

# 1 Abstract

## 2 Introduction

## 3 Methods

### 4 Data

5 To analyze pig movement patterns with respect resource availability, we used GPS collar data  
6 collected on 500 pigs in the United States of America and Canada (Fig. 1). Of the 500 pigs, X  
7 were male and Y were female. These data were a result of X unique studies and were collected  
8 from May, 2004 to November, 2017 The median collaring time per pig across all studies was X  
9 (95% quantiles) (see supplementary material for additional details). The average time between  
10 GPS fixes varied by study, with the median fix time being X across all Y studies [95% CI].

11 While the movement models that we describe below account for unequal fix times, too large  
12 of a gap between fix times for any particular pig leads to a large amount of uncertainty regarding  
13 where a pig was in between those two fixes. To account for this, for a sequence of GPS fixes  
14 for a given pig, we split the single sequence into multiple sequences at when the time between  
15 to adjacent fixes was greater than  $c$ . 130 minutes. We then discarded all sequences that had  
16 less than 200 fixes. The remaining sequences were considered independent "runs" for a given  
17 pig. We chose to only include sequences of 200 or greater fixes because at a fix time of 15  
18 minutes, a sequence of 200 GPS fixes would span just over two days. This is the minimum  
19 time span needed for inference on diel patterns of movement (i.e. two daily cycles). Finally,  
20 we cleaned trajectories for errant fixes using the non-movement criteria described in Bjørneraas  
21 *et al.* (2010). [DESCRIBE THIS MORE] Finally, we cleaned trajectories for errant fixes using  
22 the non-movement criteria described in Bjørneraas *et al.* (2010). [DESCRIBE THIS MORE and  
23 move this up.] After this cleaning, our analysis contained X pigs and a total of X fix times.

### 24 Covariates

25 The goal of our analysis was to understand how the availability of agricultural forage on a  
26 landscape affected the movement and resource selection of feral swine. Moreover, we sought to  
27 understand how the availability of natural forage resources affected the selection of agricultural  
28 forage resources, and vice-versa. To address these goals, we identified covariates related to  
29 natural forage and agricultural forage resources that were comparable across studies in our data.

30 For natural forage resources, we considered two proxies for availability of natural forage: plant  
31 productivity as measured by Normalized Difference Vegetation Index (NDVI) and density of  
32 mastig trees (Mikey citation, Table 1). We included mastig tree density as mastig events are  
33 an important component of pig diets and can have significant implications on population-level  
34 pig growth rates (e.g. Bieber & Ruf 2005). The mastig layer we used was developed by Tabak  
35 et al. and describes the density of mastig trees across the world at a 1 km by 1 km scale. NDVI  
36 is a general correlate of plant productivity in an area (Pettorelli *et al.* 2005). Given that pigs are  
37 extreme generalists and plant material often makes up a large percentage of their diet (Mayer  
38 & Brisbin 2009) [Others], NDVI provides a proxy for plant resource availability. Moreover, we  
39 consider NDVI as time-varying covariate on the monthly scale, to account for the fact that pigs  
40 often shift their foraging activity to match available resources (Mayer & Brisbin 2009).

41 For agricultural forage resources, we used agricultural data available on CropScape (X), which  
42 provides yearly data on crop production (Table 1). These data are provided yearly at a 30m by  
43 30m scale across the contiguous US and each pixel specifies the dominant type of crop grown in  
44 that pixel. While CropScape enumerates 105 types of crop, we chose not to distinguish between  
45 crop types. While previous studies have shown that feral swine can preferentially select crop  
46 types (e.g. Herrero *et al.* 2006), the generic grouping of crop allowed us to more easily explore  
47 how the effect of agricultural forage on pig movement varied across space and time. When we  
48 considered crop types in our post-hoc analyses, we grouped the 105 crop types into 11 groups:  
49 cereals, oilseed, tobacco, beverage and spice, leguminous, grasses, sugar, root and tuber, fruit  
50 and nuts, vegetables and melons, and other crops. [WHAT ABOUT CANADA?]

51 While our primary goal of this analysis was to understand the effect of forage availability on  
52 pig movement, movement is also driven by a number of other variables, including cover, water  
53 availability, temperature, pressure, human development, mammal diversity, sex, and age (Mc-  
54 Clure *et al.* 2015; Garza *et al.* 2017; Kay *et al.* 2017) [MORE]. Of these additional predictors of  
55 pig movement, pig ecology and physiology suggests that cover and water are critically important  
56 as both are necessary for pigs to thermoregulate (Choquenot *et al.* 1996) and cover, in particular,  
57 is necessary for protection from predators [citation] (citations in Mayer2009). Generally, these  
58 predictions regarding the importance of cover and water have born out in other studies exploring  
59 pig movement(e.g. Kay *et al.* 2017) and thus we sought to include these covariates in our models  
60 as well.

