

Assessment of the adequacy of available spatial data of the demographics of swine populations in Minnesota

Spencer R. Wayne & Peter R. Davies

Summary

Reliable data on the demographics of animal populations are essential for effective planning and execution of disease control programmes. To document the spatial distribution of different swine populations in Minnesota, the authors evaluated four sources of data from separate entities that maintain data independently and for different purposes. Although the total numbers of swine sites (aggregated at the county level) were significantly correlated among datasets, analysis of spatial clustering patterns demonstrated regional biases among the datasets. We used current, field verified data of farm locations in two counties to identify and quantify inaccuracies in two databases recording individual farm locations. The proportion of omitted or erroneous farm sites and the magnitude of positional inaccuracies were sufficient to limit the utility of available data for analytic purposes or for disease control efforts. There is a clear need for more current and accurate demographic data to underpin industry or government initiatives to control swine disease in Minnesota. Current efforts under the National Animal Identification System may address this concern.

Keywords

Demographics, Geographic information system, Minnesota, Spatial data, Swine, United States of America.

Valutazione dell'adeguatezza dei dati spaziali disponibili sulla demografia della popolazione suina nel Minnesota

Riassunto

Dati affidabili sulla demografia della popolazione animale risultano essenziali per una corretta pianificazione ed esecuzione di programmi di controllo delle malattie. Per documentare la distribuzione spaziale di diverse popolazioni suine del Minnesota, sono state valutate quattro fonti separate di dati che vengono conservati in maniera indipendente e per differenti scopi. Sebbene il numero totale dei punti di concentrazione degli allevamenti suini (aggregati a livello di contea) fosse significativamente correlato tra i data set, l'analisi dei modelli di aggregazione spaziale ha dimostrato bias regionali di dataset. Sono stati utilizzati di norma dati di campo verificati riguardo l'ubicazione degli allevamenti in due contee al fine di identificare e quantificare la non accuratezza in due database registranti l'ubicazione dei singoli allevamenti. La proporzione delle localizzazioni degli allevamenti omesse od errate così come la grandezza dello scostamento posizionale è stata sufficiente per limitare l'utilità dei dati disponibili a fini analitici e per il controllo delle malattie. E' evidente la necessità di dati demografici più aggiornati ed accurati per supportare le iniziative dell'industria e del governo per il controllo delle malattie dei suini nel Minnesota. In questo contesto vanno inseriti gli sforzi del sistema nazionale di identificazione animale.

College of Veterinary Medicine, University of Minnesota, 1365 Gortner Ave, St Paul, MN 55108, United States of America
wayn003@umn.edu

Parole chiave

Dati spaziali, Demographics, Minnesota, Sistema informativo geografico, Stati Uniti d'America, Suini.

Introduction

The ability to analyse disease patterns in a geographic (spatial) context is contingent upon knowledge of the demographics of the species of concern (population at risk). For transmissible diseases, the proximity of livestock populations might influence the probability of transmission between farms, and the presence of diseased animals at a given location might influence the risk of infection for susceptible populations within its vicinity. Consequently, knowledge of locations and disease status of neighbouring farms influences animal health decisions made by public or private veterinarians and producers.

In many countries, including the United States, government policy stipulates response measures to an incursion of a foreign animal disease (1, 2, 7). Geographic knowledge of the locations of animal populations is vital to containing epidemics through identification and quarantine of high-risk populations and movement restrictions. Spatial modelling techniques can be applied to predict the likely spread of disease and support resource allocation and decision-making (6). The efficiency of emergency responses is therefore partly a function of the quality of spatial data of the populations at risk. These principles equally apply to efforts to control endemic diseases. In the United States, porcine reproductive and respiratory syndrome (PRRS) is widely acknowledged as the major swine health problem. The ability of the PRRS virus to spread locally between farms, despite substantial investments in biosecurity, has prompted calls for 'regional' control efforts to reduce the impact of the disease (4, 5).

The purpose of this study was to assess available sources of data on the demographics of swine populations in Minnesota to support the planning of disease control initiatives. Specific objectives of the study were to:

- compare publicly available datasets of the spatial distribution of swine farms in Minnesota
- assess the positional accuracy of farm locations in two of these databases in relation to verified farm locations
- describe the spatial distribution of a non-commercial swine population (4-H pigs) in relation to commercial swine production.

Materials and methods

Data sources

Five available datasets recording the demographics of swine holdings in Minnesota were identified and obtained from their respective sources, as described below.

National Agricultural Statistics Census

National Agricultural Statistics (NASS) Census – 2002 of the United States Department of Agriculture (USDA) conducts a national survey of agricultural activities every five years. Data are obtained by farmer surveys and aggregated at the county level for analysis and publication. For this analysis, the most recent (2002) data for swine production were downloaded from the NASS website (www.nass.usda.gov/Census/). The NASS census data does not include point coordinates for individual farm premises and could include multiple sites as a single 'farm', because data was collected at the owner level (a single owner may operate more than one site).

Minnesota Board of Animal Health

The Minnesota Board of Animal Health (MBAH) is responsible for the control of regulated animal diseases in the state of Minnesota. The MBAH maintains data on locations of premises producing livestock for all major livestock sectors, specifically for the purpose of supporting disease control activities. The MBAH swine farm locations data has been compiled over many years and

most recently was used in the pseudorabies virus (Aujeszky's disease) eradication campaign. The data neither indicate the type of production (for example, farrowing, nursery, finishing) nor the number of animals reared on a given premises. For the purposes of this analysis, data on swine operations obtained in 2004 were used. A farm's county location was not recorded in the database, but was derived by using farm coordinates data spatially joined to counties data (ArcGIS™ version 9, Environmental Systems Research Institute, Inc., Redlands, California).

Minnesota Pollution Control Agency

The Minnesota Pollution Control Agency (MPCA) is responsible for regulating potential sources of environmental pollution in the state of Minnesota. These include livestock operations, for which the agency issues permits for manure management. Farms recorded in the MPCA database are those farms holding valid feedlot permits. Data on feedlot operations are compiled and maintained by county and state offices. They include the species and total number of animal units for which the permit is issued (measure of herd size). Exceeding a threshold capacity of 50 animal units necessitates registration with MPCA (equivalent to 125 pigs weighing more than 136 kg each). Data for premises registered as producing swine in 2004 were used for analysis.

