# Abstract

# Introduction

Pigs eat everything. But they eat some things more than others. And when those "some things" are agricultural crops, it can result in extensive economic damage. [estimate]. However, crops are one of many forage resources competing for a feral swine's attention on a landscape. As extreme generalists, feral swine consume anything from sapling trees, fungus, and masting seeds to small invertebrates and livestock. The availability of these different forage resources varies in both space and time, such that pigs will switch their foraging behavior availability of these various potential resources

# Methods

Questions: 1. How are pigs using crops and which crops are pigs using? - Summary of crop use across all 500ish pigs in the study based on the movement model. This will provide uncertainty in the amount of crop use, which will be useful (not just when they were observed in a crop field). - Summarize crop use by state, county, ecoregion, sex, etc. - Also include date in crop field (does in correspond with harvesting or when crops are likely present?) 2. How does crop use affect the movement and resource selection of pigs? - Fitting a resource selection model to understand the how patterns of movement and the use of non-crop resources might vary between pigs that use crops and don't use crops. - We hypothesize that pigs that use crops (regularly?) are utilizing a higher nutritional resource (at a higher risk), such that there general movement patterns would be different (more movement) - Second, we hypothesize that the use of crops will influence the use of non-agricultural resources (during certain points of the season), as foraging will be dominated by movement to and away from crops. 3. Are pigs using crops in different ways? - This is tough to answer because it is going to be really hard to control for season as pigs could be using crops in different ways when there are no crops on the landscape.

## GPS data

To address our questions regarding how pigs were using agricultural resources and how this affected overall patterns of resource selection, we used GPS collar data collected on 500 pigs in the United States of America (Fig. 1). These data are from X different studies and were collected from May, 2004 to November, 2017. Of the 500 pigs, X were boars and Y were sows.

Given that all of these studies were collected for different purposes, the median fix time per pig varied across all studies with a range of 15 minutes - 3 hours.

For each pig, 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 fixes in which pigs moved faster than 40 km per hour as these movements are unlikely given previously observed patterns of feral swine movement (Mayer & Brisbin 2009). After this cleaning procedure, our data consisted of X fixes. Below we describe how we accounted for unequal fix times and variable fix times across studies.

## Agricultural covariates

As the goal of our analysis was to understand both how feral swine were using crops and how the use of crops affected the resource utilization of feral swine over a continental scale, we needed an agricultural covariate that was consistent for every pigs in our study. To do this, we used the crop data available from the National Agricultural Statistics Service (NASS, available at CropScape). NASS provides raster data at the 30m by 30m scale across the contiguous US, where each pixel specifies the primary type of crop in that area (including landcover types where there are no crops, see X for more information). This data is publicly accessible and is available at the at the yearly temporal scale. The NASS crop layers identify 105 crop types. When answering our first question, how do pigs use crops?, we use all 105 of these crop types as defined by NASS. When answering our second question, Y, 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. These agricultural groups were delineated based on similarities in nutritional content and seasonal availability in a region [CHECK]. This made it more feasible to compare crop use across populations.

## Question 1: How do pigs use crops?

To answer this question, we used the GPS movement data and the NASS crop layers to determine which crops pigs were "using" on a landscape. We considered "use" of crops as any time a pig was in a pixel that contained a crop type. Note that this does not necessarily mean that pigs were foraging on or damaging this agricultural resource. However, our definition of "use" is consist with how resource use is defined in the ecological literature (Hooten *et al.* 2017) and it is not unreasonable to assume that there is a positive correlation between pigs being physically in an agricultural field and damaging the crops in that field.

While conceptually straight-forward, this definition of "use" is limited by the fact that the GPS movement data from each study in this analysis was collected on different time scales. To address this challenge, fit a continuous-time functional movement model to the GPS trajectories for each pig (Buderman *et al.* 2016). In short, this approach uses basis functions to fit a phenomenological, continuous-time movement model to a set of discrete GPS fixes (see SI for a full description Buderman *et al.* 2016). Importantly, this approach allowed us to sync-up time scales between studies and account for the uncertainty in the movement path, such that our comparisons of crop use across studies was not confounded by the the length between GPS fixes. Note that studies with larger times between fixes will tend to have larger uncertainty in the predicted crop use, as the functional movement model is less certain as to where the pigs is between to fixes farther apart in time than two fixes that are close together (Fig. X).

