.

In May 2013, a new pathogen emerged and rapidly spread in the United States swine industry, resulting in major production impacts due to a mortality rate in neo-natal piglets of up to 100% (Saif et al., 2012). The disease was caused by porcine epidemic diarrhea virus (PEDv), an RNA coronavirus in the family *Alphacoronaviridae* that previously was circulating in Asia (Huang et al., 2013; Stevenson et al., 2013). As the name suggests, clinical signs of PEDv include watery diarrhea and vomiting. By the end of June 2014, PEDv had spread to 30 states and impacted approximately 50% of breeding herds (Goede and Morrison, 2016). PEDv may have resulted in the deaths of at least 7 million piglets, and it may take up to 12 weeks for infected sow farms, where breeding and farrowing occur, to recover their pre-infection piglet production levels (Goede and Morrison, 2016). Furthermore, surviving piglets exhibit poor growth during the growing period (Alvarez et al., 2015).

Despite the rapid between-farm spread of PEDv within the U.S. swine industry, there are a number of competing hypotheses concerning the main mechanisms of between-farm transmission, with no clear resolution on the relative importance of each of these mechanisms. For example, while contaminated feed may have contributed to the rapid emergence of PEDv (Dee et al., 2014; Pasick et al., 2014), there is also substantial evidence that movement of infectious pigs contributes to spread within a single flow of animals (i.e., movement of pigs from farrowing/sow farms to nurseries, and subsequently to finishing farms where fattening occurs) (Bowman et al., 2015). On a larger spatial scale, states between which there were high rates of pig movement exhibited more synchronous PEDv epidemics in terms of weekly case incidence, suggesting that pig movements created epidemiological linkages among states (O'Dea et al., 2015).

Unrelated to animal movements, there is evidence that local spatial spread between farms also occurs as infected farms are more clustered geographically than expected by chance (Alvarez et al., 2016). PEDv virus can also be recovered from air samples collected up to 16 km from infected farms (Alonso et al., 2014). Taken together, these studies suggest that windborne aerosols could contribute to between-farm transmission, especially at distances less than 4.8 km (Alonso et al., 2014; Alvarez et al., 2016). In addition, fomites may also contribute to between-farm spread. For example, market trucks that move pigs to slaughter facilities may function as mechanical vectors for fomites if they become contaminated at the slaughter facility and subsequently transmit the disease to naïve farms (Lowe et al., 2014).

The spread of PEDv in U.S. swine represents one of the best documented emergences of a highly infectious pathogen in the United States livestock industry, providing a unique opportunity to parameterize models of pathogen spread in U.S. swine. The objective of this research is to apply epidemiological modeling approaches to simulate the spread of PEDv at the regional scale in order to answer unresolved questions concerning factors facilitating between-farm transmission. Using real-world data on observed pig movements during the initial phase of the PEDv epidemic, we develop a network-based and spatially explicit epidemiological model that simulates the spread of PEDv via both indirect and direct movement-related contact. By fitting this model to the observed epidemiological dynamics in a geographically isolated production company of nearly 400 farms, we (i) evaluate the relative contribution of each of six direct and indirect mechanisms of between-farm transmission, and (ii) determine the most likely transmission mechanisms responsible for long-distance jumps, and (iii) discuss how these methods can be used to help producers mitigate future outbreaks. This research is among the first reports of farm-level animal movements in the U.S. swine industry (Lee et al., 2017; Valdes-Donoso et al., 2017) and, to our knowledge, represents the first epidemiological model of commercial U.S. swine using actual data on farm-level animal movement.

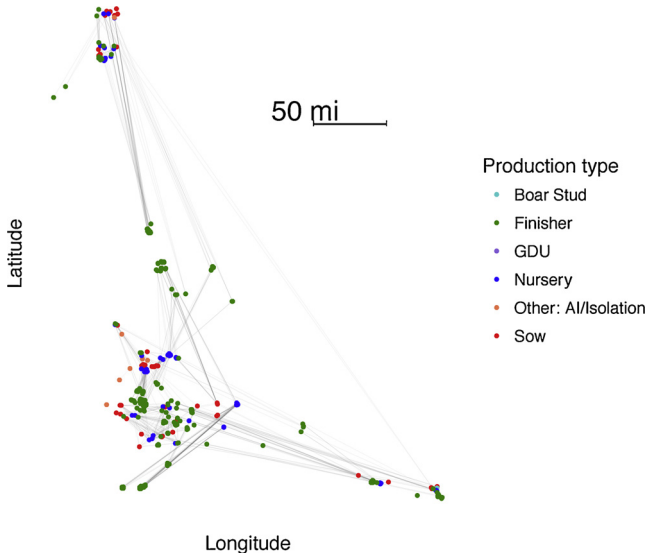

Fig. 1. Map of farm locations (colored nodes) and between-farm pig movements (gray lines) occurring between May through September 2013.

2. Materials and methods

2.1. Data source

Data on farm attributes, management, and between-farm pig movements were available for a single production company, or "system," located in the Great Plains states of the west-central United States (Fig. 1). Swine production systems in the U.S. are "vertically integrated" in that different stages of production (from birth of piglets through slaughter) occur at different premises specializing in that particular stage. Primary production types included sow farms (housing sows during gestation and farrowing [birthing] periods and preweaning piglets), nursery farms (housing weaned piglets for approximately six to eight weeks), and finishers (where pigs are moved after the nursery period to fatten them for slaughter). Replacement gilts (young females) are usually brought into sow farms from gilt development units (GDUs), which are located either on the sow farm or at a separate farm. High biosecurity production types involved in maintenance of genetic stock included boar studs (premises housing boars used as studs for artificial insemination) and AI/Isolation units. A summary of farms in the study by production type is provided in Supplementary Table S1. Farm attribute data included the geographic location as UTM (Universal Transverse Mercator) coordinates, herd size, and production type for each of 376 company-managed farms.

Farm management data were available for each farm regarding the feed mill from which feed was sourced and the organizational "flow" to which the farm belonged. Flows were defined as groups of farms that were managed as a unit and shared support services, personnel, and truck washes. This production system was geographically isolated from other swine farms; however, the locations of 84 neighboring swine farms from seven neighboring systems were also available, as these farms may play a role in local disease transmission.

The first PEDv case in this system occurred on May 9, 2013 in a sow farm. Therefore, pig movement data spanned a timeframe from May 4, 2013 to October 1, 2013, as it was assumed that the farm may have been infectious prior to detection of clinical signs. Movement data included the date of each movement, total number of animals of moved, and the premise ID of the source and destination farms. During this five-

month period, 10,709 between-farm shipments occurred with a mean of 339.7 pigs per load (range: 1–1378 pigs) and totaling 3.6 million pigs. Sow farm to nursery movements accounted for 47% of movements, followed by nursery to finisher movements (34%) and finisher to sow farms (10.5%, see Supplementary Table S2 for a full assortativity table by production type). Distance of movements ranged from < 1 to 644 km, with median distances of 15, 40, 22 km for movements originating from sow farms, nurseries, and finishers, respectively (see Supplementary Table S3 for full summary). 83% of movements were within flow, and 17% of movements were between flows. Movements from nurseries were the most common between-flow movement, accounting for 38% of all nursery movements. Movements to slaughter were not included. Movements did not occur with farms not part of the production system, and no movement data were available for neighboring farms.