4-H programme

The 4-H programme is a youth educational programme that commenced in several states in the late 1800s to provide an introduction to agricultural industries (3). Typically, 4-H youth livestock programmes allow for an annual experience that involves purchasing one or more young pigs that are reared to market weight and exhibited at various show events. Data on registered 4-H participants in Minnesota were obtained from the state 4-H office. The data included the county, but not the geographic coordinates, of premises. Additionally, a survey was sent to a random sampling of 200 4-H swine programme participants that were at least in 7th grade of school (approximately 12 years old or older). This survey tool was designed to assess the

husbandry practices of 4-H swine participants as well as their attitudes about and knowledge of swine diseases and biosecurity measures.

To evaluate the accuracy of the individual farm locations data in the MBAH and MPCA databases, recent field verified data (2006) provided a reliable 'gold standard' for specific farm locations (R. Morrison and E. Mondaca, personal communication). The data were collected as part of a pilot PRRS control project in Stevens County and the eastern half of Rice County, Minnesota. Based on NASS survey data among Minnesota counties, Rice County and Stevens County rank at the 83rd and 65th percentiles, respectively, for pig farm density and at the 76th and 57th percentiles, respectively, for the number of pig farms per county.

Software

Microsoft Office Access® 2003 was used for data aggregation and reprocessing for further analysis. ArcGIS™ 9.0 was used for warehousing and presenting the geographic data. Statistix® 8.0 (Analytical Software, Tallahassee, Florida) was used for general statistical analyses, and GeoDA® 9.5 (Luc Anselin, Illinois) was used for county-level comparisons.

Analyses

Comparison of spatial distributions of swine farm premises by county in Minnesota, as recorded in the NASS, MBAH and MPCA databases

For all three datasets, descriptive statistics were generated and the Spearman rank correlation was used to estimate correlation of county farm counts among databases. To demonstrate the agreement among datasets, GeoDA® was used to create a series of descriptive maps. Each dataset was graphed as pig farm density per county (number of farms per 100 km²) and quartiles of farm densities were mapped. County rates were then smoothed using a 1st order Queen's contiguity weighting file in GeoDA®. A 1st order weighting matrix was selected because of its inherent simplicity and the relatively large size of the counties compared to the expected range of local disease spread from individual farms.

Spatial clustering

GeoDA® was used to test for spatial clustering and to compare differences in farm counts among databases. The univariate local indicator of spatial autocorrelation (LISA) with empirical Bayes (EB) rate procedure in GeoDA® was used to determine clustering of farm density. Univariate LISA demonstrates how values for contiguous areas are spatially similar or dissimilar, and the EB rate was used to standardise the county counts over county area (100 km²). Moran's I value (with the associated pseudo *p*-value) was calculated following smoothing with 1st order Queen's contiguity weighting. Graphical representation of clustering was displayed for high-high, low-low, high-low, and low-high patterns.

Clustering of differences between datasets

To better understand possible regional biases in datasets, the differences in counts between the three datasets (NASS minus MBAH, NASS minus MPCA, and MPCA minus MBAH) were calculated for each county. The univariate LISA with EB rate was again used to describe the spatial distributions of these differences. Since none of the three datasets is a 'gold standard', this analysis will reveal potential regional bias in one dataset relative to another. The over or underestimation of a regional difference cannot be precisely known, but the regional trends in variation will be better understood.

Evaluation of the accuracy of farm location data in the MBAH and MPCA databases in two Minnesota counties

In two Minnesota counties (Stevens and the eastern half of Rice), exact farm locations are known for all active swine operations (81 in Stevens County and 53 in eastern Rice County), as of 2006. The geographic coordinates for farms in the MBAH and MPCA datasets are also known. A system relying on common name (either owner name or premises name) and proximity (points located within 10 km of each other) was used to create a common index for farms identified in all three datasets. Farms were manually matched, using knowledge of local names, potential misspellings, farm-coding differences and

ownership changes. While imperfect, this method allowed for a more complete site-by-site matching than might a rigid application of validation rules. The overall outcome of processing was that farms listed in multiple databases and possessing a common index number could be linked and compared across datasets. To evaluate the completeness and accuracy of the datasets, missing farms (existing farms not recorded in a database) and false-positive farms (recorded in a database but not currently active) of MBAH and MPCA datasets were used to derive two probability values. These were the probability that a dataset includes an actual, existent farm site and the probability that a farm listed in either the MBAH or the MPCA dataset was actually a current swine operation. For farms listed in all three datasets, the geographic error was calculated as the Euclidean distance between sites with the same index number. Basic descriptive statistics were generated including mean, standard deviation, mean centre, and standard distance. After all possible sites in the three databases were indexed, all verified farm locations were assigned identical x, y coordinates to fix the true farm location to a single point in space. The corresponding matched farm locations from the MPCA and MBAH datasets were likewise transformed such that their derived coordinates accurately represented their location relative to the respective verified 'true' farm location. This enabled both visualisation of the positional error for all commonly indexed farms in the MBAH and MPCA databases and calculation of positional accuracy and precision for each dataset relative to the true farm location.

Spatial distribution of a non-commercial swine population (4-H pigs) in relation to commercial swine production

Data on six years of 4-H swine programme participation were collected and compiled by county. Descriptive statistics were generated for these data. Spearman rank correlation was performed to measure the county-level association between the six year average 4-H participation (average participants per county) and commercial swine production (number of swine farms in the NASS dataset) at county

level. In GeoDA®, 4-H participation density (raw rate and smoothed rate) was graphically rendered for evaluation. Spatial clustering of 4-H density was assessed with univariate LISA with EB rate modification and bivariate LISA analysis was performed to test for spatial correlation of 4-H participation and spatial correlation of 4-H participation with commercial swine production (NASS data). A 1st order Queen’s contiguity weighting file was used for spatial smoothing. Bivariate LISA allows the value of one polygon (county) to be related to the different values of surrounding polygons to determine whether 4-H participation is spatially correlated clustered with commercial production.