After fitting the data to the functional movement models, we then used the predicted continuous-time trajectories along with the NASS crop layer data to determine, with uncertainty, the identity of the crops pigs were using over the course of their measured movement trajectory and how long they were using these crops for (i.e. time spent in a pixel with a certain crop type). We did this using the `ctmcmove` package in R which allows us to discretize continuous-time movement trajectories onto discrete raster grid (Hanks *et al.* 2015).

[With this data, we compared]

## Question 2: How does crop-use affect the movement patterns and resource selection of feral swine?

While the above analysis identifies which crops pigs are using, it does not provide any information on how agricultural resources affect pig movement (e.g. are pigs actively moving towards crops and slowing when they are in these crops?) or how other resources on the landscape are affecting pig movement (e.g. do pigs that use crops use other non-crop resources differently than pigs that don't use crops?).

To address these questions, we used an approach analogous to step-selection functions (Thurfjell *et al.* 2014), in which we explored how the availability of crop and non-crop resources affected pig movement on a landscape and thus influenced a pig's resource utilization distribution. This approach required identifying agricultural and non-crop resources and then analyzing how crop and non-crop resources affected continuous-time animal movement over discrete space. We describe these two steps below.

## Agricultural and non-agricultural covariates

The agricultural covariates that we used in this analysis were the same NASS crop layers described in section *Agricultural resources*. The only difference was that instead of considering all 105 crop types as defined by NASS, we placed 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. We delineated these agricultural based on similarities in nutritional content and seasonal availability in a region [CHECK]. These groupings facilitated cross-study comparison regarding how crops affected pig movement.

[Describe how we included crop rasters into these models]

The non-agricultural covariates that we included in the model were: distance-to-nearest road, distance-to-nearest perennial water source, elevation, Normalized Difference Vegetation Index (NDVI) as a measure of plant productivity, tree canopy density, masting tree density [], temperature, precipitation, and snow depth. The biological rational behind each of these covariates, their spatial and temporal resolution, and their source are described in Table 1. The following section describes how these covariates were incorporated into the modeling framework to [describe feral swine movement and resource selection].

## A resource-dependent pig movement model

To quantitatively explore how agricultural and non-agricultural resources affect pig movement and resource selection, we used the modeling framework of Hanks *et al.* (2015) and Wilson *et al.* (2018). Generally, this framework leverages fine-scale, auto-correlated animal movement data along with 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.

The first step of this approach uses a trajectory of GPS fixes (not necessarily with equal fix times) and estimates animal movement as a function of continuous time. As we described above, we used a phenomenological functional movement model (FMM) to predict continuous time paths of animal movement (Buderman *et al.* 2016; Hooten *et al.* 2017). This first step is critical as fix times varied from 15 minutes to 2 hours across the analyses in our studies. The continuous-time FMM allowed us to sync-up these time scales (with increased uncertainty as the distance between fix times increased) such that we were making inference regarding resource selection on the same scale.

The second step of the analysis requires translating our continuous movement model into discrete, rasterized space. To do this, we used the `ctmcmove` package to convert our continuous-time movement path into discrete 30m by 30m grid cells (Fig. X) (Wilson *et al.* 2018). We used a 30m by 30m scale as this was the smallest scale at which we could obtain agricultural resource data across our various studies. To account for the uncertainty in our movement model, we generated 20 continuous-time, discrete space movement paths and performed all of the analyses described below on each of the movement paths (Hanks *et al.* 2015; Buderman *et al.* 2018).

Given these continuous-time, discrete-space trajectories, we then explored how agricultural and non-agricultural forage resources on a landscape affected pig movement using a continuous-time Markov Chain (CTMC) approach Hanks *et al.* (2015) [repetitive]. Within the CTMC framework, animal movement can be considered as a series of rates of moving from cell $i$ to an adjacent cell $j$, $\lambda_{ij}$. As with any continuous-time, discrete-state 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 resource covariates in cell $i$ and $j$. This is analogous to well-known step-selection functions in movement ecology (Thurfjell *et al.* 2014).

Using this CTMC framework, Hanks *et al.* (2015) showed that inference on how resource covariates affect $\lambda_{ij}$ can be done using 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 zero 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 $\tau_{ij}$ 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. Table 1 shows which of the covariates described in the previous section we considered as location-based drivers, directional drivers, or both. Table 1 provides a description of how each covariate was calculated.

