A SPATIALLY VARYING STOCHASTIC DIFFERENTIAL EQUATION MODEL FOR ANIMAL MOVEMENT¹

By James C. Russell*, Ephraim M. Hanks † , Murali Haran † and David Hughes †

Muhlenberg College* and The Pennsylvania State University[†]

Animal movement exhibits complex behavior which can be influenced by unobserved environmental conditions. We propose a model which allows for a spatially varying movement rate and spatially varying drift through a semiparametric potential surface and a separate motility surface. These surfaces are embedded in a stochastic differential equation framework which allows for complex animal movement patterns in space. The resulting model is used to analyze the spatially varying behavior of ants to provide insight into the spatial structure of ant movement in the nest.

1. Introduction. Studying the movement of animals allows scientists to address fundamental questions in ecology and epidemiology. It can be used to understand how animals are impacted by their environment; for example, Gibert et al. (2016) studied the impact of temperature on animal movement to understand potential impacts of climate change and Watkins and Rose (2013) studied the impact of novel environments on small fish in a simulation-based analysis. Animal movement can give insight into the impact of external stimuli, as studied by Dodge et al. (2013) and Thiebault and Tremblay (2013). Further, animal movement is important in understanding the collective behavior of animal societies [Watkins and Rose (2013)]. Understanding these factors can benefit animal conservation [Killeen et al. (2014)] and can increase understanding of the impact of migration on disease dynamics [Altizer, Bartel and Han (2011)].

Improvements in technology have allowed scientists to observe animal movement at a fine temporal resolution over long periods of time [Toledo et al. (2014)]. This fine scale observation enables researchers to observe more complete individual paths revealing the animal's behavioral patterns in more detail and allowing for the fitting of more complex statistical models [Avgar et al. (2015), Bestley et al. (2015)]. Since animal movement data inherently consist of observations through time, it is intuitive to describe movement using differential equations, which are often used to describe dynamic systems. As any deterministic differential equation is unlikely to capture the movement of a single animal, we propose a stochastic differential equation (SDE) approach to modeling animal movement.

Received May 2016; revised September 2017.

¹Supported by the NSF Grant EEID 1414296.

Key words and phrases. Animal movement, stochastic differential equations, potential surface, Camponotus pennsylvanicus.

In this article, we will focus on the analysis of the movement of ants in a four chambered nest. As an entire population (colony) can be observed and tracked, ants provide an ideal system for studying animal movement [Mersch, Crespi and Keller (2013)], something rarely possible outside of laboratory systems. If the entire population has not been observed, like with any model for animal movement, conclusions should be limited to reflect the fact that the movement behavior of an animal directly impacts whether it is observed. For this same reason, ants provide an ideal system for studying the spread of information or of infectious disease in societies. Despite the fact that they live in close proximity to one another, researchers have hypothesized a "collective immunity" where collective behavior helps to prevent the spread of disease [Cremer, Armitage and Schmid-Hempel (2007), Quevillon et al. (2015)]. Proposed mechanisms, which result in this "collective immunity," include the spatial and social segregation of ants in the nest, which minimizes the number of foraging ants that come into contact with the queen. Ant behavior, however, can be challenging to model due to this social and spatial segregation within the nests [Quevillon et al. (2015)]. In this paper, we propose a spatially varying SDE model to capture ant movement behavior within the nest. Analysis of the ant movement indicates a tendency to move more quickly in the central chambers of the nest with ants utilizing the central chambers primarily as corridors for commuting. Simulation of new ants entering the nest highlights that the time needed to pass through sections of the nest varies based on spatial location.

Johnson et al. (2008) uses a continuous time correlated random walk (CTCRW) model based on integrated Brownian motion to model directional persistence in movement. Here, directional persistence refers to the tendency of animals to continue moving in the same direction and at a similar pace at nearby time points; this can, for instance, induce autocorrelation. However, Johnson et al. (2008) does not consider the case in which the behavior of an animal is dependent on its position. Quevillon et al. (2015) analyzes the movement of ants in a nest using the continuous time discrete space Markov chain model of Hanks, Hooten and Alldredge (2015). The discrete time specification used in Quevillon et al. (2015) and Hanks, Hooten and Alldredge (2015) requires the discretization of spatial location using a grid, analyzing movement by modeling the amount of time spent in each grid cell and the transition probabilities between cells. Their results reveal spatially varying movement behavior near the queen. However, movement data are typically observed in continuous space, and discretizing to a gridded space may introduce hard-to-resolve impacts on statistical inference that could result in errors in our conclusions about movement. In addition, changing the resolution of the gridded space can lead to changes in the resulting inference, and there is little guidance on how to choose a grid resolution. In contrast to Hanks, Hooten and Alldredge (2015), we model movement directly in continuous space using stochastic differential equations.

Stochastic differential equations have also been used for the movement of objects on a sphere [Brillinger (2012)]. This is essential for modeling movement

of animals such as elephant seals [Brillinger and Stewart (1998)] over large distances of the globe. Brillinger et al. (2001) and Preisler et al. (2004) propose an SDE-based model for movement on a potential surface which captures spatially varying drift in movement patterns across space. Potential functions are defined as functions of spatial location, and the negative gradient of this function determines the directional tendencies of an animal's movement at a specified location. These potential surface methods are also described in Brillinger et al. (2002) and in Brillinger et al. (2012). Potential surfaces have been used to analyze the movement of monk seals [Brillinger, Stewart and Littnan (2008)] and even the flow of play in soccer [Brillinger (2007)]. Potential functions have also been used to model constrained movement [Brillinger (2003)] and pairwise interactions in the movement of particles [Brillinger, Preisler and Wisdom (2011)].

Existing potential surface methods allow for flexible modeling of directional bias in movement. Using these models for movement with variation in the absolute speed of movement when there is no directional bias, however, is not straightforward as faster movement in a potential surface requires a steep surface which must correspond to a strong directional bias in movement. Variation in movement rate is important in ant systems, as ant nests typically contain "corridors" that connect important chambers in the nest. Ant movement through these corridors shows high velocity and directional persistence (correlated random walk movement), but little directional bias in these specific regions of the nest. Since some ants are moving in one direction while other ants are moving in the other, similar to movement of humans in a hallway, there may be no systematic directional drift. Existing models allow for directional persistence or spatially varying drift. For example, Preisler and Akers (1995) analyzes the movement of beetles and captures directional persistence using an autoregressive model for the turning angles of individuals. We propose an SDE approach to model movement behavior with all three features: (1) correlated movement patterns, (2) spatially varying drift through a potential surface and (3) spatial variation in movement rates through the inclusion of a spatially varying motility surface. We define a motility surface as a function of spatial location that determines the average overall movement rate at a specified spatial location.

