Spatial distribution of sow sites enrolled in the Swine Health Monitoring Program (SHMP)

Introduction

The swine health-monitoring program (SHMP) includes sow sites from about 15 States. Veterinarians currently share diagnostic results for PRRSv and / or PEDv. The objective of this first report is to show spatial distribution of participating sow sites by pathogen.

Materials and methods

A codified dataset constituted by geographical locations of sow sites that are enrolled in the SHMP to share PRRS, PED or PRRS/PED diagnostic information, was used to assess spatial distribution of sites. Using the software R,[1] and the packages ggplot2,[2] maps,[3] and MASS,[4] a spatial distribution of sow sites was assessed by conducting a 2d kernel density estimation (kde2d).[4] Briefly, kde2d evaluates the occurrence intensity of geographical points, i.e. sites, in a given unit of space by approximating its probability density function. This function represents the distribution of site location (latitude and longitude), and its influence distance (bandwidth) in the space.[4-6] A kde2d was run for sites that share PRRS, PED and all sites enrolled in the SHMP.

Results

1. **Sites that share PRRS-status**

   As of 3/1/2015, there were 405 sites that share their geographical locations and PRRS-status. Site density distribution (>0.02) indicates aggregation of sow sites located in the Midwest (Fig 1).

2. **Sites that share PED-status**

   As of 3/1/2015, there were 714 sites that share their geographical locations and PED status. Site density distribution (>0.02) indicates aggregation of sow sites located in the Southeast (Fig 2).

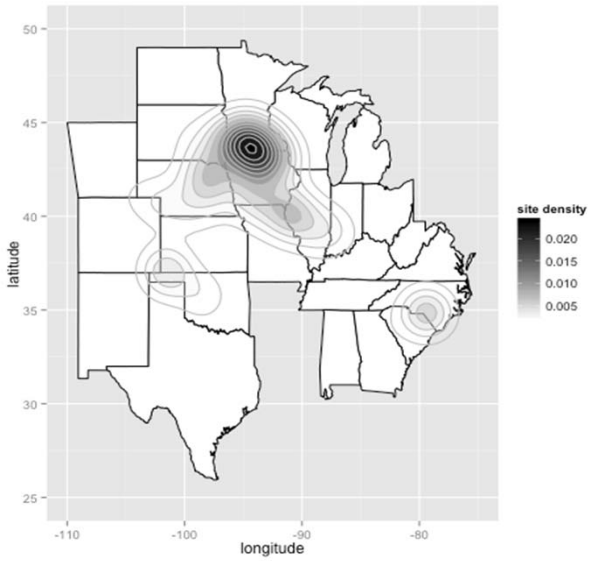

Figure 1: Kernel density estimation for sites enrolled in the SHMP that share PRRS diagnostic information

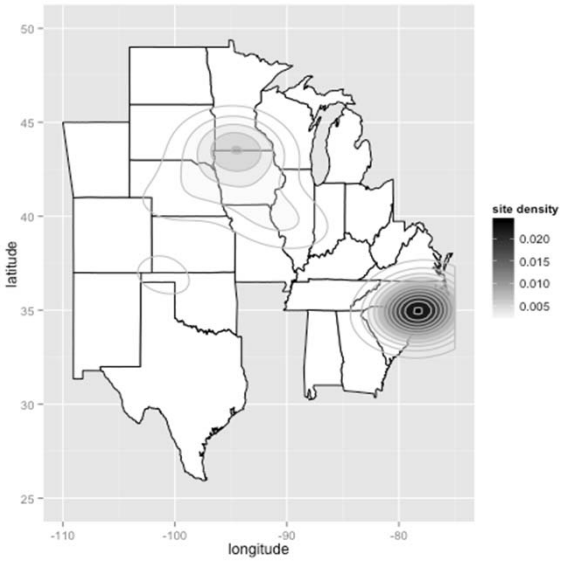

Figure 2: Kernel density estimation for sites enrolled in the SHMP that share PED diagnostic information

CONTINUED ON NEXT PAGE

3. All sites in SHMP

As of 3/1/2015, there were 774 sites that share their geographical location and PRRS, PED, or PRRS/PED status. The density distribution of site location reflects the higher number of participating sites in the Southeast (Fig 3).

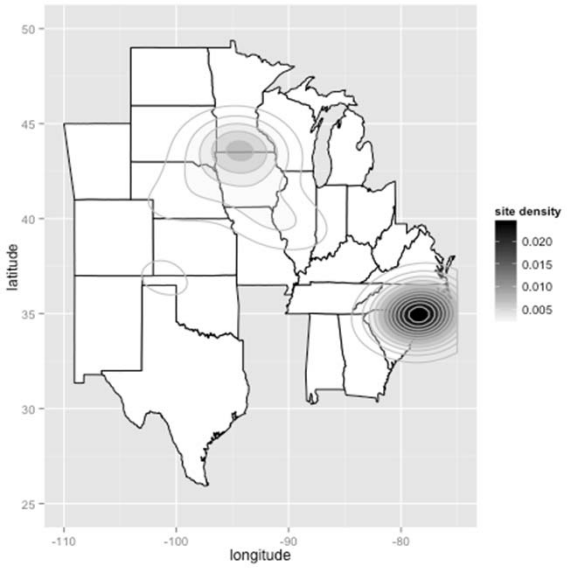

Figure 3: Kernel density estimation for all sites enrolled in the SHMP

Bibliography

1. Anon. R: A language and environment for statistical computing. In: Computing. RFfS, ed. Vienna, Austria., 2013.
2. Wickham H. *ggplot2: elegant graphics for data analysis*. New York: Springer, 2009.
3. Becker RA, Wilks AR. *maps: Draw Geographical Maps*, 2014.
4. Venables WN, Ripley BD. *Modern Applied Statistics with S*. Fourth ed. New York: Springer, 2002.
5. Silverman BW. *Density estimation for statistics and data analysis*: CRC press, 1986.
6. Duong T. ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *Journal of Statistical Software* 2007;21:1-16.

Comments

Credit for this effort goes to Pablo Valdes-Donoso, a PhD graduate student working with Andres Perez. Pablo shows us a mapping technique where we can visualize our data while keeping individual farm locations and health status confidential. Our project is constantly improving as new systems enroll and we use new techniques to analyze and present data. Today, we have 910 farms participating and several more systems putting their records together as they prepare to join. While we don't need more farms to represent incidence or prevalence for our participants, our current enrollment is biased towards certain regions as Pablo describes. That means we have to be cautious about extrapolating our observations to the industry as a whole – as discussed in last week's report. In future reports, Pablo and Andres will share maps that describe PRRS and PED prevalence.
Bob Morrison (BobM@UMN.Edu)