Results

Comparison of spatial distributions of swine farm premises by county in Minnesota, as recorded in the NASS, MBAH and MPCA databases

At the coarsest level, estimates of the total number of swine farms in Minnesota ranged from 5 499 in NASS database to 10 768 farms recorded in the MBAH database (Table I). Significant positive correlation ($P<0.0001$) was observed among county counts of swine farms in all three databases (Table II).

Table I
Minnesota swine farm counts

Farm	State total	County average	County standard deviation
NASS	5 626	65	53
MBAH	10 768	124	111
MPCA	8 331	96	115

NASS National Agricultural Statistics
MBAH Minnesota Board of Animal Health
MPCA Minnesota Pollution Control Agency

All three datasets indicate that swine farms are geographically clustered in the state. Moran’s I values ($p<0.001$) were 0.79 for the NASS data, 0.68 for the MPCA data and 0.69 for the MBAH data. All datasets indicated that the density of swine farms is greater in the southern counties of Minnesota (Fig. 1). The raw density maps and spatially smoothed

density maps both reveal that MBAH data has relatively higher numbers of swine farms recorded in the central portion of the state than the other datasets. Univariate LISA with EB rate revealed clustering of high density counties with other high density counties and vice-versa (Fig. 2). NASS and MPCA data show clustering along the southern border, while MBAH density clusters in the southern and central portion of the state. The maps indicate that the three datasets have different spatial clustering patterns for swine farm density.

Table II
Spearman rank correlation of county swine farm counts from different datasets (p -values <0.0001)

County swine farm data pairs	Spearman rank correlation R
MBAH-NASS	0.87
MPCA-NASS	0.92
MBAH-MPCA	0.86

MBAH Minnesota Board of Animal Health
NASS National Agricultural Statistics
MPCA Minnesota Pollution Control Agency

With respect to the numeric differences in county counts of farms between datasets, very clear evidence of regional differences between datasets was seen (Fig. 3). The Moran’s I values were moderate in range for the three comparisons (0.54 for NASS-MBAH, 0.47 for NASS-MPCA, and 0.40 for MPCA-MBAH) and were all significant (pseudo $p<0.001$). MBAH overestimated farm density in central Minnesota and underestimated density in northern Minnesota (relative to NASS data). MPCA overestimated farm density in the south and underestimated density across a portion of the north (again relative to NASS data). Relative to the MPCA data, MBAH data record more farms in the central portion of the state and less in the south.

Evaluation of the accuracy of farm location data in the MBAH and MPCA databases in two Minnesota counties

A substantial proportion (16% for MBAH and 25% for MPCA) of operational swine farms in both counties were not recorded in each

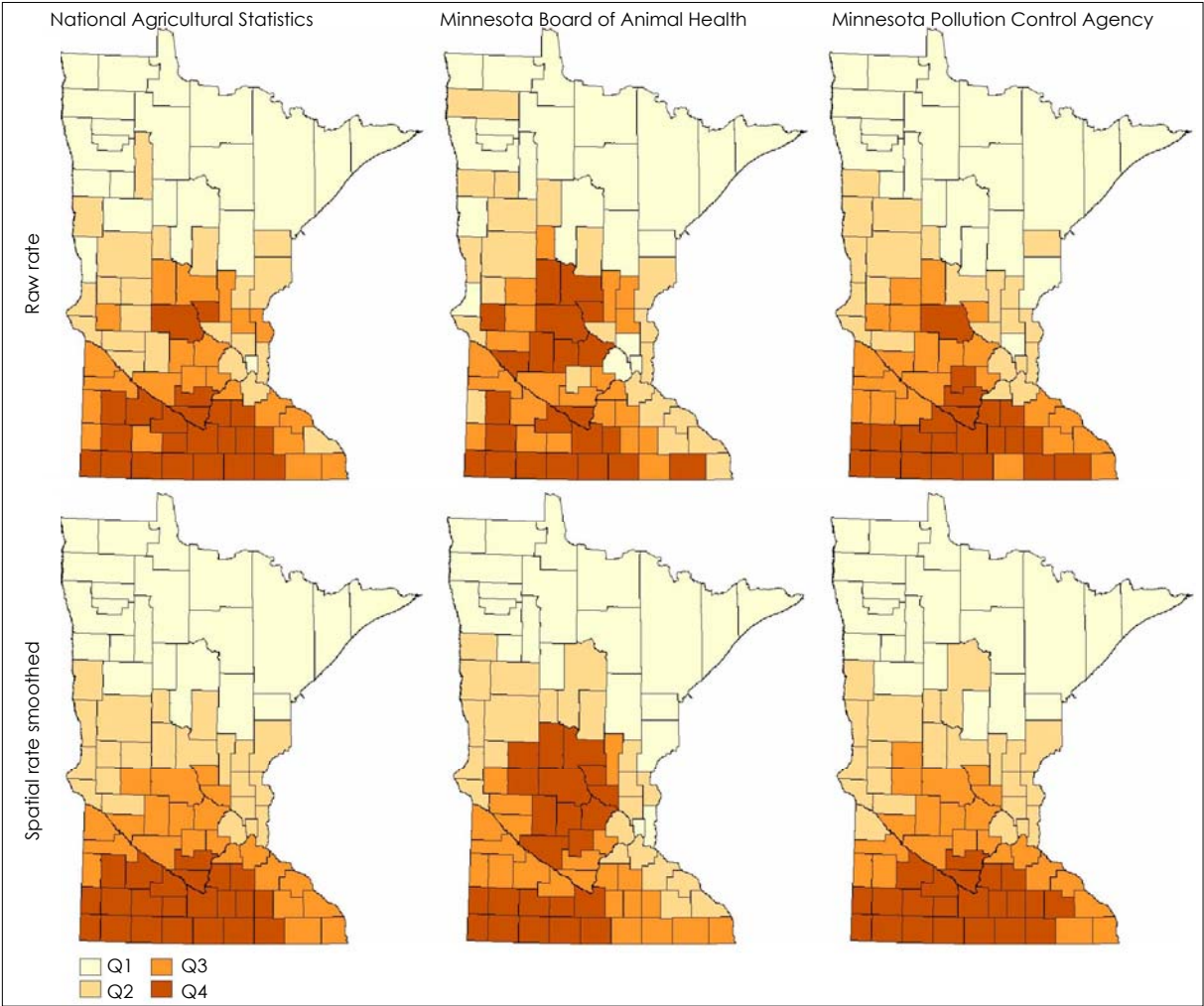


Figure 1
Swine farm density of three datasets displayed as quartiles of raw rates (number/area) and spatial rate smoothed rates (number/area)

