1 Abstract

2 Introduction

3 Methods

4 Data

To analyze how forage availability affected pig movement patterns, we used GPS collar data collected on 500 pigs in the United States of America and Canada (Fig. 1). Of the X pigs, X were male and Y were female. These data are from X different studies and were collected from May, 2004 to November, 2017. The median collaring time per pig across all studies was X (95% quantiles) (see supplementary material for additional details). The average time between GPS fixes varied by study, with the median fix time being X across all Y studies [95% CI].

For each pig used in the study, we cleaned the movement trajectory using the following criteria. First, we excluded all 2D GPS fixes from the analysis (Bjorneraas *et al.* 2010). Second, we eliminated the first and last 25 fixes for each pig to account for capture effects [citation]. Third, we eliminated all pig movements in which pigs moved faster than 50 km per hour over any distance or moved greater than 15 km per hour over a distance of 8 km or more. These movements are unlikely given previously observed patterns of feral swine movement (Mayer & Brisbin 2009). Finally, we also removed large temporal gaps between fix times as these can lead to a large amount of uncertainty regarding a pig's location in between two temporally distant fixes. To do this, we split a single pig movement trajectory into multiple sequences when the time between two adjacent fixes was greater than *c.* 130 minutes. We then discarded all sequences that had less than 150 fixes at least 130 minutes apart. The remaining sequences were considered independent "runs" for a given pig. We chose to only include sequences of 150 fixes as shorter runs would not always span the length of a day, precluding analysis of diel movement patterns, and longer runs excluded too many of the pigs from the analysis. After this cleaning, our analysis contained X pigs and a total of X fix times (Fig. 1).

26 Covariates

The goal of our analysis was to understand how the availability of anthropogenic and natural forage on a landscape affected the movement and resource selection of feral swine. To address these goals, we identified covariates related to natural forage and anthropogenic forage resources that

1

were comparable across studies. For the availability of natural forage resources, we considered two proxies: plant productivity as measured by Normalized Difference Vegetation Index (NDVI) and density of masting trees (Table 1). We included masting tree density as masting events are an important component of pig diets and can have significant implications on population-level pig growth rates (e.g. Bieber & Ruf 2005). NDVI is a general correlate of plant productivity in an area (Pettorelli *et al.* 2005) and, given that pigs are extreme generalists and plant material often makes up a large percentage of their diet (Mayer & Brisbin 2009), here we use it to reflect the availability of herbaceous forage.

For anthropogenic forage resources, we used agricultural data available from the National Agricultural Statistics Service (NASS). These data are obtained yearly at a 30m by 30m scale across the contiguous US and each pixel specifies the dominant type of crop grown in that pixel. While the NASS data enumerates 105 types of crop, we chose not to distinguish between crop types. Instead, we used a generic grouping of "crop" vs. "no-crop", which allowed us to more easily explore how the effect of agricultural forage on pig movement varied across space and time. We did consider crop-types in our post-hoc analyses, in which we grouped the 105 crop types into 11 groups: cereals, oilseed, tobacco, beverage and spice, leguminous, grasses, sugar, root and tuber, fruit and nuts, vegetables and melons, and other crops [according the criteria outlined in...]. WHAT ABOUT CANADA?]

While our primary goal of this analysis was to understand the effect of forage availability on pig movement, movement is also driven by a number of other variables, including cover, water availability, temperature, pressure, human development, mammal diversity, sex, and age (McClure *et al.* 2015; Garza *et al.* 2017; Kay *et al.* 2017) [MORE]. Of these additional predictors of pig movement, cover and water are critically important as both are necessary for pigs to thermoregulate (Choquenot *et al.* 1996) and cover, in particular, is necessary for protection from predators [citation] (citations in Mayer2009). We used tree canopy density data [check wording] from the National Landcover Database (NLCD) as a covariate for cover (Table 1). These data are available at a 30m by 30m scale for the contiguous United States. For a water, we used the National Wetland Inventory (NWI) Database which identifies X types of water bodies larger than X m2 in the contiguous United States. For this study, we only considered the water sources that were permanent or semi-permanent as defined by the NWI and computed the distance to the nearest water source at a 30m by 30m resolution (see Table 1 and Sup Mat).

We also included mean monthly temperature, total monthly precipitation, and distance to developed land/roads as additional covariates (Table 1). Note that while the inclusion of these co-

variates is import for capturing pig movement, our questions were focused the foraging resources, such that we considered non-foraging covariates "blocking" covariates and did not exhaustively explore their potential relationships and interactions beyond what had already been shown in the literature.

Finally, we also considered a number of study-level variables to allow us to explore how and why the effect of natural forage resources and anthropogenic forage resources varied across different studies. To this end, we explored study-level covariates such as drought severity, ecoregion, local pig density, and MORE (see Table 1).

Continuous-time, discrete-space movement model

To understand role of forage resources on pig movement, we used the modeling framework of Hanks *et al.* (2015) and Wilson *et al.* (2018). Generally, this framework leverages auto-correlated animal movement data and gridded raster covariates to make inference about the resource utilization of an animal (Hanks *et al.* 2015; Buderman *et al.* 2018; Wilson *et al.* 2018).