Data on PEDv status in all company-managed sow farms was available on a weekly basis via the Morrison Swine Health Monitoring Project (MSHMP). PEDv status was not documented for non-sow farms. Uninfected farms were classified as PEDv negative. A sow farm was considered "positive-unstable" when clinical signs consistent with PEDv were observed. Status was confirmed by PCR. Weaned piglets moved from the farm during this period would likely be infected. Infected farms were considered to have reached stability (classed as "positive-stable") when clinical signs were absent for > 21 days, and at least 30 litters were PCR negative.

We defined two geographic areas (Fig. 1) that were geographically removed from the area in which the index case occurred: Region I was 450 km from the index case area and the earliest infection in this area in the observed epidemic was day 22. Region II was 260 km from Region I and even farther removed from the index area. Region II was not infected during the observed epidemic. There was virtually no swine production, either by this production system or by others, in the intermediate areas between these regions.

2.2. Model description

2.2.1. Overview

We developed a stochastic between-farm transmission model that mechanistically simulated six types of transmission between farms, including one direct mechanism (animal movement) and five indirect mechanism. Mechanisms included: movement of infected animals (1), local spatial spread (2), contaminated feed (3), and fomites that could spread among farms receiving feed from the same mill via feed trucks (4), among finishing farms via movement of market trucks to and from slaughter facilities (5), and/or among farms in the same flow via shared support services (6). Shared support services include things like personnel, vehicles, truck washes, etc., but these are not independently account for. Because 80% of sows may be clinical within a two to three days post-introduction of PEDv onto a farm (Bowman et al., 2015), we considered the farm to be the epidemiological unit. The model followed a compartmental framework, where farms were classified as Susceptible, Latent (exposed but not yet infectious), or Infectious (positive-unstable; presence of clinical animals). The model operated on a daily time step for five months, beginning with the first observed case of PEDv in this system.

When a farm was infected by any of the transmission mechanisms (see sub-sections below for details), the transmission mechanism was recorded and the farm is classified as Latent. Movements from latent farms were considered potentially infectious, but it was assumed that farms in the latent period did not contribute to other modes of transmission (fomites or local spatial spread) due to low shedding. Indeed, the latent period from exposure to shedding for an individual animal ranges from 24 to 72 h (Crawford et al., 2015; Madson et al., 2014; Song et al., 2006). The latent period at the farm level prior to detection of clinical disease has been estimated through field investigations to be between 2 and five days, after which up to 80% of sows may be clinical affected (Bowman et al., 2015; CAHFS, 2013; Saif et al., 2012). Thus, the number of days each farm remained in the latent period was drawn from a PERT distribution with a minimum of 1, mode of 2, and maximum of 5 days, after which a farm becomes infectious (Bowman et al., 2015; CAHFS, 2013; Saif et al., 2012).

Based on analysis of field data on the length of time required for a farm to no longer produce PEDv positive piglets, the infectious period length was drawn from a PERT distribution with a minimum of 7 weeks, mode of 23 weeks, and maximum of 64 weeks (Goede and Morrison, 2016). After the infectious period, sow farms are classified as positive-stable. Infection in sow farms was generally managed through intentional exposure of gilts to increase immunity levels, thus although virus was present, large outbreaks with high levels of shedding did not occur and weaned piglets moved from these farms were negative. Positive-stable sow farms thus are no longer infectious to others, and due to ongoing disease management within the farm, they are not considered susceptible for the purposes of the model. Unlike sow farms, nurseries and finishers generally utilized all-in-all-out management, meaning that these farms were completely emptied of live animals before new animals were introduced. Thus, after the infectious period, non-sow farms were moved to the susceptible class.

2.2.1.1. Direct transmission via animal movements. For each daily time step t, all movements involving latent and infected farms at time t were extracted from the movement database. The probability that no infected animals are moved from infected farm i to susceptible farm j is $(1-\mu)^{b_{ij}}$, where μ is the probability that an animal is infected (based on the within-farm prevalence) and b_{ij} is the batch size (the total number of animals moved between farm i and j). Thus, the probability that farm j becomes infected in time t given that it receives movements from i farms is given by:

$$P(farm_j \ becomes \ infected) = 1 - \prod_{i}^{1} (1-\mu)^{b_{ij}}$$

Movement of weaned pigs from infected sow farms may be reduced as a result of high neonatal mortality rates. An assessment of the observed movement data from PED-positive farms revealed that although the number of pigs per shipment did not change as a result of PED infection in a sow farm, the number of shipments leaving an infected sow farm was reduced by 70% during the first five weeks post-infection, and by 15% during weeks six through ten post-infection. After ten weeks, movements returned to their pre-infection frequency. Thus, in the model, 70% and 15% of shipments from sow farms were randomly excluded for 0–5 weeks and 6–10 weeks post-infection, respectively. This was done because sow farms that were uninfected in the observed data could be infected in the model, and not accounting for their change in movement frequency could result in overestimation of spread via animal movements.

2.2.1.2. Indirect transmission via local spread. Local spread accounts for localized processes that could contribute to transmission, but unlike movements, are not directly measured. Inherent to local spread is the assumption that transmission between farms is more likely between farms that are geographically closer. Mechanisms include windborne dissemination of virus aerosols, contaminated sewage or water, and movement of vehicles/personnel. Local spread is captured via a spatial transmission kernel (Supplementary Methods S1)(Keeling et al., 2004; Szmaragd et al., 2009), where the probability of transmission decreases as two farms become farther apart:

$$P(farm_j \ becomes \ infected) = 1 - \prod_{i}^{1} \left(1 - \phi \frac{\alpha}{\pi} d_{ij}^{-\sqrt{\alpha}}\right)$$

Where farm j is an uninfected farm and d_{ij} is the distance between farm j and every infected farm i within 50 km. Distances of > 50 km were not considered as local spread over long distances is unlikely. $\phi \frac{\alpha}{\pi} d_{ij}^{-\sqrt{\alpha}}$

represents the probability of transmission between farm i and j given distance apart in kilometers, and thus 1- $\phi\frac{\alpha}{\pi}d_{ij}^{-\sqrt{\alpha}}$ is the probability that transmission will not occur between those two farms. ϕ and α control the shape of the transmission kernel, with ϕ representing the basal transmission coefficient (at $d_{ij} = 0$), α controlling the steepness with which probabilities decline with distance. The power-law shape of the transmission kernel and rate at which transmission declined with increasing distance ($\alpha = 0.312$) were based on an analysis of the recovery of PEDv in air samples at various distances from infected farms (Supplementary Methods S1, Fig. S1-1) (Alonso et al., 2014). It should be noted that Alonso et al. (2014) was conducted in the same company and during the same time period as the modeled study period, making their data highly relevant for inferring airborne dissemination of virus in this system. Although PEDv recovered from air samples collected from the field were not infectious in bioassays (Alonso et al., 2014), the sample size and sampling duration (30 min) may have been insufficient to capture a low probability event (i.e., recovery of infectious virus in air). Samples were also collected during the day, when UV radiation and temperature would be expected to reduce survival. However, infectiousness of airborne PEDv has been shown experimentally (Alonso et al., 2014) and the possibility of this mechanism has been supported in epidemiological investigations (Alvarez et al., 2016; Beam et al., 2015). Thus, we believe that it is a reasonable assumption that the likelihood of infectious virus to be found in air would be proportional to the total concentration of virus (inactivated or otherwise) found in the air samples.