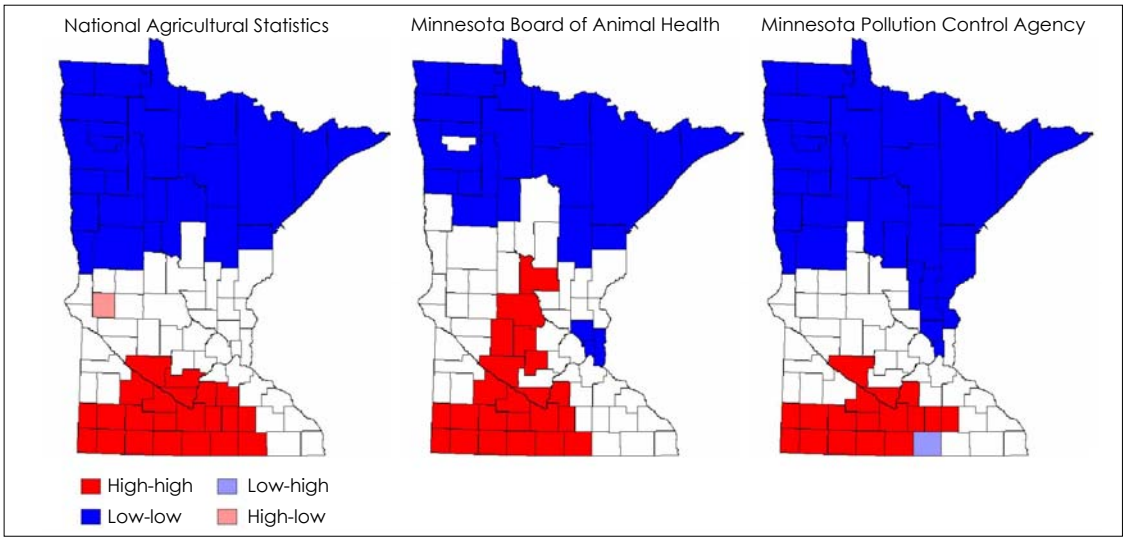


Figure 2
Cluster maps (univariate local indicator of spatial autocorrelation with empirical Bayes rate modification) of swine farm density (number/area) for three datasets
Red and blue areas indicate patterns similar to those seen in the descriptive maps

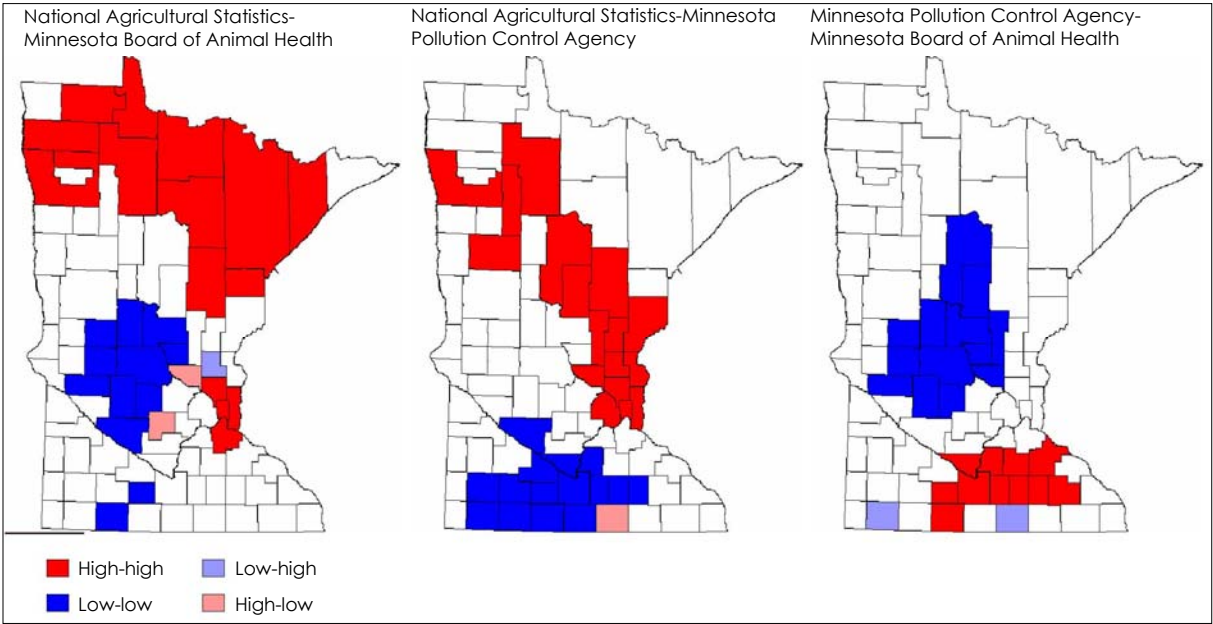


Figure 3
Clustering of differences in swine farm counts between counties standardised by area (bivariate local indicator of spatial autocorrelation with empirical Bayes rate)

database (Tables III and IV), with the probability that a dataset includes an actual, existent farm site being fairly consistent between counties and datasets. In contrast, the probability that a farm listed in either the MBAH or the MPCA dataset was actually an existent swine operation was highly variable (36% and 74%) among databases and counties with no evident pattern. These observations demonstrate considerable inaccuracy in both available databases of swine farm location with respect to the presence of operational swine farms. Although the NASS data did not include farm identities, the NASS database recorded 69 farms in Stevens County where 81 current operations were identified.