The remainder of the paper is organized as follows. In Section 2, we introduce the carpenter ant system in detail. In Section 3, we discuss our proposed model for the ant movement data that incorporates autocorrelation, spatially varying drift and spatially varying absolute movement rate. In Section 4, we describe a discrete approximation of the model and our Bayesian inferential approach. The results of the application of our approach to ant movement data are presented in Section 5. We conclude, in Section 6, with a discussion and potential directions for future work.

2. Carpenter ant movement. We begin with a description of the ant movement data and describe an exploratory data analysis that motivates our modeling



FIG. 1. Movement paths of three sample ants.

approach. We analyze the movement of the ants in a nest constructed to mimic natural conditions, as closely as possible, in a lab setting. We used the common black carpenter ant, *Camponotus pennsylvanicus*, which nests in wood in temperate forests in the Eastern USA. We collected colonies between May and June 2015. The ants were placed in a nest structure consisting of four distinct chambers. Each chamber is divided into two sections by an internal barrier, creating a small passageway 12 mm across between the upper and lower halves of the chamber. Each of the four chambers measures 65 mm by 40 mm resulting in a total nest size of 65 mm by 160 mm. Each doorway between chambers is 6 mm across. There is an exit from the nest in chamber IV, leading to an area with food and water. The queen resides primarily in chamber I, far from the nest exit.

A plot containing three examples of individual ant movement paths, using linear interpolation between observed locations at each second, is given in Figure 1. The nest exit is marked with an "X" below chamber IV in each plot. The movement paths indicate that ant behavior is different within different chambers of the nest with faster and more directed movement happening in chambers II and III. This type of spatially varying behavior would be difficult to capture using the potential surface SDE approach [e.g., Brillinger et al. (2002) and Preisler et al. (2004)] as the ants move quickly in some regions of the nest but do not show consistent drift. To capture this type of behavior using potential surfaces alone would require estimation of potential surfaces that differ over time, which would complicate inference and increase the computational burden. Alternatively, our approach allows for the estimation of an additional motility surface that captures differences in the average movement rate of ants throughout the nest.

The data consist of the two-dimensional location coordinates $(x_{t_i,j}, y_{t_i,j})$, a time index t_i , which ranges from 1 to 3600 for the one hour observation period. In this

case the observations are made at evenly spaced intervals; however, this is often not the case in movement data. Additionally, there is a unique ant identifier *j* for each of 32 ants. One camera was positioned over each chamber of the nest. Ants were filmed under infrared lighting with GoPro Hero3 and Hero3+ cameras with modified infrared filters (RageCams, Sparta, MI) to simulate natural lighting conditions. Each ant was individually tagged with a unique identifier allowing a human observer to record their position. Observations were recorded at every second by clicking on the location of the ant using a custom software package. Measurement error is small, and there are no missing observations as the recorded videos allow for careful location selection and, if necessary, review of the movements of each individual. Locations were recorded for the entire hour for all ants that enter chamber IV at any time during the observation window. The dimensions of the nest were recorded by clicking on predetermined corners in the nest at the beginning of the observation period. Each camera records a separate section of the nest, so that there are four separate sets of observation for each ant. The observations in each of the four chambers can then be combined, using a common time index, to represent movement across the four chambers.

Several challenges arise in combining the data from the four cameras. At some time points (0.02% of observations) an ant is positioned at the door between two chambers and is observed by two cameras in two chambers at the same time. This could be due to different segments of the ants being visible in different chambers at the same time point. Further complications arise when the ants are not observed for a span of time. This can happen when the ant has exited the nest structure (5.46% of observations) or when the ant is situated in chamber doorways (1.08% of observations). When the ant has exited the nest structure, no cameras capture its movement. In what follows, we assume that when the ant re-enters the nest its movement is independent from prior in-nest movements. If the ant is between chambers and not visible on any camera, the ant's locations are linearly interpolated from the observations before and after the ant is in the entryway.

To explore the spatial movement behavior of the ants, a kernel density estimate for all observed ant locations and kernel density estimates of all observed empirical velocities in each of the nest chambers are plotted in Figure 2. In both cases, the kernel density estimates are calculated using the density function of the base stats package in R [Ihaka and Gentleman (1996)] using a Gaussian kernel and the default bandwidth. From Figure 2(a), we can see that the ants spend most of their time on the right-hand side of chamber IV; this indicates that there are certain preferred regions of the nest that ants tend to move toward, hinting at spatially varying drift. Figure 2(b) indicates that the average velocities in chambers I and IV tend to be less than the average velocities in chambers II and III. These exploratory results indicate that both the drift and the absolute ant velocity may vary spatially and suggest that the center chambers are primarily used for higher-velocity transit. To explore the temporal autocorrelation in movement, an autoregressive model of order 1 was fit to the empirical velocities of a randomly selected ant. The resulting



FIG. 2. Exploratory data analysis for carpenter ant movement.

estimated autoregressive parameter, 0.81 (standard error 0.015), indicates the need to model directional persistence. Our exploratory data analysis suggests that an appropriate model for these data should include temporal autocorrelation, spatially varying drift and a spatially varying movement rate. We develop a model which allows for this type of behavior in Section 3.

3. Spatially varying SDE model. In this section, we propose an SDE model for animal movement that captures:

- 1. directional persistence via a continuous time correlated random walk (CTCRW),
- 2. spatial variation in drift through a potential surface,
- 3. spatial variation in overall movement rate using a motility surface.