Alvarez et al. (Alvarez et al., 2016) demonstrated that the risk of spread at short distances (1–2 km) was three to twelve times higher during the first week post-infection, most likely due to high morbidity and shedding within the farm. Therefore, we included a parameter, ϕ_7, which acted as a multiplier on ϕ during the first seven days that a farm was infected to account for potential rapid spread due to this initially high shedding.

2.2.1.3. Indirect transmission via contaminated feed. Feed contamination has been suggested to play a role in the dissemination of PEDv in North America (Dee et al., 2014; Pasick et al., 2014). Here, we assumed that infection due to contaminated feed would function as a point-source outbreak; all farms that received feed from the same mill would be likely to become infected near simultaneously. Thus, we identified weeks during which a point-source outbreak appeared to be possible in the group of sow farms receiving feed from each of the four mills (Supplementary Methods S2, Fig. S2-1). We identified two time periods in which sow farms associated with two separate mills experienced temporal clustering of cases consistent with a potential point-source outbreak in feed (Supplementary Methods S2-1). One occurred in Mill A between May 14 and May 27, during which half of the 12 susceptible sow farms that received feed from Mill A became infected, and the other in Mill B between June 11 and June 24, during which over 80% of 17 susceptible sow farms receiving feed from Mill B became infected.

We modeled feed contamination as a two-step process. First, we included a 0/1 parameter (F_A and F_B) in the model controlling whether a contamination event occurred during the temporal clustering periods in Mill A and B. Second, if an event occurred, the virus would likely not be homogenously distributed in feed, hence all farms receiving feed from those mills became infected with probability $p_{con.feed}$. In this way, we captured the potential for feed contamination events to drive sudden increases in the number of infected farms. However, such increases could be explained by other mechanisms. We explore the ability of alternate mechanisms to explain observed patterns during the process of selecting the best-performing models described below.

2.2.1.4. Indirect transmission via feed trucks: fomite transmission among farms receiving feed from the same mill via feed trucks. Regardless of contaminated feed, all farms receiving feed from the same mill were assumed to be in indirect contact, with potential risk for transmission via fomites spread through the movement of feed trucks between farms. However, the actual network of movement of trucks between farms was not known. Thus, a farm's risk of becoming infected via this mechanism was assumed to be proportional to the number of farms infected within its feed truck network. Thus, we defined the feed truck force-of-infection as the per capita risk at which susceptible farms become infected. For each feed truck network, this was represented as a daily probability:

$$P(farm_j \ becomes \ infected) = \beta_f \frac{I_k}{N_k}$$

Where β_f is the transmission coefficient for feed truck related spread, N_k and I_k are the total number and number of infected farms in the k^{th} feed truck network. Although it's possible that trucks remain contaminated even after farms recover, we made the simplifying assumption that fomite contamination would be far higher when farms are actively infected and would be negligible after farms are no longer infectious. This also applies to other fomite-related transmission mechanisms.

2.2.1.5. Indirect transmission via market trucks: fomite transmission among finishing farms via movement of trucks to and from slaughter facilities. Finishing farms were assumed to be in indirect contact with one another due to the movement of market trucks to and from slaughter facilities (Lowe et al., 2014). We assumed that as more finishing farms became infected, the amount of contamination at slaughter facilities increases. This would subsequently lead to an increased likelihood that a market truck returning from a slaughter facility would be contaminated. Similar to feed trucks, a finishing farm's risk of becoming infected via this mechanism was based on the force-of-infection amongst finishing farms. The daily probability of a finishing farm becoming infected via market trucks was defined as:

$$P(farm_j \ becomes \ infected) = \beta_m \frac{I_{fin}}{N_{fin}}$$

Where β_m is the transmission coefficient for market truck related spread, N_{fin} and I_{fin} are the total number and number of infected finishing farms. All finishing farms in this company utilized a single slaughter facility owned by the company therefore were all indirectly connected in our model.

2.2.1.6. Indirect transmission within-flows: fomite transmission via shared support services. Farms that were managed together as a "flow" were assumed to be in indirect contact due to sharing of support services. Similar to feed and market trucks, a farm's risk of becoming infected via fomites from shared support services was based on the force-of-infection within the flow. The daily probability of a farm becoming infected via within-flow transmission was defined as:

$$P(farm_j becomes infected) = \beta_c \frac{I_c}{N_c}$$

Where β_c is the transmission coefficient for within-flow spread, and N_c and I_c are the total number and number of infected farms in each flow.

2.2.2. Model simulation and calibration

Because most of the parameter values controlling the likelihood of each type of transmission are unknown, we conducted a multivariate calibration exercise on uncertain parameters (μ, F_A, F_B, $p_{con.feed}$, ϕ, ϕ_7,β_f, β_m, β_c) using Latin Hypercube Sampling (LHS) and random forest analyses (see Table 1 for definitions of symbols). This approach has often been used for global sensitivity analyses in disease models and agent-based models (Blower and Dowlatabadi, 1994; Legrand et al., 2008; White et al., 2017; Wu et al., 2013). We generated 1000 parameter sets through sampling a Latin Hypercube, which is expected to efficiently cover the parameter space. Simulation results from an initial exploratory model calibration (e.m.c.) revealed that large portions of parameter space yielded unrealistic models with very low fitness. Thus,

Table 1
Parameter definitions and minimum and maximum values in which parameter values were sampled from in the Latin Hypercube Sampling (LHS) analysis. Some parameter ranges were based on published research. Where literature values were unavailable, parameter ranges were based an initial exploratory model calibration (e.m.c.).

Symbol	Definition [units]	Where used	Min (LHS)	Max (LHS)	Citation
b_{ij}	batch size [number pigs]	Animal movements transmission	–	–	(observed)
μ	P(a pig moved from an infected herd is infected)	Animal movements transmission	0.5	1.0	(e.m.c.)
Φ	Spatial transmission coefficient	Airborne spatial spread	0.001	0.02	(e.m.c.)
Φ_7	Multiplier on local spread during first week post-infection	Airborne spatial spread	1	12	Alvarez et al. (2016)
d_{ij}	Distance between infected farm and farm$_j$ [km]	Airborne spatial spread	–	–	(observed)
β_f	Transmission coefficient – fomites – feed-trucks	Feed trucks	0	0.01	(e.m.c.)
β_m	Transmission coefficient – fomites – market trucks	Market trucks (finishers)	0	0.01	(e.m.c.)
β_c	Transmission coefficient – fomites – within-flow	Within-flow force of infection	0	0.01	(e.m.c.)
F_A	P(Feed contamination event in Mill A)	Feed-borne transmission	0/1 on day 10		(observed/e.m.c.)
F_B	P(Feed contamination event in Mill B)	Feed-borne transmission	0/1 on day 38		(observed/e.m.c.)
$p_{feed.con}$	P(Farm$_i$ infected | Feed contamination)	Feed-borne transmission	0	0.82	(observed/e.m.c.)

calibration was repeated with a new Latin Hypercube generated from parameter value ranges refined from the initial exploratory calibration (see Table 1 for minimum and maximum values for each parameter). The infection was always seeded in the farm that was the observed index case for this system.