Regarding the positional accuracy of farm coordinates, overall the MBAH data was less accurate in locating farms (greater 1 standard deviational ellipse) than the MPCA data (Figs 4 and 5). These two figures show all matched farm locations plotted in reference to their true farm location. The direction and distance from the ‘true farm location’ represents the error for every farm that was matched and commonly indexed to its true farm location, assuming no bias in those farms that were successfully matched. In Rice County, the MBAH data included two points that deviated similarly in magnitude (approximately 9.8 km) and direction from the verified sites and are probably attributable to

Table III
Stevens County data accuracy for two datasets of farm locations matched to a verified farm location

Farms	Status	Verified farms		Probability estimate 1 ^(a)	Probability estimate 2 ^(b)
		Present	Absent		
MBAH farms	Present	68	96	84%	41%
	Absent	13	0		
MPCA farms	Present	60	21	74%	74%
	Absent	21	0		

(a) probability that a dataset includes an actual, existent farm site
(b) probability that a farm listed in either the MBAH or the MPCA dataset was actually a current swine operation
MBAH Minnesota Board of Animal Health
MPCA Minnesota Pollution Control Agency

Table IV
Rice County data accuracy for two datasets of farm locations matched to a verified farm location

Farms	Status	Verified farms		Probability estimate 1 ^(a)	Probability estimate 2 ^(b)
		Present	Absent		
MBAH farms	Present	45	20	85%	69%
	Absent	8	0		
MPCA farms	Present	40	72	75%	36%
	Absent	13	0		

(a) probability that a dataset includes an actual, existent farm site
(b) probability that a farm listed in either the MBAH or the MPCA dataset was actually a current swine operation
MBAH Minnesota Board of Animal Health
MPCA Minnesota Pollution Control Agency

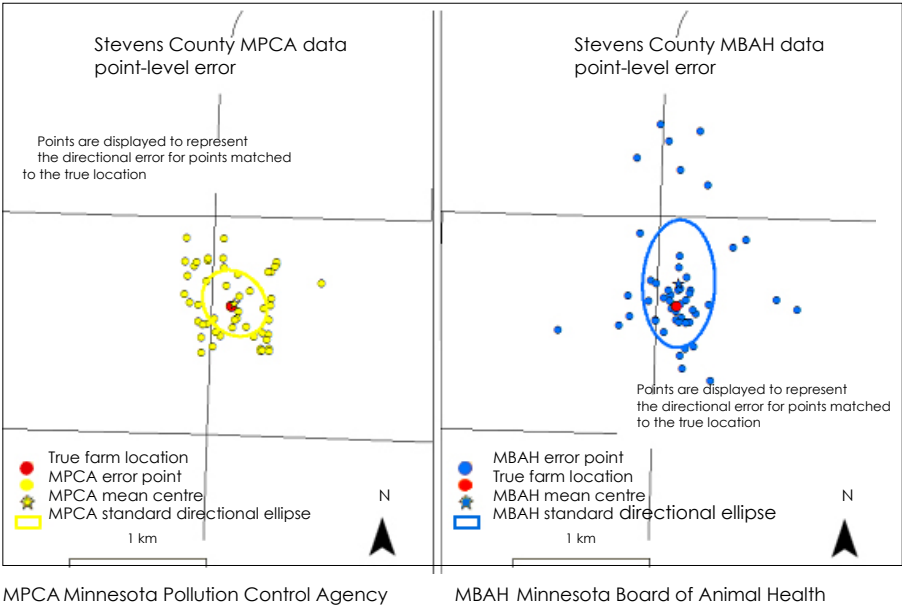


Figure 4
Point-level errors in Stevens County

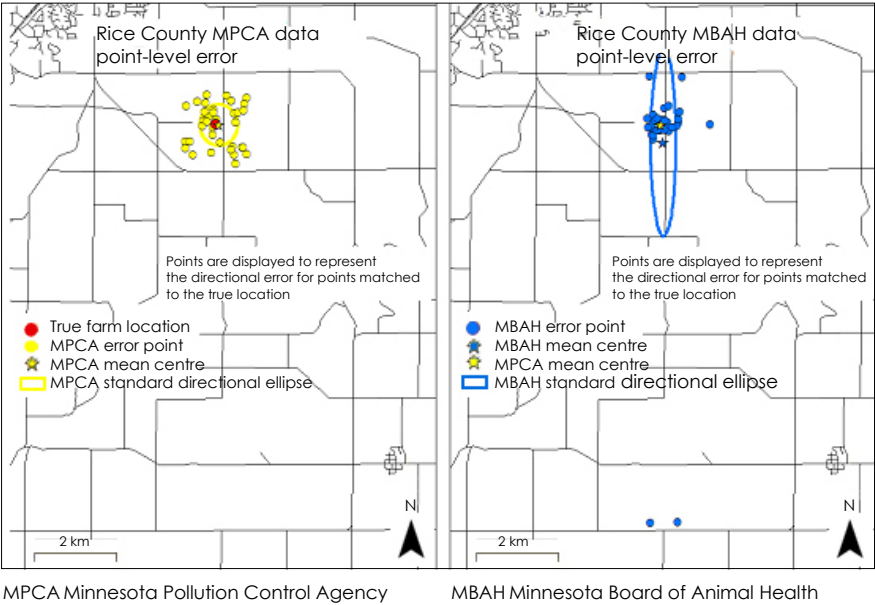


Figure 5
Point-level errors in eastern Rice County

errors in data entry. Removal of these two points would greatly reduce the average error in MBAH dataset and reduce the obvious N-S axis of the standard deviational ellipse. Histograms of the error distributions show the MBAH data to be right skewed in both counties (Fig. 6). The MPCA data appeared to have a more uniform distribution. Nearest

neighbour analysis indicated that MBAH points were clustered around the verified sites and MPCA points distributed randomly (Table V).