These three features have not, to our knowledge, been jointly utilized in any previous continuous space models. In previous work, potential surfaces have been used to model spatially varying movement [Brillinger et al. (2012), Preisler et al. (2004)]. When estimating potential surfaces, temporal autocorrelation has often been ignored to simplify the model and improve computational feasibility. We extend these methods by modeling directional persistence through a CTCRW, and expand the model to allow for variability in movement rate. This model will be used to analyze within-nest ant movement, where an individual's location at time t is denoted (x(t), y(t)) and its velocity is denoted $(v_x(t), v_y(t))$. We will begin by describing the movement of one individual, and in Section 4 we will generalize for multiple independent ants. 3.1. Continuous time correlated random walk. Animal movement is often autocorrelated through time, resulting in smooth movement paths. The CTCRW model presented by Johnson et al. (2008) provides one method for modeling dependence through time. The CTCRW model specifies an Ornstein–Uhlenbeck (O–U) process on an animal's velocity with directional drift μ . The O–U process is defined as a stochastic process that is stationary (the joint density remains the same for uniform shifts in time), Gaussian, Markovian and has continuous paths [Klebaner (2005)]. Alternatively, the O–U process can be derived using linear stochastic differential equations [Gardiner (1986)].

We consider a process that describes the movement of a particle in two dimensions. The movement in each dimension is assumed to follow an independent CTCRW. This is a reasonable assumption for animal movement since positive correlation in the x-dimension and y-dimension would result in bias toward movement in the North-East or South-West directions. The CTCRW model to describe the velocities $v_x(t)$, $v_y(t)$ and the locations x(t), y(t) is defined as

$$dv_x(t) = \beta (\mu_x - v_x(t)) dt + \sigma dW_{v_y}(t),$$

$$dv_y(t) = \beta (\mu_y - v_y(t)) dt + \sigma dW_{v_x}(t),$$

$$dx(t) = v_x(t) dt + \kappa dW_x(t),$$

$$dy(t) = v_y(t) dt + \kappa dW_y(t).$$

In the velocity equations, μ_x represents the mean drift in the x direction, μ_y represents the mean drift in the y direction, β controls the autocorrelation in movement, σ is related to the variability in velocity and $W_{v_x}(t)$, $W_{v_y}(t)$ represent independent Brownian motion processes with unit variance. In the equation for location, κ is related to the additional variability in location (i.e., measurement error, which relates to the error in "clicking location" when recording the locations of the ant from the recorded videos) and $W_x(t)$, $W_y(t)$ are again Brownian motion with unit variance. The case where $\kappa = 0$ gives intuition on the position of the object over time

(1)
$$x_t = x_0 + \int_0^t v_x(s) \, ds$$

The CTCRW process results in a model where the conditional mean of the discretized velocity is a weighted average of the previous velocity and the directional drift term μ . Several R packages have been developed due to the popularity of the CTCRW model for animal movement. For example, Albertsen et al. (2015) introduces the Template Model Builder (TMB) package which incorporates non-Gaussian error in a computationally efficient manner through Laplace approximation. Other examples of R packages to fit the CTCRW model include crawl [Johnson (2013)] and bsam [Jonsen (2015)]. Examples of the use of this model include Baylis et al. (2015), which analyzes the decline in population of southern sea lions, Northrup, Anderson and Wittemyer (2015), which studies the impact of habitat loss on the movement of mule deer and Rode et al. (2015), which analyzes the movement of polar bears in reaction to the decrease in sea ice habitat. The CTCRW model, however, assumes that the movement behavior is homogeneous in space. This assumption can be relaxed by allowing the drift term for the individual's velocity to be a function of the individual's location (x(t), y(t)). Doing so unifies the CTCRW model with the potential function models of Brillinger et al. (2002) and Preisler et al. (2004).

3.2. *Potential surfaces*. Potential surfaces are described in relation to animal movement modeling by Brillinger et al. (2002) and Preisler, Ager and Wisdom (2013). There are several examples of analyses using potential surfaces in movement modeling, including Preisler et al. (2004), which studies the influence of roads on the movement of elk, and Brillinger et al. (2012), which analyzes the movement of elk in a fenced-in experimental forest. In a potential function approach, the individuals, or ants in our case, are considered to be moving around on a surface with regions of attraction and repulsion.

First consider a potential surface, H(x, y), that is only a function of the current two-dimensional location of the object (x, y). This function can be thought of as a topological surface where objects are drawn to lower regions—similar to a marble moving on a curved surface. The mean direction of movement is, therefore, down the slope of the surface, or equivalently in the direction of the negative gradient of the potential surface. Using this analogy, the expectations of the x and y components of velocity, denoted $v_x(t)$ and $v_y(t)$ respectively, of an individual can each be calculated by taking the negative of the x and y derivatives of H(x, y)respectively:

$$E(v_x(t)) = -\frac{dH(x(t), y(t))}{dx},$$
$$E(v_y(t)) = -\frac{dH(x(t), y(t))}{dy}.$$

This spatially varying drift in movement is incorporated in the CTCRW model from Section 3.1 by defining the mean drift μ as the negative gradient of H(x, y). The result is a system of stochastic differential equations for correlated velocity with spatially varying movement bias defined by the potential surface H(x, y)

$$dv_x(t) = \beta \left(-\frac{dH(x(t), y(t))}{dx} - v_x(t) \right) dt + \sigma \, dW_{v_x}(t),$$

$$dv_y(t) = \beta \left(-\frac{dH(x(t), y(t))}{dy} - v_y(t) \right) dt + \sigma \, dW_{v_y}(t).$$

(2)

The SDE model for velocity (2) jointly models spatially varying drift through the potential function H(x, y) and temporal autocorrelation through the finite autoregressive parameter β . In previous work [e.g. Brillinger et al. (2002) and Preisler et al. (2004)], only the over-damped case is considered, where $\beta \rightarrow \infty$. Taking the limit as $\beta \rightarrow \infty$ eliminates directional persistence in movement, as it puts all of the importance on the gradient of the potential surface and eliminates the dependence on the animal's current velocity in (2). This simplifies estimation of the potential surface but fails to model the temporal autocorrelation, which is normally present in animal movement.

We model the potential surface using B-spline basis functions [de Boor (1978)]. Parametric models of the potential surface are also possible [Hanks, Hooten and Alldredge (2015), Quevillon et al. (2015)], but our goal in this analysis is to use the potential surface to flexibly model spatially varying directional drift within the nest. We thus assume the potential surface is given by

(3)
$$H(x, y) = \sum_{k,l} \gamma_{kl} \phi_k^{(M)}(x) \psi_l^{(M)}(y),$$

where $\phi_k^{(M)}(x)$ and $\psi_l^{(M)}(y)$ are B-spline basis functions of order M, with K basis functions in the *x*-direction and L basis functions in the *y*-direction so that $k \in \{1, ..., K\}$ and $l \in \{1, ..., L\}$. We chose to set M = 4, so that the potential surface H(x, y) has two continuous derivatives. As a result, the gradient of the potential surface (3) also has a continuous derivative, meaning it is relatively smooth. The potential surface H(x, y) only impacts movement through its gradient. Thus, only contrasts of the B-spline coefficients $\{\gamma_{kl}\}$ are identifiable, and we elect to subject them to the constraint $\sum_{k,l} \gamma_{kl} = 0$. Rather than using tuning to select the number of knots directly, the number of basis functions is set to 560, selected to keep the scale in the dimensions (65 mm by 160 mm) approximately equal (K = 16, L = 35) and penalized [Eilers and Marx (1996)] by using a zero mean multivariate normal prior on the coefficients $\{\gamma_{kl}\}$. Details of this prior will be given in Section 4.