Model outputs included the day on which each farm became infected and the mechanism by which they became infected. From this, we compared the simulated to observed PEDv dynamics by a) calculating the Spearman's rank correlation, ρ, in the order in which sow farms became infected, and b) the percent deviation, Dev_t, in the day in which sow farms became infected. For the latter, this was calculated as:

$$Dev_t = \frac{1}{n}\sum_{i}^{1}\frac{abs(Time.\ Inf_{s,i} - Time.\ inf_{o,i})}{T}$$

Where $Time.Inf_{S,i}$ and $Time.Inf_{o,i}$ represent the time point (in days) in which sow farm i became infected in the simulated and observed datasets, respectively. T represents the total number of days in which the model ran (151 days), and n indicates the total number of sow farms. We then calculated the "fitness" of each parameter set as $\rho - Dev_t$. A higher fitness value indicates that a simulation produced dynamics that were more similar to the observed dynamics, with high correlations in the order in which sow farms were infected and low deviation in the day they became infected. The average fitness was calculated for each parameter set. Two hundred simulations were conducted per parameter set, yielding 200,000 simulations. Two hundred was based on an assessment of the number of simulations required to stabilize the coefficient of variation in fitness values (standard deviation/mean fitness).

2.2.3. Sensitivity analysis

We conducted a sensitivity analysis utilizing a random forest approach, a commonly used approach for global sensitivity analysis in epidemiological and ecological models that is particularly useful when the relationship between the outcome and parameters is non-linear (Harper et al., 2011; Hultquist et al., 2014; Wang et al., 2016; White et al., 2017). In this analysis, the outcome was the average fitness values from the simulation sets and parameter values were used as explanatory variables (White et al., 2017). Random forest is an ensemble machine learning method used for classification and regression which is based on the consensus of hundreds of randomized decision trees built using the explanatory variables (Criminisi et al., 2011). Unlike regression approaches, random forests can handle more complex data with numerous interacting and multicollinear variables with non-linear effects on the outcome (Hultquist et al., 2014; Wang et al., 2016), making them a highly suitable approach for understanding the marginal contribution of variation in each parameter on the overall fitness after controlling for the effects of all other parameters.

Briefly, the parameter sets were randomly divided into training (75%) and testing (25%) datasets, as recommended by (Kuhn and Johnson, 2013). The random forest was built using the training dataset with 250 trees and the number of variables considered per split set to three. In this analysis, a randomly selected 66% of observations were used to build each tree, and the remaining 33% were considered "out-of-bag" (OOB) and used to evaluate the importance of each predictor variable in that tree. Variable importance was assessed by calculating the percent increase in mean squared error (MSE) of predictions made for OOB observations relative the MSE of predictions when the variable is permuted. A greater increase would indicate that the variable is important for predicting fitness because randomizing the predictor variable results in greater prediction error (Breiman, 2001; Culter et al., 2007). To assess the overall performance of the model, the overall percent of variance in fitness explained by the random forest was calculated. The trained random forest was used to predict fitness values for the testing dataset, and the mean squared error was used to evaluate model performance. Partial dependence plots were constructed to assess the marginal effect of model parameters on model outputs after accounting for the effects of all other parameters.

2.2.4. Best-performing models

We selected five parameter sets as the best-performing models. Parameter sets were first ranked by their mean fitness in decreasing order, and the two parameter sets with the highest fitness values were included in the best-performing sets. In addition, from a visual assessment of the relationship between rank and fitness (Supplementary Fig. S1), there was a sudden drop in fitness values for models that ranked below 30. Given that there were potentially multiple combinations of parameters that produced high fitness models, we performed a k-means analysis on the top 30 sets to identify clusters of parameter sets that were similar to one another. All parameter values were centered and re-scaled for k-means analysis. Defining three clusters produced results with good separation among clusters, as shown by a classical discriminant coordinate plot (Supplementary Fig. S2). The mean parameter values for each of the three clusters were used as best-performing sets, for a total of five best parameter sets (Table 2).

One thousand simulations were run using parameters from each of the five best scenarios. For each scenario, the number of infected farms and infected sow farms was tabulated over time, and the overall proportion of transmission events due to each mechanism was recorded. In addition, we examined the proportion of "high consequence" transmission events were due to each mechanism, where high consequence events were defined as the first introduction of PEDv into new geographic areas that were previously spatially isolated from the epidemic (Region I and II, Fig. 2).

All modeling and analysis was performed in R version 3.3 (Team, 2013).

Table 2
Parameter values for five best-perfoming parameter sets.

Parameter set	μ	ϕ	ϕ_7	β_c	β_m	β_f	$p_{con.feed}$	F_A	F_B	Selection method	Rank correlation (ρ)	Dev_t	Fitness
A	0.593	0.008	8.685	~0.000	0.002	0.001	0.347	0	1	Top rank	0.43	0.20	0.23
B	0.506	0.017	1.902	~0.000	0.003	0.001	0.496	0	1	Top rank	0.45	0.20	0.25
C	0.812	0.011	6.730	0.002	0.002	0.004	0.513	0	1	k-means	0.39	0.23	0.17
D	0.651	0.007	8.033	0.002	0.001	0.001	0.556	0	1	k-means	0.40	0.21	0.19
E	0.637	0.016	3.344	0.002	0.004	0.003	0.529	0	1	k-means	0.40	0.23	0.17

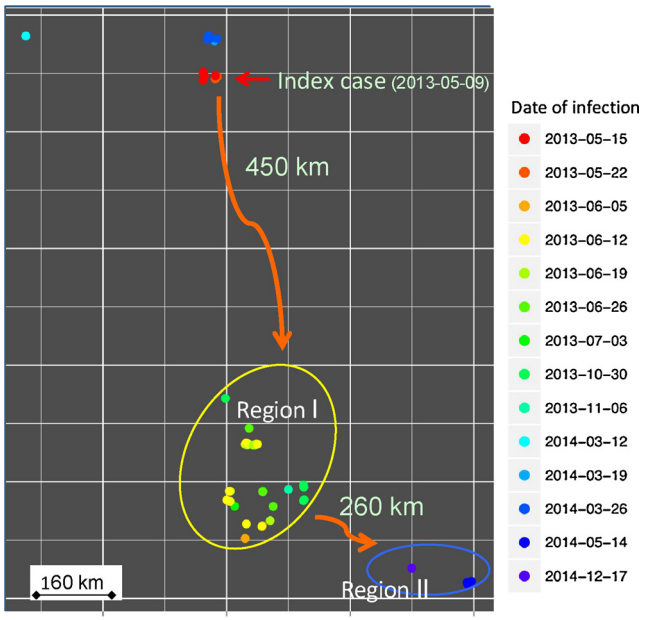

Fig. 2. Map of sow farms and dates of infection or "break".