Specifically, this approach can be broken into two distinct steps. Given a trajectory of GPS fixes (not necessarily with equal fix times), the first step of this approach estimates animal movement as a function of continuous time at some particular temporal grain (e.g. 15 minutes). To do this, we used a phenomenological functional movement model (FMM) (Buderman *et al.* 2016; Hooten *et al.* 2017), which is a non-mechanistic, continuous-time movement model that can capture an animal's movement patterns at some desired-level of detail (Buderman *et al.* 2016). The phenomenological FMM can be represented as a series of basis functions, which allow for large flexibility in animal movement patterns. In particular, we used a B-spline basis expansion to model the longitude and latitude of an animal as a function of time (see Supplementary Material for additional detail).

After fitting the phenomenological FMM to each pig trajectory, we used this model to predict a pig's location at 15 minute intervals. We chose 15 minutes as this provided a reasonable trade-off between computational time and computing the exact amount of time spent in each cell, which is equivalent to letting the time between fix times go to 0 [NOTE: Would this be hard to do? Double Check interpretation of 15 minute intervals...]. We repeated this 20 times to account for the uncertainty in the movement path (Hanks *et al.* 2015; Buderman *et al.* 2018). All the analyses described next were repeated on each of the 20 imputed data sets to account for uncertainty in the movement trajectory.

Given the FMM-predicted trajectories, we then explored how agricultural and natural forage

resources on a landscape affected pig movement using a continuous-time Markov Chain (CTMC) approach Hanks *et al.* (2015). The CTMC approach considers continuous-time animal movement (i.e. the FMM model described above) through a discrete, rasterized landscape. From this point-of-view, animal movement can be considered as a series of rates of moving from cell i to an adjacent cell j, λ_{ij}. As for any continuous-time Markov Chain, the process can be decomposed into the waiting time before a state change occurs (i.e. the time an animal spends in a cell) and the new state once a change occurs (i.e. the new cell to which the animal has moved) (Allen 2003). With this interpretation, one can then model the rate of moving between cell i and j λ_{ij} as a function of the environmental covariates in cell i and j. Hanks *et al.* (2015) showed that this type of inference can be re-expressed as a latent-variable, Poisson Generalized Linear Model, where the response variable for adjacent cell j is one if a pig moved to that cell from cell i, and 0 otherwise. Specifically, let z_{ij} be the zero/one latent variable, then

$$z_{ij} \sim \text{Poisson}(\lambda_{ij}) \tag{1}$$

$$\log \lambda_{ij} = \log \tau_{ij} + \beta \mathbf{X} \tag{2}$$

where τ_{ij} is the waiting time before moving from cell i to cell j, $\mathbf{X}$ is a vector of landscape covariates, and β is the effect of these covariates on movement.

Considering $\mathbf{X}$, we explored two classes of covariates: location-based drivers and directional drivers of movement (Hanks *et al.* 2015). Location-based drivers are a result of the cell that an animal is currently in and affect how long an animal remains in the current cell. For example, if masting tree density was a negative location-based driver of pig movement then a pig in a cell with high masting tree density would tend to remain in that cell longer than a cell with lower masting tree density. Directional drivers of movement determine the direction that a pig might move once it leaves the cell it is currently occupying. For example, if masting tree density was a positive directional driver of pig movement then, upon leaving the currently occupied cell, a pig tends to move in the direction of increasing masting tree density, relative to its current position [TODO: THIS IS A LOCAL GRADIENT EFFECT...SHOULD WE ALSO EXPLORE GLOBAL EFFECTS AT SOME POINT? IN OTHER WORDS, does the presence of masting trees some where in the vicinity affect how a pig uses crop?]. Table 1 shows which of the covariates described in the previous section we considered as location-based drivers, directional drivers, or both.

Model specification and fitting

We analyzed three variations of the CTMC model with the goal of understanding the role of foraging resources on the movement of feral swine across populations. The each model is described below

Model 1: Main effects model

The first, and simplest, model that we analyzed only considered time-invariant main effects of anthropogenic and natural foraging resources and "blocking" factors. The model is defined as

$$\log(\lambda_{ij}) = \log(\tau_{ij}) + \beta_0 + \beta_1 \text{directional persistence} + \beta_2 \text{male} \tag{3}$$

$$+ \beta_3 \text{canopy cover}_{\text{loc}} + \beta_4 \text{canopy cover}_{\text{grad}} \tag{4}$$

$$+ \beta_5 \text{distance to water}_{\text{loc}} + \beta_6 \text{distance to water}_{\text{grad}} \tag{5}$$

$$+ \beta_7 \text{distance to development}_{\text{loc}} + \beta_8 \text{distance to development}_{\text{grad}} \tag{6}$$

$$+ \beta_9 \text{distance to crops}_{\text{loc}} + \beta_{10} \text{distance to crops}_{\text{grad}} \tag{7}$$

$$+ \beta_{11} \text{masting tree density}_{\text{loc}} + \beta_{12} \text{masting tree density}_{\text{grad}} \tag{8}$$

$$+ \beta_{13} \text{NDVI}_{\text{loc}} + \beta_{14} \text{NDVI}_{\text{grad}} \tag{9}$$