The movement of the ants in our analysis is restricted by the locations of the walls in the ant chambers (Figure 1). Various models that account for restricted animal movement have been proposed [Brillinger (2003), Brost et al. (2015)]. In this case, we restrict ant movement near walls by augmenting the spatially smooth potential function (3) with an additive exponential potential R(x, y)

$$R(x, y, \mathbf{r}, r_1) = \exp\{-r_1(x - r_l^x)\} + \exp\{r_1(x - r_u^x)\} + \exp\{-r_1(y - r_l^y)\} + \exp\{r_1(y - r_u^y)\},$$

where $\mathbf{r} = (r_u^x, r_l^x, r_u^y, r_l^y)$; r_u^x and r_l^x represent the upper and lower wall boundaries in the *x*-dimension; and r_u^y and r_l^y represent the upper and lower wall bound-

1320

aries in the y-dimension. The parameter r_1 controls the rate of decay of the wall repulsion function. Other formulations for this basis are possible, but estimation using more flexible wall basis functions may require observations with a higher temporal frequency near the boundaries. The combined potential surface we consider is

(4)
$$H(x, y) = \sum_{k,l} \gamma_{kl} \phi_k^{(M)}(x) \psi_l^{(M)}(y) + R(x, y, \mathbf{r}, r_1).$$

3.3. Motility surface. In previous studies of the movement of ants in a nest, researchers have found that ants tend to move at different speeds in different areas of their nest [Quevillon et al. (2015)]. To incorporate this behavior in our SDE model, we propose a spatially varying motility surface M(x(t), y(t)), which scales the overall rate of movement in different parts of the nest. The motility surface scales the absolute movement rate in different spatial locations, allowing for flexible modeling of animal movement in which velocities can depend on the individual's environment. This results in the following stochastic differential equation model, where the location equation from the CTCRW model of Johnson et al. (2008) in each dimension has been adjusted to account for the spatially varying motility

(5)
$$dx(t) = M(x(t), y(t))v_x(t) dt + \kappa dW_x(t),$$
$$dy(t) = M(x(t), y(t))v_y(t) dt + \kappa dW_y(t).$$

When $\kappa = 0$, we get the physical interpretation of velocity scaled by the motility surface $x(t) = \int_0^t M(x(s), y(s))v_x(s) ds$, and when M(x, y) = 1, we get back (1).

For an example of movement behavior induced by the motility surface, consider a habitat separated into two distinct regions—one in which ants move fast and another in which ants move more slowly. If $M(\cdot) = 2$ in the fast region, and $M(\cdot) = 1$ in the slow region, the ants will move approximately twice as fast in the fast region as they will in the slow region. Equations (2) and (5) together define a continuous time two-dimensional stochastic movement process. Note that the v_x and v_y can no longer be interpreted directly as velocity since they must be multiplied by the motility surface M(x, y) to obtain the overall velocity of the ant at that specific location. The actual velocities are a scaled version of v_x and v_y ; at (x(t), y(t)) the animal's mean movement rate vector ignoring autocorrelation is $(M(x, y) \cdot v_x(t), M(x, y) \cdot v_y(t))$.

For the motility surface M(x(t), y(t)), we again utilize penalized B-spline basis functions,

(6)
$$M(x, y) = \sum_{q,r} \alpha_{qr} \zeta_q^{(P)}(x) \xi_r^{(P)}(y),$$

where $\zeta_q^{(N)}(x)$ and $\xi_r^{(N)}(y)$ are B-spline basis functions of order *P*, with *Q* basis functions in the *x*-direction and *R* basis functions in the *y*-direction so that $q \in$

 $\{1, \ldots, Q\}$ and $r \in \{1, \ldots, R\}$. To facilitate computation, we set P = 4, Q = K and R = L so that the motility surface and the potential surface share the same basis functions.

3.4. *Identifiability.* Without imposing constraints, there are multiple (infinitely many) combinations of parameters, which will yield an identical formula for the model defined by equations (2) and (5). First, the potential surface H(x, y) only enters the model through its gradient. Therefore, adding some constant c_1 to the surface results in an equivalent model $[H^*(x, y) = H(x, y) + c_1]$. As discussed in Section 3.2, the constant c_1 is fixed to 0 by constraining the sum of the basis coefficients $\{\gamma_{kl}\}$ to equal 0.

Second, the motility surface and velocity terms are only identifiable up to a multiplicative constant c_2 . Multiplying the motility surface by c_2 and the velocities by $\frac{1}{c_2}$ yields an equivalent model since the following pairs of stochastic differential equations yield identical distributions for the observed locations of the ants

$$\begin{aligned} \frac{1}{c_2} dv_x(t) &= \beta \left(-\frac{1}{c_2} \frac{dH(x(t), y(t))}{dx} - \frac{1}{c_2} v_x(t) \right) dt + \frac{1}{c_2} \sigma \, dW_{v_x}(t), \\ \frac{1}{c_2} dv_y(t) &= \beta \left(-\frac{1}{c_2} \frac{dH(x(t), y(t))}{dy} - \frac{1}{c_2} v_y(t) \right) dt + \frac{1}{c_2} \sigma \, dW_{v_y}(t), \\ dx(t) &= c_2 M(x(t), y(t)) \frac{1}{c_2} v_x(t) \, dt + \kappa \, dW_x(t), \end{aligned}$$

$$(7) \qquad dy(t) &= c_2 M(x(t), y(t)) \frac{1}{c_2} v_y(t) \, dt + \kappa \, dW_y(t). \end{aligned}$$

To obtain identifiability, the model is constrained by setting $c_2 = \sigma^2$, or, equivalently, fixing $\sigma^2 = 1$.