3. Results

During the first five months (May through September 2013) of the observed PEDv epidemic in the studied production system, 30 of 47 sow farms were infected. The epidemic began slowly with just a few sow farms infected, followed by a sudden increase in infection in mid-June, which was associated with the introduction and spread of the virus in Region I (Fig. 2). No new sow farms were infected between July 3 until October 30, 2013, which is beyond the time period of our model.

3.1. Model calibration and sensitivity analysis

Fitness for each model run was calculated based on comparing the simulated and observed data in terms of the deviation in the day and rank correlation in the order in which sow farms become infected. The overall fitness of the model ranged from −0.39 to 0.25, with a median value of −0.02. Rank correlations in the order in which sow farms became infected in the observed versus simulated data ranged from 0.09 to 0.45, with a median of 0.28. Reported as a proportion, the percent deviation in the day on which sow farms became infected, Dev_t, ranged from 0.19 to 0.48, with a median of 0.29. Overall, the random forest explained 91.1% of the variance in fitness in the training data set, and validation with the testing dataset showed strong performance (mean squared error = 0.0013). The parameters that were most important for determining how well the PEDv model matched the observed data (i.e., models with higher fitness) were whether or not a feed contamination event occurred in feed mill B (F_B), values controlling the probability of local spatial spread (ϕ, ϕ_7), followed by the proportion of farms affected by a feed contamination event ($p_{con.feed}$ Supplementary Fig. S4). Model outputs, as measured by fitness, were less sensitive to the remaining parameters (μ, F_A, β_f, β_m, β_c)

After controlling for the effects of other parameters, the fitness of the PEDv model increased at lower values of β_c and β_m(Supplementary Fig. S5), indicating that models that simulated the observed epidemic tended to have low rates of within-flow and market truck-related transmission. Fitness also increased with higher values of ϕ, ϕ_7, β_f, and $p_{con.feed}$, indicating that more realistic models tended to have high rates of local spatial spread and feed truck-related transmission as well as a high numbers of farms affected by feed contamination events. Fitness was higher when a contamination event occurred in Mill B, but was not affected by Mill A. Similarly, the five best parameter sets that produced PEDv dynamics most resembling observed data all included a feed contamination event for Mill B and not for Mill A (Table 2). The model was not sensitive to the animal-level probability that an individual pig was infected in a batch of shipped (μ), most likely because the large batch sizes (median batch size of 284 head) meant that even a low animal-level infection probability translated into an almost certain probability of moving at least one infected pig.

3.2. Analysis of best-performing models

Simulations using the best-performing parameter sets generally showed a good match to the observed epidemiological curve (Fig. 3), with the number of sow farms infected slightly underestimated at the start of the outbreak and slightly overestimated at the end of the outbreak. In general, our model matched the observed data, showing a

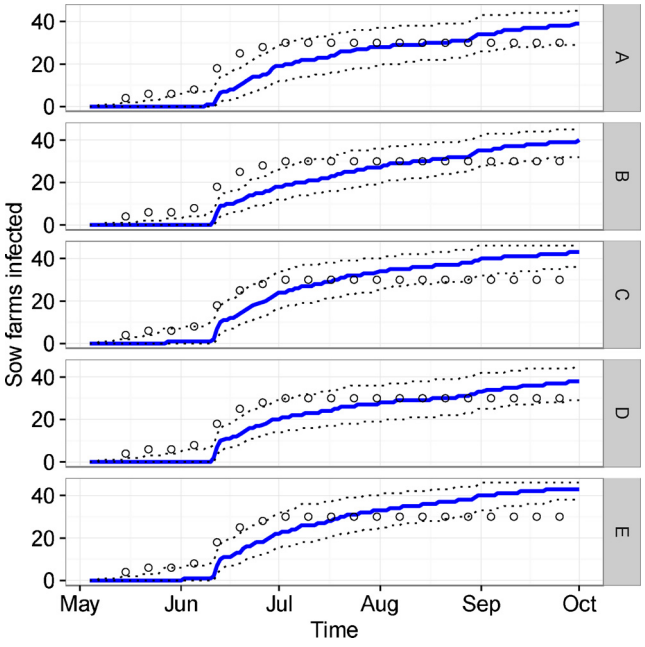

Fig. 3. Observed (circles) and simulated epidemic curves of the number of sow farms infected over time for the five best-performing models. The median, interquartile range, and 95% prediction interval of model simulations are represented by the blue line, shaded area, and dotted lines respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

slow initial rise in the number of infected farms (mainly non-sow farms), followed by a fairly sudden jump during June, which was related to the occurrence of the feed contamination event in Mill B (Fig. 3, Supplementary Fig. S6 for non-sow farms). Parameter sets A and D matched the data particularly well, with nearly all observed points falling within the 95% prediction interval of the model. However, it was difficult to match the exact order and time in which sow farms became infected. When comparing the interquartile range of predicted infection dates for every sow farm (Supplementary Fig. S3), the predicted interval encapsulates the observed dates most often for sets A and D. In addition, there is a general correspondence between the observed and predicted infection dates, with farms predicted to be infected in the middle or later in the study period also having observed dates that fell in middle or later for sets A and D. Mismatches occurred in the early study period, which lowered ρ. The model struggled to capture which specific farms should be infected early, but performance improved after the early period. For all five scenarios, the model predicted that over 65% of all farms (sow farms and other production types) were infected by October.

Across all best-performing models, the majority of between-farm transmission events were due to local spatial spread and the movement of pigs (Fig. 4, Supplementary Table S4-S6). When considering the mechanism of infection in all types of farms (Fig. 4a), pig movement was responsible for approximately one quarter to one third of transmission events, whereas this proportion was nearly half when considering only transmission to sow farms (Fig. 4b). Local spread was responsible for half of transmission events when considering all types of farms, and closer to one third when considering sow farms only (Fig. 4b). Thus, it appears that animal movements were relatively more important for the infection of sow farms whereas local spatial spread was relatively more important for non-sow farms, which could be related to the fact that many non-sow farms were located in areas with higher farm density (Fig. 1). Contaminated feed was responsible for a greater proportion of sow farm infections compared to the proportion attributed to this mechanism for all farms. The remaining transmission mechanisms accounted for a relatively low percentage of events, regardless of whether we are examining the responsible mechanisms for the infection of sow farms or all farms (Fig. 4). Transmission mechanism for non-sow farms only were very similar to those of all farms (Supplementary Table S6 and Fig. S7).