[TODO: add in random effects if necessary and more description]. We fit this model in two steps. First, we performed LASSO regularization with X-fold cross-validation using the `glmnet` package to select the estimate the best-fit, regularized model [MORE]. However, as this LASSO approach does not provide easy access to parameter uncertainty or random effects, we then re-fit the best-fit regularized model using automatic differentiation variational inference (ADVI), which allowed us to approximate the uncertainty on the model parameters and include random effects of individual pigs on movement. Due to the size of the datasets under-consideration in this study (i.e. millions of datapoints) ADVI allowed us to perform these analyses in a reasonable amount of time (i.e. 2- 4 hours) compared to a full-scale Bayesian analysis.

[Describe additional models here]

Results

Discussion

References

Allen, L.J. (2003) *An Introduction to Stochastic Processes with Applications to Biology.* Pearson Education, Inc, Upper Saddle River, New Jersey.

Bieber, C. & Ruf, T. (2005) Population dynamics in wild boar Sus scrofa: Ecology, elasticity of growth rate and implications for the management of pulsed resource consumers. *Journal of Applied Ecology*, **42**, 1203–1213.

Bjorneraas, K., Moorter, B.V., Rolandsen, C.M. & Herfindal, I. (2010) Screening Global Positioning System Location Data for Errors Using Animal Movement Characteristics. *Journal of Wildlife Management*, **74**, 1361–1366.

Buderman, F.E., Hooten, M.B., Alldredge, M.W., Hanks, E.M. & Ivan, J.S. (2018) Predatory behavior is primary predictor of movement in wildland-urban cougars. *bioRxiv*.

Buderman, F.E., Hooten, M.B., Ivan, J.S. & Shenk, T.M. (2016) A functional model for characterizing long-distance movement behaviour. *Methods in Ecology and Evolution*, **7**, 264–273.

Choquenot, D., McIlroy, J. & Koen, T. (1996) Managing vertebrate pests: feral pigs. *Beureau of Resource Sciences*, pp. 1–171.

Garza, S.J., Tabak, M.A., Miller, R.S., Farnsworth, M.L. & Burdett, C.L. (2017) Abiotic and biotic influences on home-range size of wild pigs (Sus scrofa). *Journal of Mammalogy*, pp. 1–11.

Hanks, E.M., Hooten, M.B. & Alldredge, M.W. (2015) Continuous-time discrete-space models for animal movement. *Annals of Applied Statistics*, **9**, 145–165.

Hooten, M.B., Johnson, D.S., McClintock, B.T. & Morales, J.M. (2017) *Animal Movement: Statistical Models for Telemtry data.* CRC Press, New York, USA.

Kay, S., Fischer, J., Monaghan, A., Beasley, J., Boughton, R., Campbell, T., Cooper, S., Ditchkoff, S., Hartley, S., Kilgo, J., Wisely, S., Wyckoff, A., VerCauteren, K. & Pepin, K. (2017) Quantifying drivers of wild pig movement across multiple spatial and temporal scales. *Movement Ecology*, **5**, 1–15.

Mayer, J.J. & Brisbin, I.J.J. (2009) Wild Pigs: Biology, Damage, Control Techniques and Management. Technical report, Savannah River National Laboratory, Aiken, South Carolina.

McClure, M.L., Burdett, C.L., Farnsworth, M.L., Lutman, M.W., Theobald, D.M., Riggs, P.D., Grear, D.A. & Miller, R.S. (2015) Modeling and mapping the probability of occurrence of invasive wild pigs across the contiguous United States. *PLoS ONE*, **10**, 1–17.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J. & Stenseth, N.C. (2005) Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology and Evolution*, **20**, 503–510.

Wilson, K., Hanks, E. & Johnson, D. (2018) Estimating animal utilization densities using continuous-time Markov chain models. *Methods in Ecology and Evolution*, **In press**.

Table 1: Description of covariates used in analysis

Covariate	Description	Data Source	Spatial Resolution	Temporal Resolution	Location or gradient?
Normalized Difference Vegetation Index (NDVI)	A proxy for plant productivity and natural forage availability.	MODIS and NASS	250 m × 250 m	Monthly	Both
Density of hard-masting trees		Tabak et al.	1 km × 1 km	Time-invariant	Both
Distance to crops	The distance to the nearest crop field. A measure of anthropogenic forage.	NASS	30 m × 30 m	Yearly	Both
Distance to water	The distance the nearest permanent or semi-permanent water source	NWI	30 m × 30 m	Time-invariant	Both
Canopy density [CHECK]	A proxy for habitat cover	NLCD	30 m × 30 m	Time-invariant	Both
Distance to developed-land	A measure of human presence	NLCD	30 m × 30 m	Time-invariant	Both
Temperature	Mean monthly temperature	NOAA	50 km × 50 km	Monthly	Location
Precipitation	Total monthly precipitation	NOAA	50 km × 50 km	Monthly	Location