4. Inference. There is no clear method to analytically solve the nonlinear SDEs given by equations (2) and (5); so, instead, we consider approximate solutions using the Euler–Maruyama method [Kloeden and Platen (1992)], which is based on a first-order Taylor expansion. We approximate the continuous time processes (2) and (5) at a set of observed (or simulated) discrete times t_i , $i = 1, 2, ..., N_{sim}$, where N_{sim} is the number of points in the path (observed or simulated). Higher-order Euler–Maruyama approximations could be used to improve accuracy, and these higher-order methods may be interesting directions for future research—both in the context of simulation and inference. Higher-order approximations are not necessary in this case due to the high temporal frequency of discrete observations—every second for the ants. An analysis of the impact of the time discretization in the Euler–Maruyama method is presented in the Appendix

[Russell et al. (2018)]. The Euler–Maruyama method results in the following discrete approximations for equations (2) and (5)

(8)

$$v_{t_{i+1}}^{x} = v_{t_{i}}^{x} + \beta \left(-\nabla H(x_{t_{i}}, y_{t_{i}}) - v_{t_{i}}^{x}\right) \Delta_{t} + \sigma \varepsilon_{2x},$$

$$v_{t_{i+1}}^{y} = v_{t_{i}}^{y} + \beta \left(-\nabla H(x_{t_{i}}, y_{t_{i}}) - v_{t_{i}}^{y}\right) \Delta_{t} + \sigma \varepsilon_{2y},$$

$$x_{t_{i+1}} = x_{t_{i}} + M(x_{t_{i}}, y_{t_{i}}) v_{t_{i}}^{x} \Delta_{t} + \kappa \varepsilon_{1x},$$

$$y_{t_{i+1}} = y_{t_{i}} + M(x_{t_{i}}, y_{t_{i}}) v_{t_{i}}^{y} \Delta_{t} + \kappa \varepsilon_{1y},$$

where Δ_t is the time step of the discretized solver and ε_{1x} , ε_{1y} , ε_{2x} , ε_{2y} are independent Gaussian random variables with mean 0 and variance Δ_t . These discrete Euler–Maruyama approximations will be used for inference. This is similar to Wikle and Hooten (2010), which uses a second-order Taylor expansion to approximate nonlinear dynamics in statistical models with differential equations. A two-dimensional simulation generated using the Euler–Maruyama method is presented in the Appendix [Russell et al. (2018)] to illustrate the range of behavior possible under our SDE model.

We use Bayesian inference to estimate the model parameters. The Euler– Maruyama approximations of the SDEs are used to provide an approximation for the likelihood function in the posterior distribution. Markov chain Monte Carlo is used to draw samples from this approximate posterior distribution. Vague priors are used for the the location variability parameter $\kappa^2 \sim \text{Gamma}(0.001, 0.001)$, and for the autocorrelation parameter $f_{\beta}(\beta) \propto \exp(\frac{-(\beta-1)^2}{20,000})I(\beta > 0)$, a normal distribution truncated to be positive. To test sensitivity to prior specification, inference was performed again using priors with twice the variance, and there was no significant difference in the results. For identifiability σ^2 is fixed at 1. To induce spatial smoothness in the potential and motility surfaces, a proper CAR covariance structure [Banerjee, Carlin and Gelfand (2014), Chapter 3] is used for the coefficients of both of the B-spline expansions in (9) and (6). The model was also fit with an uncorrelated prior on all of the spatial surface parameters, and the resulting surface estimates were very jagged and indicated overfitting (results not shown). Thus, for the potential surface coefficients, the prior selected is

$$\boldsymbol{\gamma} \sim N(\boldsymbol{0}, (\boldsymbol{\tau}_{\gamma}(\boldsymbol{D} - \boldsymbol{\rho}_{\gamma} \boldsymbol{Q}))^{-1}),$$

where $(D)_{jj} = \{\text{number of neighbors of } j\}; (D)_{ij} = 0 \text{ if } i \neq j; (Q)_{jj} = 0 \text{ and } Q_{ij} = I(i, j \text{ are neighbors}).$ For the smoothness parameter we use the prior $\rho_{\gamma} \sim \text{Uniform}(0.01, 0.99)$. The parameter τ_{γ} is a tuning parameter that determines the scale of the potential surface. The prior distribution for τ_{γ} is set to be an exponential distribution $\tau_{\gamma} \sim \exp(\mu_{\alpha}^2)$ so that the observed data will inform the level of tuning, and the distribution of the potential surface will scale with the level of the motility surface through μ_{α} . As discussed in Section 3.4, only contrasts of the B-spline coefficients $\{\gamma_{kl}\}$ are identifiable, so we utilize the constraint $\sum_{k,l} \gamma_{kl} = 0$.

This is done through a linear transformation of a set of unconstrained basis function coefficients $\tilde{\gamma}$ [Gelfand et al. (2010), Chapter 12]

(9)
$$\boldsymbol{\gamma} = \boldsymbol{\tilde{\gamma}} - \boldsymbol{\Sigma} \mathbf{1}^t (\mathbf{1} \boldsymbol{\Sigma} \mathbf{1}^t)^{-1} (\mathbf{1} \boldsymbol{\tilde{\gamma}} - \mathbf{0}),$$

where Σ represents the covariance matrix of the unconstrained coefficients.

Similarly, the coefficients of the B-spline basis functions for the motility surface were assigned a CAR prior

$$\boldsymbol{\alpha} \sim N(\mu_{\alpha} \mathbf{1}, \mu_{\alpha}^{2}(\tau_{\alpha}(\boldsymbol{D} - \rho_{\alpha} \boldsymbol{Q}))^{-1}).$$

For the motility surface the prior for the smoothing parameter is again set to $\rho_{\alpha} \sim \text{Uniform}(0.01, 0.99)$. For identifiability, as discussed in Section 3.4, the tuning parameter τ_{α} is set to 9 so that 99.8% of the prior mass is positive. The parameter μ_{α} adds flexibility to the model as it scales the motility surface. The inclusion of μ_{α} allows for approximating c_2 in equation (7), and, therefore, it can be used to approximate σ_2 if the identifiability constant in equation (7) had instead been fixed at 1. Additionally, μ_{α} alters the smoothing such that the penalization on variability by the multivariate Gaussian prior in the motility surface coefficients is less in regions of greater relative motility. A Gaussian prior is used for the scaling parameter $\mu_{\alpha} \sim N(1, 1)$. A lognormal distribution is used for the prior on the wall repulsion parameter $r_1 \sim \log Norm(10, 1)$. The analysis was rerun with sparse priors (the variability of each prior distribution was doubled), and there was no significant change in the posterior means or credible intervals for all parameters. The full conditional distributions for most parameters are amenable to Gibbs sampling and are presented in the Appendix [Russell et al. (2018)].