The initial introduction of PEDv into region I was attributable to a contaminated feed event in 69.4% of simulations and to movement in 20.7% of simulations (Table 3). However, the model predicted that this introduction would occur after day 34 in 75% of simulations, whereas the observed date of introduction to region I was day 22. Thus, we restricted the summary of transmission mechanisms to only those simulations (n = 827) where introduction occurred prior to day 22 and evaluated the transmission mechanism responsible for the initial introduction. In this case, animal movement and feed trucks were responsible for 72.9% and 20%, respectively, of the initial transmission of PEDv to Region I. Potentially, there could have been multiple introductions into region I beginning, with the earlier introduction prior to day 22 followed by a contaminated feed event. Thus, our results do not rule out a possible role of contaminated feed in long-distance transmission events. Introduction of PEDv to region II was attributable to market trucks and within-flow transmission in 59.2% and 21.8% of simulations where region II was infected, respectively (Table 3).

Table 3
Proportion of introductions to new geographic areas due to each transmission mode.

	Region I	*Region I < 22 days*	*Region II*
Distance from infected area (mi)	280	280	160
Contaminated feed	0.694	0.000	0.000
Feed trucks	0.063	0.200	0.055
Within-flow	0.000	0.000	0.218
Local	0.011	0.029	0.016
Market trucks	0.024	0.042	0.592
Movement	0.207	0.729	0.051

4. Discussion

Here, we developed a mechanistic model to simulate the rapid spread of PEDv within a swine production system to better understand the drivers of between-farm viral transmission during the emergence of PEDv in the U.S. swine industry. Prior to 2013, PEDv did not occur in the United States. As a result of the rapid spread and production impact of this formerly exotic disease, PEDv is considered an industry game-changer and highlights the vulnerability of the U.S. swine industry to foreign pathogens. However, mechanisms driving between-farm spread are still poorly understood. By calibrating our PEDv spread model such that simulated data matched observed epidemiological dynamics, we demonstrated that while all hypothesized mechanisms for direct and indirect transmission between-farms likely occurred to some extent, most transmission events were attributable to movement of infected

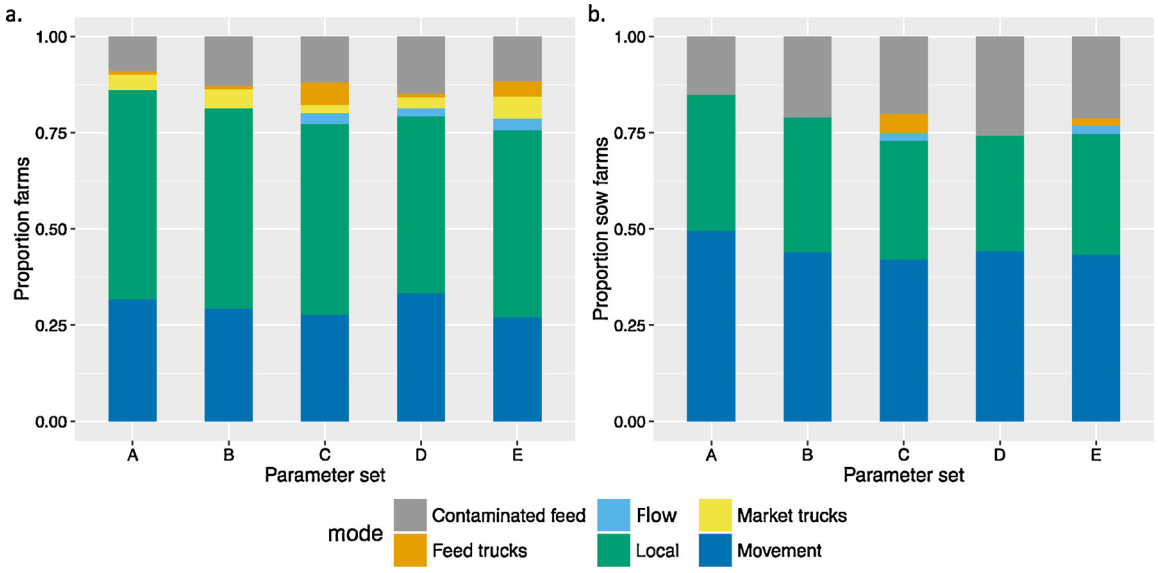

Fig. 4. Proportion of transmission events attributed to each mechanism for infection of a) all farms and b) sow farms.

animals (direct transmission) and local spatial spread (indirect transmission). However, other indirect transmission mechanisms, such as fomites related to feed trucks, market trucks, and shared support services within flows, became more important for high consequence, long distance transmission events.

Overall, our model provided adequate but not outstanding match to the observed data. The slight under- and over-estimate of the number of sow farms infected during the early and later parts of the epidemic, respectively (Fig. 3), potentially indicate that our model did not capture some of the underlying changes in how PEDv was managed at the farm-level. During the early part of the epidemic, there was a delay in identifying the infectious agent, and PEDv was novel to swine veterinarians and producers. However, the industry adapted to PEDv and became better at mitigating between-farm spread through enhanced biosecurity, and this may explain why our model overestimates the number of infected farms in the latter stage of the modeled period.

Alternatively, even in the best-performing parameter sets, such as A and D, our model struggled to capture which sow farms should be infected early in the epidemic. This suggests that the model did a poorer of capturing epidemiological processes at the beginning of the epidemic, though model performance improved in the middle to late portions of the study period. This contributed to the moderate rank correlation values ($\rho \sim 0.39$–0.49) and model fitness. It may have been easier to match the model to generalized patterns of infection dates (i.e., early, middle, late) rather than the exact order of infection.

Another source of discrepancy between the simulated and observed epidemics is the infection of Region II. Region II contains most of the genetic stock (i.e., high value breeding stock) for this company, and thus these farms are subject to enhanced biosecurity measures. Although Region II eventually became infected in the observed epidemic, Region II did not become infected during the first five months. In contrast, the model predicted that Region II would become infected, and this discrepancy may be due to a failure to account for enhanced biosecurity measures utilized for genetic stock. Despite this discrepancy, it is still useful to track the mechanisms by which Region II became infected as this may highlight potential vulnerabilities in biosecurity measures. Region II usually became infected by either market trucks or within-flow transmission, which emphasizes the importance of decontaminating market trucks and preventing mechanical vectors or fomites associated with shared support services within flows. Also, our model's structure likely overestimated the role market truck transmission to Region II, given that all animals going to slaughter from Region II passed through transfer stations (market trucks did not directly visit Region II farms).

More generally, we did not account for other variabilities in biosecurity measures at the farm-level. For example, sow farms typically have higher biosecurity than nurseries and finishers, which may decrease their likelihood of becoming infected via local spatial spread or fomite-related mechanisms. Similarly, biosecurity and mitigation measures employed by infected farms may lead to fluctuating infectiousness over time, which we did not allow for beyond the first week. This could lead to an overestimation of how long farms were able to transmit the virus. Our assumptions about the length of the infectious period of all farms were also based on data from sow farms, as those were the only data that were available.