Table V
Nearest neighbour analysis

County	Dataset	Cluster Z-score
Stevens	MBAH	-3.65 clustered
	MPCA	-0.17 random
Rice	MBAH	-5.86 clustered
	MPCA	1.00 random

MBAH Minnesota Board of Animal Health
MPCA Minnesota Pollution Control Agency

Spatial distribution of a non-commercial swine population (4-H pigs) in relation to commercial swine production

Participation in 4-H programmes in Minnesota varies among counties. The six-year mean participation per county ranged from 0 students to 112 students enrolled per year (mean of 30 participants per county per year). Analysis of year-to-year participation numbers by county revealed very little variation. Much of the variation in county-level participation can be attributed to the previous year's participation, where correlation coefficients ranged from 0.83 to 0.98 (Table VI). We used the six-year mean number of 4-H swine participants per county to visualise the spatial distribution of the 4-H swine population in the state (Fig. 7). The four quartiles of density presented on the map show an increase in density for south-eastern counties. The

Table VI
Annual Spearman rank correlation for 4-H swine participation by county (p-values <0.0001)

Year	Spearman rank correlation
	R
2000-2001	0.83
2001-2002	0.93
2002-2003	0.97
2003-2004	0.97
2004-2005	0.97

NASS National Agricultural Statistics
MBAH Minnesota Board of Animal Health

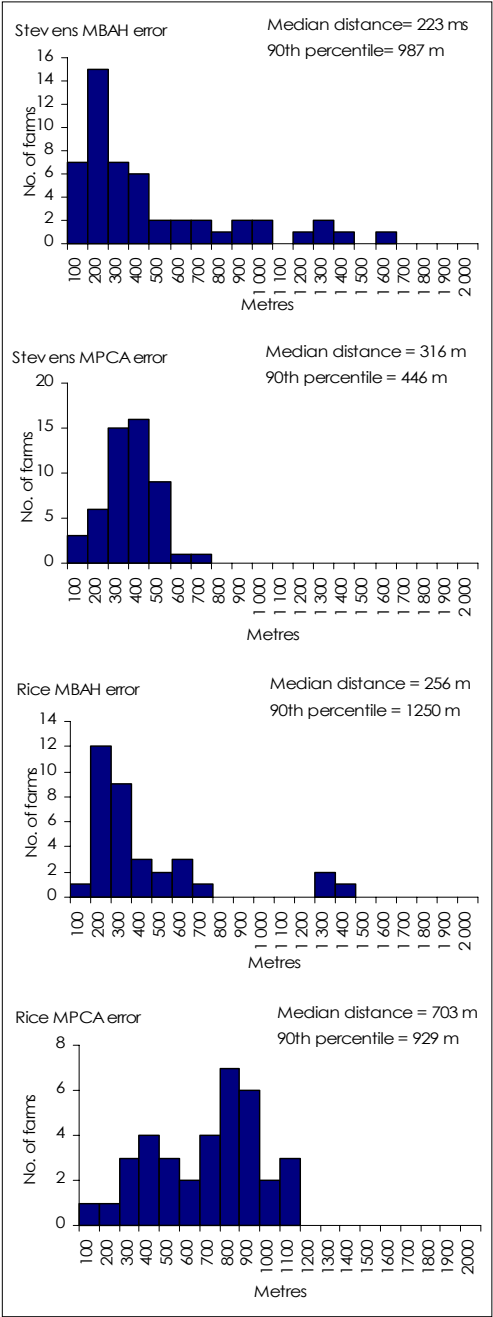


Figure 6
Point-level error histograms for two datasets in Stevens County and Rice County
Distance from presumed location to verified true location

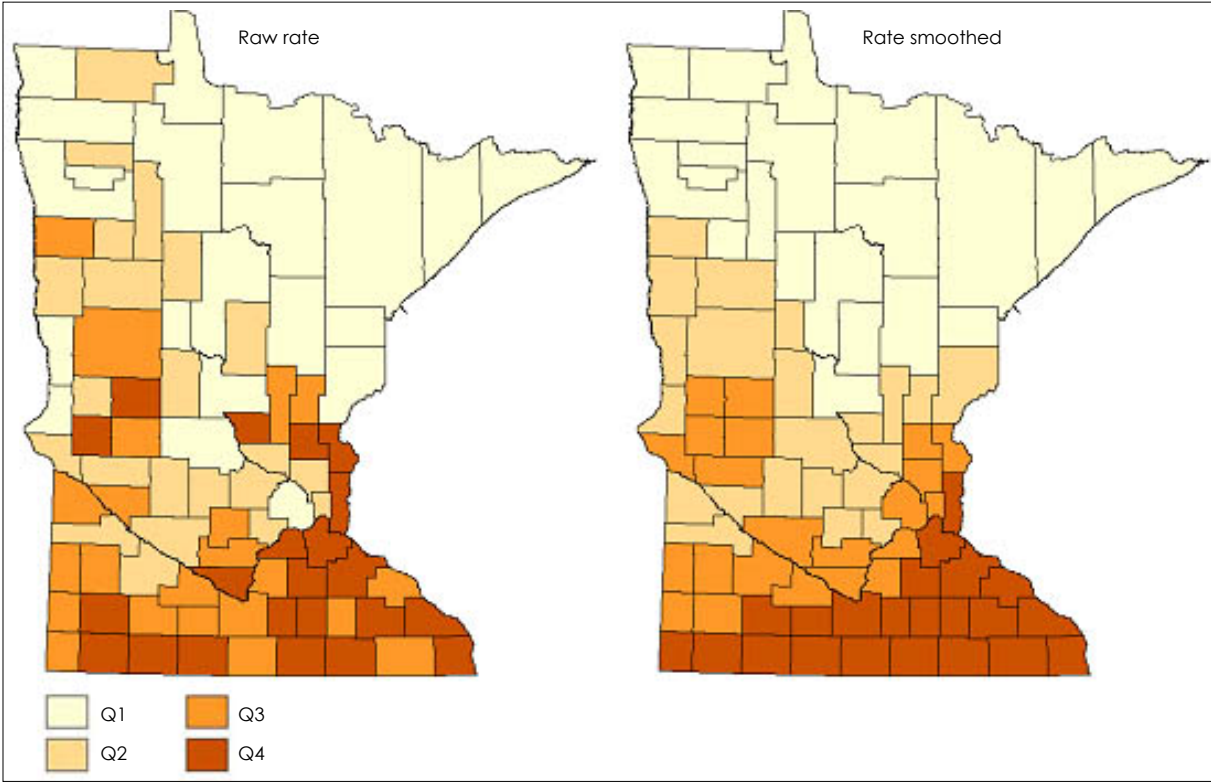


Figure 7
Six-year mean 4-H participation displayed as quartiles of raw rates (number/area) and spatial rate smoothed rates (number/area)