Block update MCMC is used to sample from the posterior distributions of model parameters. When available, the updates are drawn from full conditional distributions (a Gibbs sampler). The coefficients of the potential and motility surfaces, γ and α , are each updated as a separate block. 10^5 samples are drawn from the posterior distribution and convergence of the Markov chains is determined by monitoring Monte Carlo standard errors using the batch means procedures [Flegal, Haran and Jones (2008), Jones et al. (2006)]. The initial 20,000 values of the chain are discarded as burn in—as the initial estimates for the potential surface and motility surface are difficult to select—resulting in inaccurate parameter estimates at the beginning of the chain. Multiple chains with different starting values were run to ensure estimates are robust across initial values.

Inference here is computationally taxing, as it takes approximately six days to generate 10^5 samples from the posterior distribution on a single core of a 2.7 GHz Intel Xeon Processor with code written in R. The computing time scales at a linear rate with increases in observations. Reducing the computation time needed for inference is one of the primary goals of ongoing research. In this case, the computational complexity of the model is driven by the large number of latent variables (213,534 velocities $v_{t,j}^{(x)}$ and $v_{t,j}^{(y)}$), and the number of basis functions used for the

1324

potential and motility surfaces. Simulations under various settings show that developing a method to accurately approximate latent velocities and reducing the number of basis functions for the surfaces greatly improves computational efficiency. Bayesian implementation allows for straightforward estimation of tuning parameters for the smoothness of our B-spline surfaces, and avoids the additional difficulties of selecting the number of basis functions using other methods such as generalized cross validation. Improving the computational efficiency of inference is an important direction for future work, especially due to the availability of data sets with more individuals and observation over longer time periods.

5. Results. Posterior sample means and 95% credible intervals for the model parameters are presented in Table 1. There is significant autocorrelation in the ant's velocity as the credible interval for β does not include 1, but the autocorrelation is not strong. Note that in a continuous time framework (where $\Delta \rightarrow 0$) no autocorrelation corresponds to $\beta \rightarrow \infty$, but in the discrete approximation with $\Delta = 1$, no autocorrelation in velocity corresponds to $\beta = 1$. This is reasonable for ant movement, as the ant paths are not smooth since the ants tend to change direction suddenly inside the nest. The estimate for κ is small, which indicates that there is not much additional location variability after conditioning on latent velocity. Our estimate for r_1 indicates that there is a repulsion behavior from the wall, but the size of the effect is relatively small.

Estimates of the posterior mean potential surface and motility surface are constructed by taking the pointwise posterior mean of each coefficient and plotting the resulting surfaces over a fine grid. Estimating the variability in the potential surface is complicated by the fact that we are only interested in the relative height. Shifting the level of the entire surface by a constant has no impact on movement since the potential surface only impacts behavior through its gradient. The potential surface, plotted in Figure 3(a), reveals a tendency to move away from the walls on the left and right sides of the nest, particularly in the center chambers (II and III). This is consistent with ants that turn as they approach the walls while traveling between chambers.

The motility surface, plotted in Figure 3(b), reveals a tendency to move faster in the center of the nest—in chambers II and III. In chambers I and IV, where the ants

TABLE 1Ant results		
Parameter	Posterior mean	Credible interval
β	0.872	(0.866, 0.877)
κ	0.00133	(0.00131, 0.00134)
r_1	0.078	(0.044, 0.114)

Parameter estimates for ant movement data.



FIG. 3. Ant posterior mean surfaces.

spend most of their time, the movement rate is slower. This suggests that the central chambers (II and III) are mainly used by the ants for commuting between the entrance chamber (IV) and the deepest chamber (I) of the nest. Additionally, there is low relative motility in each of the doorways between chambers. The combined gradient field, which is estimated by multiplying the gradient of the estimated potential surface by the magnitude of the estimated motility surface over a spatial grid, plotted in Figure 3(c), reveals both the tendency to move away from walls and the tendency to move faster in the central chambers.

The 95% point wise credible intervals for the motility surface estimated from the ant movement observations, along with a black line denoting the movement path of an arbitrarily selected ant, are plotted in Figure 4. The upper limits in most regions in chambers I and IV are below the lower limits in parts of chambers II and III, indicating that the difference in movement rates in the different areas of the nest are statistically significant. Additionally, the regions with low relative motility in the doorways are apparent, particularly in the doorway connecting the chambers I and II and in the doorway connecting chambers II and III. This might be due to the linear interpolation procedure described in Section 2, or it could reveal a tendency to move slower in the doorways between these chambers. A simpler model for animal movement would not be able to capture this variation, nor would it model the statistically significant autocorrelation presented in Table 1.

To assess model fit, the one step ahead mean prediction error for the ant locations is calculated. At each time step, the next location $(\hat{x}_{t_i,j}, \hat{y}_{t_i,j})$ (for all 106,767 observations of ants in the nest) is predicted using the model by drawing predicted locations at each iteration of the Markov chain. The average prediction error, which is the distance between the predicted location and the observed location for each of the 106,767 observations, at each MCMC iteration is calculated using

$$\frac{1}{106,767}\sum_{i}\sum_{j}\sqrt{(\hat{x}_{t_{i},j}-x_{t_{i},j})^{2}+\hat{y}_{t_{i},j}-y_{t_{i},j})^{2}}.$$



FIG. 4. Ant posterior motility surface.

The result is then averaged over all draws from the Markov chain. The prediction error was estimated for three models—the full model (described above), a model with a constant potential surface $[H(x, y) \equiv 0]$ and a model with a constant motility surface $[M(x, y) \equiv 1]$. The mean prediction error for each of the models is, respectively 0.09 mm, 0.11 mm and 0.64 mm. This indicates that in this case the model that estimates both motility and potential surfaces is the best of the three at predicting animal movements. Further, the model with a constant motility surface performs the worst of the three models that we considered. This reflects the importance of incorporating variability in movement rate in different spatial locations within the nest.