Our generalized approach to fomite-related spread was based on the force-of-infection amongst farms that had indirect connections via feed trucks, market trucks, or shared within-flow support services. In reality, the likelihood of infection by any of these routes would not be shared equally among farms; companies carefully manage the order in which farms are visited in the course of a week such that high-priority farms (i.e., farms where it is more important to prevent infection, such as sow farms) are visited early in the week, thus minimizing the likelihood of fomite contamination of personnel, vehicles, and equipment. In addition, there is often down-time between visits to "clean" and "dirty" farms (in terms of infection) to decrease the likelihood of transporting fomites between farms. These factors were not accounted for in our model, and may have resulted in sow farms becoming infected too easily. Our model also predicted that few transmission events were due to fomite-related mechanisms. However, it is possible that the roles of fomite-related mechanisms were underestimated given that groups of farms that share the same mill or flow tend to be spatially clustered. This means that it is possible that local spatial spread soaked up at least some of the transmission that could have actually been mediated by fomites, and that our model could not distinguish between fomite transmission and local spatial spread. That being said, many of the flows and mills had farms that were highly dispersed and/or overlapped with other flows/mills, suggesting that indirect transmission via feed trucks or within flows and local spatial spread were not analogous.

Results of our analysis indicated that feed contamination played a minor but substantial role in the rapid spread of PEDv, accounting for ~11% and 20% of transmission events in non-sow and sow farms respectively. Further, long-distances jumps of PEDv in simulated outbreaks were often attributed to this mechanism (Table 3). However, our results are specific to transmission within one production system. Contaminated feed may have played a larger role in the initial introduction of the virus to production systems, which is not addressed here, and more generally to the introduction of PEDv to the United States. In addition, our approach for modeling contaminated feed and feed milling were very generalized, and we did not account for the nuances of various feed ingredients that may be more or less likely to harbor live virus (Dee et al., 2015).

Our work should be interpreted in the context in which the model was calibrated, and care should be taken in extrapolating these results to other regions of the country or to the other time periods. For example, the landscape in the Great Plains is markedly more windy and open than in other swine dense regions of the United States, which may have facilitated windborne transmission. In addition, this model focused on the initial spread of PEDv, and between-farm transmission dynamics during the emerging phase of the epidemic may differ from the current endemic situation.

5. Conclusions

During the early phase of the PEDv epidemic in U.S. swine, results of our model indicate that between-farm transmission was primarily driven by direct mechanisms via the movement of animals and indirect mechanisms related to local area spread based on geographic proximity between farms. However, other forms of indirect contact among farms were responsible for high consequence transmission events resulting in the introduction of the virus into new geographic areas. Lessons learnt from the PEDv epidemic in regards to how best represent the complexity of the U.S. swine industry in epidemiological models are valuable to monitoring other emerging swine pathogens. Our model is a more realistic representation to the realities of the swine industry than other between-farm modeling approaches given its ability to account for management practices and patterns of direct and indirect contact among farms unique to the vertically integrated nature of the swine industry and it could be readily adapted to model the spread of other foreign and endemic pathogens of swine in the US. Thus, this work provides insights into the spread of an exotic pathogen in US swine and provides a foundation for future infectious disease modeling efforts, highlighting the potential explosive nature of an epidemic in a production system with high levels of both direct and indirect connectivity among farms.

Acknowledgements

This material is based upon work supported by the Cooperative State Research Service, U.S. Department of Agriculture, under Project No. MINV-62-044. We would like to thank company veterinarians and the Morrison Swine Health Monitoring Project team for sharing data,

insights, feedback relevant to the spread of this disease.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.epidem.2018.04.001.

References

Alonso, C., Goede, D.P., Morrison, R.B., Davies, P.R., Rovira, A., Marthaler, D.G., Torremorell, M., 2014. Evidence of infectivity of airborne porcine epidemic diarrhea virus and detection of airborne viral RNA at long distances from infected herds. Vet. Res. 45, 73.

Alvarez, J., Sarradell, J., Morrison, R., Perez, A., 2015. Impact of porcine epidemic diarrhea on performance of growing pigs. PLoS One 10, e0120532.

Alvarez, J., Goede, D., Morrison, R., Perez, A., 2016. Spatial and temporal epidemiology of porcine epidemic diarrhea (PED) in the Midwest and Southeast regions of the United States. Prev. Vet. Med. 123, 155–160.

Arruda, A.G., Friendship, R., Carpenter, J., Greer, A., Poljak, Z., 2016. Evaluation of control strategies for porcine reproductive and respiratory syndrome (PRRS) in swine breeding herds using a discrete event agent-based model. PLoS One 11, e0166596.

Bajardi, P., Barrat, A., Savini, L., Colizza, V., 2012. Optimizing surveillance for livestock disease spreading through animal movements. J. R. Soc. Interface 9, 2814–2825.

Beam, A., Goede, D., Fox, A., McCool, M.J., Wall, G., Haley, C., Morrison, R., 2015. A porcine epidemic diarrhea virus outbreak in one geographic region of the United States: descriptive epidemiology and investigation of the possibility of airborne virus spread. PLoS One 10, e0144818.

Blower, S.M., Dowlatabadi, H., 1994. Sensitivity and uncertainty analysis of complex-Models of disease transmission – an Hiv model, as an example. Int. Stat. Rev. 62, 229–243.

Bowman, A.S., Krogwold, R.A., Price, T., Davis, M., Moeller, S.J., 2015. Investigating the introduction of porcine epidemic diarrhea virus into an Ohio swine operation. BMC Vet. Res. 11, 38.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brooks-Pollock, E., de Jong, M.C., Keeling, M.J., Klinkenberg, D., Wood, J.L., 2015. Eight challenges in modelling infectious livestock diseases. Epidemics 10, 1–5.

CAHFS, 2013. Porcine Epidemic Diarrhea Virus (PEDV). Center for Animal Health and Food Safety, University of Minnesota, St. Paul, MN.

Craft, M.E., 2015. Infectious disease transmission and contact networks in wildlife and livestock. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, 20140107.

Crawford, K., Lager, K., Miller, L., Opriessnig, T., Gerber, P., Hesse, R., 2015. Evaluation of porcine epidemic diarrhea virus transmission and the immune response in growing pigs. Vet. Res. 46, 49.

Criminisi, A., Shotton, J., Konukoglu, E., 2011. Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning. Microsoft Research Cambridge, Tech. Rep. MSRTR-2011-114, vol. 5. pp. 12.

Culter, D.R., Edwards, T.C.J., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.

Dee, S., Clement, T., Schelkopg, A., Nerem, J., Knudsen, D., Christopher-Hennings, J., Nelson, E., 2014. An evaluation of contaminated complete feed as a vehicle for porcine epidemic diarrhea virus infection of naive pigs following consumption via natural feeding behavior: proof of concept. BMC Vet. Res. 10, 176.

Dee, S., Neill, C., Clement, T., Singrey, A., Christopher-Hennings, J., Nelson, E., 2015. An evaluation of porcine epidemic diarrhea virus survival in individual feed ingredients in the presence or absence of a liquid antimicrobial. Porcine Health Manage. 1, 9.

Goede, D., Morrison, R.B., 2016. Production impact and time to stability in sow herds infected with porcine epidemic diarrhea virus (PEDV). Prev. Vet. Med. 123, 202–207.

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. Lond. B 273, 2729–2735.