smoothed county-level farm density map suggests regional concentration in the south-eastern corner of the state. The spatial clustering of 4-H density seen with univariate LISA with EB rate modification further supports the observation that there is a general trend for increased 4-H participation density in the south-eastern counties (Fig. 8). Clustering was significant but the spatial autocorrelation was less than seen with commercial production (Morans $I = 0.48$ versus 0.79 for NASS data). Using the bivariate LISA with the EB rate method, it was observed that high density 4-H counties co-cluster geographically with high-density commercial production counties along the south-central portion of the state (Fig. 9). While there was statistically significant correlation between the datasets of 4-H production, the actual degree of correlation was only moderate ($R \sim 0.6$) for all comparisons (Table VII).

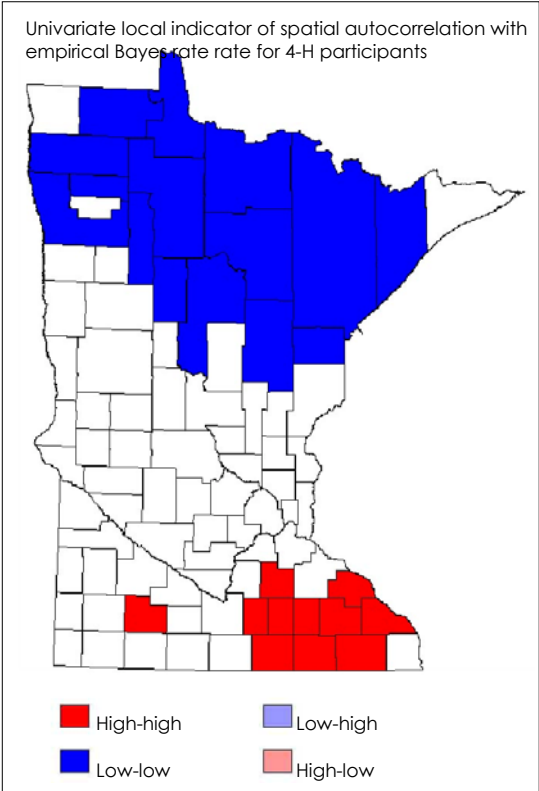


Figure 8
Spatial clustering of 4-H participation rate

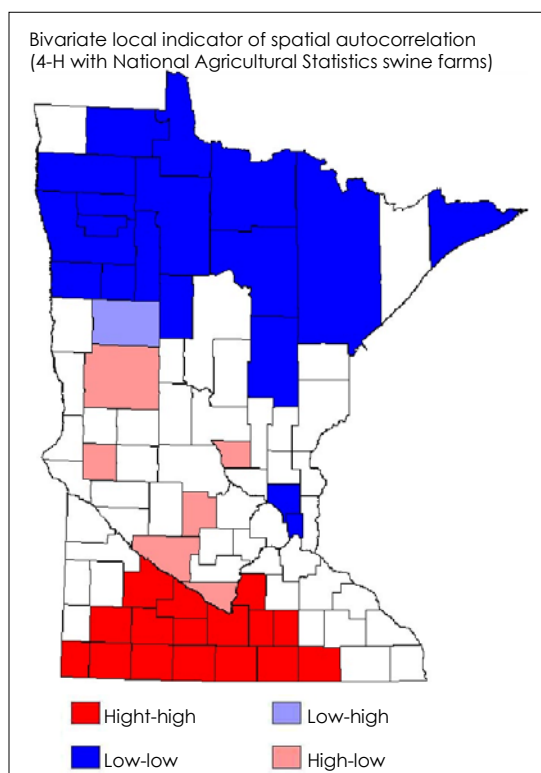


Figure 9
Clustering of 4-H participation with commercial
production
National Agricultural Statistics swine farms

Table VII
Spearman rank correlation of 4-H participants
versus commercial production
(p -values <0.0001)

Commercial production	Spearman rank correlation R
NASS farms	0.663
MBAH farms	0.595
MPCA farms	0.673

NASS National Agricultural Statistics
MBAH Minnesota Board of Animal Health
MPCA Minnesota Pollution Control Agency

The response rate for the 4-H survey was 61% (121/200). Thirty-six percent of respondents kept 4-H pigs at the site throughout the year and 39% indicated that their family also raised pigs commercially. Of these, 83% responded that the commercial pigs were within 0.5 miles (approximately 0.8 km) of their 4-H pigs and 32% reared their 4-H pigs in the same barn as commercial pigs. For respondents whose families did not raise commercial pigs, 20% of respondents indicated that the nearest commercial pigs were located within 0.5 miles.

Discussion

Accurate data on the spatial distribution of susceptible populations is an invaluable resource for planning and executing emergency responses or other programmes to control or eradicate animal diseases. There are logistic difficulties in maintaining high quality spatial databases, particularly for rapidly evolving industries, such as the swine industry in the United States. Until recently, the United States has not attempted to establish a national system for registration of premises raising animals. This responsibility has historically rested with the individual states. Current efforts led by the United States Department of Agriculture to establish a national premises database as part of the National Animal Identifications System aim to address this shortcoming, but completion of this task is unlikely in the short term. The accuracy of existing spatial farm databases is likely to vary considerably among states and our observations in Minnesota cannot be extrapolated beyond the state.