6. Discussion. The SDE-based model we have developed allows for autocorrelated movement and flexible spatially varying drift and velocity. Ant movement observed every second within a nest for an hour is investigated, and the transit behavior of ants between different chambers is captured. The motility surface impacts the rate at which ants spread through the nest. Modeling this is important as the spatially varying rate of movement has important implications for the spread of resources, pathogens and pesticides within the nest. The spatially varying model presented is used here for experimental data. The applicability of this model to observational data should be considered on a case-by-case basis.

This model is flexible, allowing for temporal autocorrelation and movement, which is dependent on an animal's environment. The environmental dependency can be based on covariates, such as resources in the area; it can be nonparametric, using a variety of basis functions, or it can be a semiparametric structure combining both environmental features and the additional spatially varying behavior. For example, our results for the movement of ants indicate that the motility surface plotted in Figure 3(b) could be effectively modeled using dummy covariates for each "chamber." As it is often impossible to observe all environmental covariates when analyzing animal movement over large areas, the ability to estimate the surface using nonparametric methods is important. Thus, this model provides a flexible framework to model movement to help understand a variety of processes. Examples include resource selection, impacts of environmental changes and the spread of agents—such as pathogens or nutrients—through animal societies.

Some parts of the model are only identifiable up to a multiplicative constant. Therefore, interpretation of the estimated surfaces should be limited on the scale of the motility and potential surfaces. Further, computation becomes more burdensome as the number of basis functions and observations increase. Thus, more efficient methods for inference would be necessary to estimate surfaces with high resolution features, such as walls or fences, in very large regions. The model requires a high temporal rate of observation of animal locations due to the Euler–Maruyama approximation to the SDEs. To correct this issue, locations between observation times can be imputed from the model, however, this may become computationally taxing since the motility and potential surface estimates require using B-spline basis representation for these imputed observations. Updating the basis representation of these imputed locations at every iteration in a Metropolis–Hastings algorithm is often slow. Alternatively, higher-order approximations to the SDEs can be used, which may provide more accurate numerical approximations of the underlying SDE models (2) and (5).

In future work, temporally varying behavior will be considered. Distinct potential surfaces at different time points may capture the movement behavior of ants without the need for a motility surface. A state-space model with distinct potential surfaces and movement parameters in different states could be utilized, however, these states may be difficult to identify if the behavior in different states is similar [Beyer et al. (2013)]. This could also allow for the interaction of ants through combining a latent interaction network model [Scharf et al. (2016)] with potential surfaces consisting of directed movement toward specific neighboring ants.

SUPPLEMENTARY MATERIAL

Supplement to "A spatially varying stochastic differential equation model for animal movement" (DOI: 10.1214/17-AOAS1113SUPP; .pdf). We provide additional information including prior distribution specification, full conditional distributions, analysis of the discretization error, an application to simulated data and an application to the spread of pathogens.

REFERENCES

- ALBERTSEN, C. M., WHORISKEY, K., YURKOWSKI, D., NIELSEN, A. and MILLS, J. (2015). Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. *Ecology* 96 2598–2604.
- ALTIZER, S., BARTEL, R. and HAN, B. A. (2011). Animal migration and infectious disease risk. *Science* **331** 296–302.
- AVGAR, T., BAKER, J. A., BROWN, G. S., HAGENS, J. S., KITTLE, A. M., MALLON, E. E., MCGREER, M. T., MOSSER, A., NEWMASTER, S. G., PATTERSON, B. R. et al. (2015). Spaceuse behaviour of woodland caribou based on a cognitive movement model. *J. Anim. Ecol.* 84 1059–1070.
- BANERJEE, S., CARLIN, B. P. and GELFAND, A. E. (2014). *Hierarchical Modeling and Analysis* for Spatial Data. CRC Press, Boca Raton.
- BAYLIS, A. M., ORBEN, R. A., ARNOULD, J. P., CHRISTIANSEN, F., HAYS, G. C. and STANI-LAND, I. J. (2015). Disentangling the cause of a catastrophic population decline in a large marine mammal. *Ecology* 96 2834–2847.
- BESTLEY, S., JONSEN, I. D., HINDELL, M. A., HARCOURT, R. G. and GALES, N. J. (2015). Taking animal tracking to new depths: Synthesizing horizontal–vertical movement relationships for four marine predators. *Ecology* **96** 417–427.
- BEYER, H. L., MORALES, J. M., MURRAY, D. and FORTIN, M.-J. (2013). The effectiveness of Bayesian state-space models for estimating behavioural states from movement paths. *Methods Ecol. Evol.* **4** 433–441.
- BRILLINGER, D. R. (2003). Simulating constrained animal motion using stochastic differential equations. In Probability, Statistics and Their Applications: Papers in Honor of Rabi Bhattacharya. Institute of Mathematical Statistics Lecture Notes—Monograph Series 41 35–48. IMS, Beachwood, OH. MR1999413
- BRILLINGER, D. R. (2007). A potential function approach to the flow of play in soccer. J. Quant. Anal. Sports **3** 3. MR2304568
- BRILLINGER, D. (2012). A particle migrating randomly on a sphere. In *Selected Works of David Brillinger* 73–87. Springer, New York.
- BRILLINGER, D. R. and STEWART, B. S. (1998). Elephant-seal movements: Modelling migration. Canad. J. Statist. 26 431–443.
- BRILLINGER, D., PREISLER, H. and WISDOM, M. (2011). Modelling particles moving in a potential field with pairwise interactions and an application. *Braz. J. Probab. Stat.* 25 421–436.
- BRILLINGER, D. R., STEWART, B. S. and LITTNAN, C. L. (2008). Three months journeying of a Hawaiian monk seal. In *Probability and Statistics: Essays in Honor of David A. Freedman* 246–264. IMS, Beachwood, OH.
- BRILLINGER, D. R., PREISLER, H. K., AGER, A. A., KIE, J. and STEWART, B. S. (2001). Modelling movements of free-ranging animals. Univ. Calif. Berkeley Statistics, Technical Report 610.
- BRILLINGER, D. R., PREISLER, H. K., AGER, A. A., KIE, J. G. and STEWART, B. S. (2002). Employing stochastic differential equations to model wildlife motion. *Bull. Braz. Math. Soc. (N.S.)* 33 385–408.
- BRILLINGER, D. R., PREISLER, H. K., AGER, A. A. and KIE, J. (2012). The use of potential functions in modelling animal movement. In *Selected Works of David Brillinger* 385–409. Springer, Berlin.
- BROST, B. M., HOOTEN, M. B., HANKS, E. M. and SMALL, R. J. (2015). Animal movement constraints improve resource selection inference in the presence of telemetry error. *Ecology* 96 2590–2597.
- CREMER, S., ARMITAGE, S. A. O. and SCHMID-HEMPEL, P. (2007). Social immunity. *Curr. Biol.* **17** R693–R702.