Harper, E.B., Stella, J.C., Fremier, A.K., 2011. Global sensitivity analysis for complex ecological models: a case study of riparian cottonwood population dynamics. Ecol. Appl. 21, 1225–1240.

Huang, Y.W., Dickerman, A.W., Pineyro, P., Li, L., Fang, L., Kiehne, R., Opriessnig, T., Meng, X.J., 2013. Origin, evolution, and genotyping of emergent porcine epidemic diarrhea virus strains in the United States. MBio 4, e00737–00713.

Hultquist, C., Chen, G., Zhao, K., 2014. A comparison of Gaussian process regression, random forests and support vector regression for burn severity assessment in diseased forests. Remote Sens. Lett. 5, 723–732.

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. Interface 4, 907–916.

Kao, R.R., 2002. The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. Trends Microbiol. 10, 279–286.

Keeling, M.J., Brooks, S.P., Gilligan, C.A., 2004. Using conservation of pattern to estimate spatial parameters from a single snapshot. Proc. Natl. Acad. Sci. 101, 9155–9160.

Kim, Y., Yang, M., Goyal, S.M., Cheeran, M.C., Torremorell, M., 2017. Evaluation of biosecurity measures to prevent indirect transmission of porcine epidemic diarrhea virus. BMC Vet. Res. 13, 89.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York.

Lee, K., Polson, D., Lowe, E., Main, R., Holtkamp, D., Martinez-Lopez, B., 2017. Unraveling the contact patterns and network structure of pig shipments in the United States and its association with porcine reproductive and respiratory syndrome virus (PRRSV) outbreaks. Prev. Vet. Med. 138, 113–123.

Legrand, J., Sanchez, A., Le Pont, F., Camacho, L., Larouze, B., 2008. Modeling the impact of tuberculosis control strategies in highly endemic overcrowded prisons. PLoS One 3.

Lowe, J., Gauger, P., Harmon, K., Zhang, J., Connor, J., Yeske, P., Loula, T., Levis, I., Dufresne, L., Main, R., 2014. Role of transportation in spread of porcine epidemic diarrhea virus infection, United States. Emerg. Infect. Dis. 20, 872–874.

Madson, D.M., Magstadt, D.R., Arruda, P.H.E., Hoang, H., Sun, D., Bower, L.P., Bhandari, M., Burrough, E.R., Gauger, P.C., Pillatzki, A.E., Stevenson, G.W., Wilberts, B.L., Brodie, J., Harmon, K.M., Wang, C., Main, R.G., Zhang, J., Yoon, K.J., 2014. Pathogenesis of porcine epidemic diarrhea virus isolate (US/Iowa/18984/2013) in 3-week-old weaned pigs. Vet. Microbiol. 174, 60–68.

Martinez-Lopez, B., Ivorra, B., Ramos, A.M., Sanchez-Vizcaino, 2011. A novel spatial and stochastic model to evaluate the within- and between-farm transmission of classical swine fever virus. I. Genetal concepts and description of the model. Vet. Microbiol. 147, 300–309.

O'Dea, E., Snelson, H., Bansal, S., 2015. State-Level Transport Flows are Predictive of the Dynamics of Porcine Epidemic Diarrhea Virus. bioRxiv.

Pasick, J., Berhane, Y., Ojkic, D., Maxie, G., Embury-Hyatt, C., Swekla, K., Handel, K., Fairles, J., Alexanderson, S., 2014. Investigation into the role of potentially contaminated feed as a source of the first-detected outbreaks of Porcine Epidemic Diarrhea in Canada. Transbound. Emerg. Dis. 61, 397–410.

Rossi, G., De Leo, G.A., Pongolini, S., Natalini, S., Vincenzi, S., Bolzoni, L., 2015. Epidemiological modelling for the assessment of bovine tuberculosis surveillance in the dairy farm network in Emilia-Romagna (Italy). Epidemics 11, 62–70.

Saif, L.J., Pensaert, M.B., Sestak, K., Yeo, S., Jung, K., 2012. Coronaviruses. In: Zimmerman, J.J., Karriker, L.A., Ramirez, A., Schwartz, K.J., Stevenson, G.W. (Eds.), Diseases of Swine, 10th ed. John Wiley & Sons Ltd, Ames, IA, pp. 501–524.

Song, D.-S., Kang, B.-K., Lee, S.-S., Yang, J.-S., Moon, H.-J., Oh, J.-S., Ha, G.-W., Jang, Y.-S., Park, B.-K., 2006. Use of an internal control in a quantitative RT-PCR assay for quantitation of porcine epidemic diarrhea virus shedding in pigs. J. Virol. Methods 133, 27–33.

Stevenson, G.W., Hoang, H., Schwartz, K.J., Burrough, E.R., Sun, D., Madson, D., Cooper, V.L., Pillatzki, A., Gauger, P., Schmitt, B.J., Koster, L.G., Killian, M.L., Yoon, K.J., 2013. Emergence of Porcine epidemic diarrhea virus in the United States: clinical signs, lesions, and viral genomic sequences. J. Vet. Diagn. Invest. 25, 649–654.

Szmaragd, C., Wilson, A.J., Carpenter, S., Wood, J.L., Mellor, P.S., Gubbins, S., 2009. A modeling framework to describe the transmission of bluetongue virus within and between farms in Great Britain. PLoS One 4, e7741.

Team, R.C, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria(URL http://www.r-project.org/).

Thakur, K.K., Sanchez, J., Hurnik, D., Poljak, Z., Opps, S., Revie, C.W., 2015. Development of a network based model to simulate the between-farm transmission of the porcine reproductive and respiratory syndrome virus. Vet. Microbiol. 180, 212–222.

Valdes-Donoso, P., VanderWaal, K., Jarvis, L.S., Wayne, S.R., Perez, A.M., 2017. Using machine learning to predict swine movements within a regional program to improve control of infectious diseases in the US. Front. Vet. Sci.: Vet. Epidemiol. Econ. 19, 1–13.

VanderWaal, K., Enns, E.A., Picasso, C., Alvarez, J., Perez, A., Gil, A., Fernandez, F., Craft, M.E., Wells, S., 2017. Optimal surveillance strategies for bovine tuberculosis in a low-prevalence country. Sci. Rep. 7, 4140.

Wang, B., Oldham, C., Hipsey, M.R., 2016. Comparison of machine learning techniques and variables for groundwater dissolved organic nitrogen prediction in an urban area. Procedia Eng. 154, 1176–1184.

White, L.A., Torremorell, M., Craft, M.E., 2017. Influenza A virus in swine breeding herds: combination of vaccination and biosecurity practices can reduce likelihood of endemic piglet reservoir. Prev. Vet. Med. 138, 55–69.

Wu, J.Y., Dhingra, R., Gambhir, M., Remais, J.V., 2013. Sensitivity analysis of infectious disease models: methods, advances and their application. J. R. Soc. Interface 10.

Yadav, S., Olynk Widmar, N.J., Weng, H.Y., 2016. Modeling classical swine fever outbreak-related outcomes. Front. Vet. Sci. 3, 7.