For an infectious disease in a geographic area, the relevant populations at risk include all susceptible animal species. This study was limited to swine because the predominant swine health problems (notably PRRS and porcine circovirus associated disease) affecting the United States industry appear to be highly host-specific. For logistic reasons, this study was restricted to commercial swine enterprises and a single non-commercial sector (4-H pigs) for which relevant data were accessible. This dataset demonstrated not only that a niche swine population (4-H pigs) was geographically correlated with commercial production, but also that this niche swine population was relatively stable over sequential years. Further, the 4-H survey findings demonstrate the potential interaction between commercial and non-commercial swine populations. In this, they emphasise the likely importance of non-commercial swine populations as reservoirs of swine pathogens and the need to assess the risk that such populations present when planning regional control programmes. Clearly, when

investigating disease control options, all commercial and non-commercial swine populations (e.g. feral, backyard, pet, etc.) in a region need to be included.

It should be noted that the sources of data that we evaluated were compiled for very different purposes. Adequacy of data is therefore a relative concept and standards for data quality according to use are arguably more appropriate than absolute measures of accuracy. The clustering of error points in the MBAH dataset (Figs 4 and 5), suggests they are 'seeking' to be accurate to the true location. The MPCA error point-pattern, in contrast, does not intensify at its cluster centre; the points exist randomly in the study area (Figs 4 and 5). This is probably an artefact of the respective data collection procedures. For MBAH data, farm locations are assigned to locations deemed to be the actual farm location. This fits with the data element coding laid out for Animal and Plant Health Inspection Service (APHIS) surveillance data standards on the APHIS website (www.aphis.usda.gov/vs/nahss/resources.htm). However, these definitions of data elements delineate methods of geographic data collection (geographic positioning system [GPS], address geo-coding, etc.), but no standards for precision or accuracy in reference to actual livestock locations are presented. In contrast, MPCA farm locations are assigned to the centroids of the quarter-section (square administrative unit equal to approximately 65 ha) or full section, depending on county. Since actual farm sites may be anywhere within a quarter-section, the relationship between the actual location and the MPCA assigned location appears random within the study area. Furthermore, geographic correlation and autocorrelation of county-level farm density among datasets indicate that differences among datasets are not spatially random and that one data resource may under- or overestimate the density of swine farms in a multi-county region of the state.

Conclusions

The databases that we obtained for this project were not accompanied by metadata files that would normally convey some data quality parameters and document procedures used to compile the data. Appropriate metadata should provide timely information to facilitate data sharing across organisations that could be critical in managing emergency responses. Therefore, future consideration of data quality standards for spatial databases to support animal health activities should include metadata standards.

Our study primarily considered completeness and positional accuracy. In comparing the three databases of swine production across Minnesota, we were limited by the absence of a true 'gold standard'. The unique purposes of each dataset and their differing definitions of what constitutes a 'swine farm' will, no doubt, be cause for differences. However, when used for the purposes of estimating the true locations of swine farms, it can logically be concluded that there are significant accuracy concerns in at least two of the three databases (only one can be accurate). Furthermore, the existence of spatial correlation in the patterns of discrepancies among these databases suggests local administrative factors, such as local office compensation for enrolment, or regional shifts in livestock production over time, may influence the accuracy of data. Substantial variability in the adequacy of available data can be expected even among counties within the state. This was confirmed in the analysis of verified 'gold standard' locations in Stevens and Rice counties where the probability that a farm listed in either the MBAH or the MPCA dataset was actually a current swine operation (and to a lesser extent the probability that a dataset includes an actual, existent farm site) calculated for the MPCA and MBAH databases varied substantially from county to county. We consider that the completeness of these databases observed in these two counties is substantially below what should be achievable and necessary to support effective emergency responses or other animal health programmes.

Similarly, with respect to positional accuracy of individual farms, deviations of the order of 10 km observed for a small number of farms would be problematic when managing emergency responses. The positional inaccuracy observed in the MPCA data could largely be attributed to the system for recording locations (by section rather than specific location), and could be considered acceptable for the purposes of their responsibilities (i.e. the management of livestock waste and air quality). However, the greater positional inaccuracy of the MBAH locations is of greater concern as locations are intended to indicate specific farm locations that would be used in cases of emergency response to foreign disease outbreaks. In the two counties considered, 10% of farms had a positional error of more than 900 m, a radius which would typically include other multiple farm sites. If resources were to be allocated to improve the accuracy of this database, priority should first be given to improving the completeness of the data more than the

positional accuracy of recorded farms. The current inadequacy of reliable spatial data of commercial swine enterprises in Minnesota is compounded by the presence of substantial populations of swine in non-commercial arenas. Planning for emergency response preparedness should include formal assessment of non-commercial populations and their potential role in disease transmission within a region.

Acknowledgments

The authors are grateful to staff of the Minnesota Board of Animal Health, Minnesota Pollution Control Agency, and Minnesota 4-H office for their cooperation in providing the respective datasets.

Grant support

This study was funded by a grant from the Minnesota Pork Board.

References

1. Animal and Plant Health Inspection Service (APHIS) 2002. Emergency response: foot-and-mouth disease and other foreign animal diseases. APHIS Factsheet, January. United States Department of Agriculture APHIS, Veterinary Services, Washington, DC, 3 pp (www.aphis.usda.gov/lpa/pubs/fsheet_faq_notice/fs_ahfmdres.pdf accessed on 20 July 2007).
2. Animal and Plant Health Inspection Service (APHIS) 2005. National veterinary accreditation program. Reference guide. United States Department of Agriculture APHIS, Veterinary Services, Washington, DC, APHIS Bulletin 91-55-082, 200 pp.
3. Anon. 2006. History of 4-H (www.4hcentennial.org/ accessed on 30 July 2007).
4. American Association of Swine Veterinarians (AASV) 2005. AASV establishes position statement on PRRS eradication (www.aasv.org/news/story.php?id=1610 accessed on 20 July 2007).
5. Minnesota Pork Board 2006. PRRS Eradication Task Force. Task force long-term goal is PRRS eradication from swine herds. Minnesota Pork Board, North Mankato (www.mnpork.com/producer/prrs.php accessed on 20 July 2007).
6. Sanson R. 1993. The development of a decision support system for an animal disease emergency. PhD thesis. Massey University, Palmerston North, 1-290.
7. Williams J.L. & Sheesley D. 2000. Response to bio-terrorism directed against animals. *Ann N Y Acad Sci.*, **916**, 117-120.