61 For a cover covariate, we used tree canopy density data [check wording] from the National  
62 Landcover Database (NLCD) (Table 1). These data are available at a 30m by 30m scale for

the contiguous United States. For a water covariate, we used the National Wetland Inventory (NWI) Database which identifies X types of water bodies larger than X m<sup>2</sup> 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. For a given landscape on which a pig was moving, we then 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 covariates is important 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 affect of natural forage resources, agricultural forage resources, and their interactions 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 natural forage resources and agricultural forage resources on pig resource selection, 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 the continuous-time Markov Chain (CTMC) approach described in 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  $j$ ,  $\lambda_{ij}$ . 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  $j$   $\lambda_{ij}$  as a function of the environmental covariates on the landscape 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

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

$$\log \lambda_{ij} = \log \tau_{ij} + \beta \mathbf{X} \quad (2)$$

where  $\tau_{ij}$  is the waiting time before moving from cell  $i$  to cell  $j$ ,  $\mathbf{X}$  is a vector of landscape covariates, and  $\beta$  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 four variations of the model described by equation X in order to understand the the role of foraging resources on the movement of feral swine across populations. In order to minimize computational and model complexity, we performed each of the following steps when fitting the models below. First, for each population separately, we fit the full model using a regularized we fit each model described below separately for all studies and used a post-hoc [regression analysis] to compare population-level effects across studies. This was necessary as the imputed data for each model averaged 1 million data points with between 20-60 parameters being fit and regularized [fix wording and lingo].

#### *Model 1: Main effects model*

The first, and simplest, model that we analyzed only considered time-invariant main effects of agricultural and natural foraging resources and “blocking” factors. Specifically, the model was defined as

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

$$+ \beta_3 \text{canopy cover}_{\text{loc}} + \beta_4 \text{canopy cover}_{\text{grad}} \quad (4)$$

$$+ \beta_5 \text{distance to water}_{\text{loc}} + \beta_6 \text{distance to water}_{\text{grad}} \quad (5)$$

$$+ \beta_6 \text{distance to crops}_{\text{loc}} + \beta_7 \text{distance to crops}_{\text{grad}} \quad (6)$$

$$+ \beta_8 \text{masting tree density}_{\text{loc}} + \beta_9 \text{masting tree density}_{\text{grad}} \quad (7)$$

$$+ \beta_{10} \text{NDVI}_{\text{loc}} + \beta_{11} \text{NDVI}_{\text{grad}} \quad (8)$$

[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.
- Herrero, J., García-Serrano, A., Couto, S., Ortuño, V.M. & García-González, R. (2006) Diet of wild boar *Sus scrofa* L. and crop damage in an intensive agroecosystem. *European Journal of Wildlife Research*, **52**, 245–250.

- 181 Hooten, M.B., Johnson, D.S., McClintock, B.T. & Morales, J.M. (2017) *Animal Movement:*  
182 *Statistical Models for Telemetry data*. CRC Press, New York, USA.
- 183 Kay, S., Fischer, J., Monaghan, A., Beasley, J., Boughton, R., Campbell, T., Cooper, S.,  
184 Ditchkoff, S., Hartley, S., Kilgo, J., Wisely, S., Wyckoff, A., VerCauteren, K. & Pepin, K.  
185 (2017) Quantifying drivers of wild pig movement across multiple spatial and temporal scales.  
186 *Movement Ecology*, **5**, 1–15.
- 187 Mayer, J.J. & Brisbin, I.J.J. (2009) Wild Pigs: Biology, Damage, Control Techniques and Man-  
188 agement. Technical report, Savannah River National Laboratory, Aiken, South Carolina.
- 189 McClure, M.L., Burdett, C.L., Farnsworth, M.L., Lutman, M.W., Theobald, D.M., Riggs, P.D.,  
190 Grear, D.A. & Miller, R.S. (2015) Modeling and mapping the probability of occurrence of  
191 invasive wild pigs across the contiguous United States. *PLoS ONE*, **10**, 1–17.
- 192 Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J. & Stenseth, N.C. (2005) Us-  
193 ing the satellite-derived NDVI to assess ecological responses to environmental change. *Trends*  
194 *in Ecology and Evolution*, **20**, 503–510.
- 195 Wilson, K., Hanks, E. & Johnson, D. (2018) Estimating animal utilization densities using  
196 continuous-time Markov chain models. *Methods in Ecology and Evolution*, **In press**.