### Model specification and fitting

We fit three CTMC models of the form given by equation 2, where the $\mathbf{X}$ varied between models. We fit the model to each pig separately using the covariates described in Table 1. For the agricultural resources, we only included crop types in the model that the pig used at some point during the time it was collared. If no crops were used, a crop covariate was not included in the model. The first model included the main effects of all of the location-based and direction covariates listed in Table 2. Moreover, we also included an interaction between directional persistence and whether or not a pig was in a particular crop type. This term accounted for possible changes in turning angle once a pig entered a crop field, e.g. we might a reduction in directional persistence if a pig starts foraging upon entering an agricultural field.

The second model we fit allowed for the influence of particular agricultural and non-agricultural resources to vary with time of day (see Table 1). We hypothesized that overall movement rate and directional persistence would change with time of day, as well as the tendency of pigs to move toward and away from agricultural resources. This hypothesis was based on previously observed behavior of wild boar actively seeking out agricultural fields at night before returning to cover during the day (Thurfjell *et al.* 2009). Moreover, we also allowed the tendency of pigs to toward or away from increasing plant productivity (e.g. a proxy for non-agricultural forage), canopy cover, distance to water, and elevation to vary with time of day. To be as consistent across populations that were sampled at different times of the year and different latitudes, we defined time of the day based on the sun position at a particular date and a particular location. In particular we broke the day into four solar periods: dawn to solar noon (morning), solar noon to dusk (afternoon), dusk to solar midnight (night), and solar midnight to dawn (early morning). [Biological relevance of these categories].

The third model we fit as identical to the second model, except that we now allowed the following covariates of the model to vary seasonally: overall movement rate, directional persistence, the tendency to move toward agricultural fields, the tendency to move toward increasing plant productivity, and the tendency to move toward water. [Give rationale for why we did this]. We defined season in terms of crop season: planting and harvest dates...

We fit each model with a LASSO Poisson GLM using the package `glmnet` in the R computing language (Friedman *et al.* 2010). The LASSO Poisson GLM shrinks variables that describe a small portion of variation in the data to zero allowing for smaller, more interpretable models with better out-of-sample performance [Site statistical learning]. We selected the the best fit regularization parameter as the one that minimized the five-fold cross-validated deviance (Friedman *et al.* 2010). We compared the three models using the minimum deviance from each model.

[Model validation?]

### Comparison of resource selection

After fitting the three models described above, we then qualitatively and quantitatively explored the roll of agricultural and non-agricultural resources on pig movement. We did this by analyzing the movement parameters predicted by the models described above [but which ones?]. These comparisons are compounded by the fact that pigs are not consistently sampled across seasons! [This is where is gets tricky 1. Made box plots across individuals]

# Results

Figures

1. Map of studies 2. Description figure of modeling framework 3. - 4. Summary of crop use across pigs 5. - 6. Some summary plots that show consistent effects of resources on pig movement. Also want to show how crop-users and non-crop users move differently. This is again confounded by season, though we can point out that this might no necessarily be an intrinsic property of the pig, but a property of the season. How should we group these pigs? Ecoregion? Study? Crops used? Landscape heterogeneity? 7. We could include some maps? Tough with time-varying effects

## Question 1: How do pigs use agricultural resources?

Summary plots of how pigs are using agricultural resources across different studies. This is more of an exploratory analysis based on the continuous time movement model.

### Question 2: How does crop-use affect the movement patterns and resource selection of feral swine?

# Discussion

Limitations 1. Some studies not directly comparable 2. Use doesn't mean damage 3.

# References


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

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.

Friedman, J., Hastie, T. & Tibshirani, R. (2010) Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, **33**, 1–22.

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.

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.

Thurfjell, H., Ball, J.P., Åhlén, P.A., Kornacher, P., Dettki, H. & Sjöberg, K. (2009) Habitat use and spatial patterns of wild boar Sus scrofa (L.): agricultural fields and edges. *European Journal of Wildlife Research*, **55**, 517–523.

Thurfjell, H., Ciuti, S. & Boyce, M.S. (2014) Applications of step-selection functions in ecology and conservation. *Movement Ecology*, **2**, 1–12.

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