- DE BOOR, C. (1978). A Practical Guide to Splines. Applied Mathematical Sciences 27. Springer, New York–Berlin. MR0507062
- DODGE, S., BOHRER, G., WEINZIERL, R., DAVIDSON, S. C., KAYS, R., DOUGLAS, D., CRUZ, S., HAN, J., BRANDES, D. and WIKELSKI, M. (2013). The environmental-data automated track annotation (Env-DATA) system: Linking animal tracks with environmental data. *Mov. Ecol.* **1** 3.
- EILERS, P. H. and MARX, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statist. Sci.* 89–102.
- FLEGAL, J., HARAN, M. and JONES, G. (2008). Markov chain Monte Carlo: Can we trust the third significant figure? *Statist. Sci.* 23 250–260.
- GARDINER, C. (1986). Handbook of stochastic methods for physics, chemistry and the natural sciences. *Springer Ser. Synergetics* **13** 149–168.
- GELFAND, A. E., DIGGLE, P., GUTTORP, P. and FUENTES, M. (2010). Handbook of Spatial Statistics. CRC Press, Boca Raton.
- GIBERT, J. P., CHELINI, M.-C., ROSENTHAL, M. F. and DELONG, J. P. (2016). Crossing regimes of temperature dependence in animal movement. *Glob. Change Biol.* **22** 1722–1736.
- HANKS, E. M., HOOTEN, M. B. and ALLDREDGE, M. W. (2015). Continuous-time discrete-space models for animal movement. Ann. Appl. Stat. 9 145–165.
- IHAKA, R. and GENTLEMAN, R. (1996). R: A language for data analysis and graphics. J. Comput. Graph. Statist. 5 299–314.
- JOHNSON, D. (2013). crawl: Fit continuous-time correlated random walk models to animal movement data. R package version 1.4-1.
- JOHNSON, D., LONDON, J., LEA, M. and DURBAN, J. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology* **89** 1208–1215.
- JONES, G., HARAN, M., CAFFO, B. and NEATH, R. (2006). Fixed-width output analysis for Markov chain Monte Carlo. J. Amer. Statist. Assoc. 101 1537–1547.
- JONSEN, I. (2015). bsam: Bayesian state-space models for animal movement. R package version 0.43.1.
- KILLEEN, J., THURFJELL, H., CIUTI, S., PATON, D., MUSIANI, M. and BOYCE, M. S. (2014). Habitat selection during ungulate dispersal and exploratory movement at broad and fine scale with implications for conservation management. *Mov. Ecol.* 2 15.
- KLEBANER, F. C. (2005). Introduction to Stochastic Calculus with Applications. Imperial College Press, London.
- KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin.
- MERSCH, D. P., CRESPI, A. and KELLER, L. (2013). Tracking individuals shows spatial fidelity is a key regulator of ant social organization. *Science* **340** 1090–1093.
- NORTHRUP, J. M., ANDERSON, C. R. and WITTEMYER, G. (2015). Quantifying spatial habitat loss from hydrocarbon development through assessing habitat selection patterns of mule deer. *Glob. Change Biol.* **21** 3961–3970.
- PREISLER, H. K., AGER, A. A. and WISDOM, M. J. (2013). Analyzing animal movement patterns using potential functions. *Ecosphere* 4 art32.
- PREISLER, H. K. and AKERS, R. P. (1995). Autoregressive-type models for the analysis of bark beetle tracks. *Biometrics* 259–267.
- PREISLER, H. K., AGER, A. A., JOHNSON, B. K. and KIE, J. G. (2004). Modeling animal movements using stochastic differential equations. *Environmetrics* 15 643–657.
- QUEVILLON, L. E., HANKS, E. M., BANSAL, S. and HUGHES, D. P. (2015). Social, spatial, and temporal organization in a complex insect society. Scientific Reports.
- RODE, K. D., WILSON, R. R., REGEHR, E. V., MARTIN, M. S., DOUGLAS, D. C. and OLSON, J. (2015). Increased land use by Chukchi Sea polar bears in relation to changing sea ice conditions. *PLoS ONE* 10 e0142213.

- RUSSELL, J. C., HANKS, E. M., HARAN, M. and HUGHES, D. (2018). Supplement to "A spatially varying stochastic differential equation model for animal movement." DOI:10.1214/17-AOAS1113SUPP.
- SCHARF, H. R., HOOTEN, M. B., FOSDICK, B. K., JOHNSON, D. S., LONDON, J. M. and DUR-BAN, J. W. (2016). Dynamic social networks based on movement. *Ann. Appl. Stat.* 10 2182–2202. MR3592053
- THIEBAULT, A. and TREMBLAY, Y. (2013). Splitting animal trajectories into fine-scale behaviorally consistent movement units: Breaking points relate to external stimuli in a foraging seabird. *Behav. Ecol. Sociobiol.* **67** 1013–1026.
- TOLEDO, S., KISHON, O., ORCHAN, Y., BARTAN, Y., SAPIR, N., VORTMAN, Y. and NATHAN, R. (2014). Lightweight low-cost wildlife tracking tags using integrated transceivers. In *Education* and Research Conference (EDERC), 2014 6th European Embedded Design 287–291.
- WATKINS, K. S. and ROSE, K. A. (2013). Evaluating the performance of individual-based animal movement models in novel environments. *Ecological Modelling* **250** 214–234.
- WIKLE, C. K. and HOOTEN, M. B. (2010). A general science-based framework for dynamical spatio-temporal models. *Test* **19** 417–451.

J. C. RUSSELL MUHLENBERG COLLEGE 2400 CHEW ST Allentown, Pennsylvania 18104 USA E-MAIL: JamesRussell@muhlenberg.edu E. M. HANKS M. HARAN D. HUGHES THE PENNSYLVANIA STATE UNIVERSITY OLD MAIN STATE COLLEGE, PENNSYLVANIA 16801 USA