MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **12**, Number **3**, June **2015**

doi:10.3934/mbe.2015.12.585

pp. **585–607**

AN AGGREGATE STOCHASTIC MODEL INCORPORATING INDIVIDUAL DYNAMICS FOR PREDATION MOVEMENTS OF *ANELOSIMUS STUDIOSUS*

Alex John Quijano
Department of Mathematics & Statistics
East Tennessee State University
Johnson City, TN, 37614, USA

Michele L. Joyner[*][†], Edith Seier[*][†], Nathaniel Hancock[‡]
Michael Largent[‡] and Thomas C. Jones[‡]

[*]Department of Mathematics & Statistics
[†]Institute for Quantitative Biology
[‡]Department of Biological Sciences
East Tennessee State University
Johnson City, TN, 37614, USA

(Communicated by H. T. Banks)

Abstract. In this paper, we discuss methods for developing a stochastic model which incorporates behavior differences in the predation movements of *Anelosimus studiosus* (a subsocial spider). Stochastic models for animal movement and, in particular, spider predation movement have been developed previously; however, this paper focuses on the development and implementation of the necessary mathematical and statistical methods required to expand such a model in order to capture a variety of distinct behaviors. A least squares optimization algorithm is used for parameter estimation to fit a single stochastic model to an individual spider during predation resulting in unique parameter values for each spider. Similarities and variations between parameter values across the spiders are analyzed and used to estimate probability distributions for the variable parameter values. An aggregate stochastic model is then created which incorporates the individual dynamics. The comparison between the optimal individual models to the aggregate model indicate the methodology and algorithm developed in this paper are appropriate for simulating a range of individualistic behaviors.

1. **Introduction.**

The mathematical modeling of systems in which the underlying dynamics are probabilistic in nature arises in many applications such as pharmacokinetics, ecology, and electromagnetics [1]. In this paper, we develop a methodology for creating a stochastic model of movement in which the parameters in the model behave according to an underlying unknown probability distribution which must be estimated. The approach in this paper is similar to the nonlinear mixed effects modeling (NLME) or hierarchical nonlinear modeling approach discussed in the

---

2010 *Mathematics Subject Classification.* Primary: 92B99, 60H10, 60H35; Secondary: 65C30, 97M60.

Key words and phrases. Stochastic differential equation, stochastic model, probability distribution, *anelosimus studiosus*, animal movements, parameter estimation, hierarchical modeling.

This research was partially supported by the National Science Foundation grant DMS-1128954.

paper by Davidian, et. al. [3] in which a population model is constructed using individual-leveled models. In both approaches, NLME and the one presented in this paper, it is assumed that individuals behave differently and hence are modeled using different parameter values. The individual parameter values are then used to develop a population-based model. In this paper, we use the individual models to estimate probability distributions for and relationships between the parameter values. Although this paper focuses on a specific application to predation movements of the spider *Anelosimus studiosus*, the same methodology developed in this paper can be used for a variety of modeling applications in which the parameters are assumed to be variable across individuals.

Grinstead et. al. [6] and Pruitt et. al. [14] gave evidence that spiders participate differently in attacking prey depending on whether or not they exhibit the behavioral characteristics for boldness. Those spiders with the characteristic for boldness are more likely to actively participate in prey capture when compared to other adult spiders in the colony. In this paper, we focus on predation movements of the subsocial spider *Anelosimus studiosus* with the eventual goal of using this model to determine the probability of successful prey capture as a function of individual predation dynamics and varied spacing within the web. *Anelosimus studiosus* is a spider from the family Theridiidae that ranges from Argentina to northeastern United States [10]. As such, these spiders are cobweb weavers which means their webs are not sticky enough for a prey to be immobilized; their prey cannot fly away and must walk to the edge of the web to escape [9, 10].

In section 2, we discuss the original model for the movement of *A. studiosus* during predation and compare simulations from the previous model to current data to motivate the need for an improved model which more accurately models an individual spider. In Section 3, we discuss the collection and processing of the data. Section 4 describes the updated model including the essential features of *A. studiosus* movement during prey capture which must be integrated within the model as well as pseudocode for running the model. The method for estimating parameters for the model of an individual spider is given in Section 5 and then a comparison of these parameters across spiders in discussed in Section 6. Simulations of the aggregate model are given in Section 7 and compared to each of the individual data sets. Conclusions are discussed in Section 8.

2. **Original model.**

Joyner et. al. [11] previously developed a stochastic model of the form

$$d\mathbf{r}(t) = \mu\{\mathbf{r}(t), t\}\, dt + \Sigma\{\mathbf{r}(t), t\}\, d\mathbf{W}(t) \tag{1}$$

that simulates the movement of *A. studiosus* during feedings. This model for animal movement is based on Lagrangian modeling strategies [16] and the work by Brillinger et. al. [2]. In Equation (1), $\mathbf{r}(t) = [X_s(t), Y_s(t)]'$ represents the location of the spider at time t, dt is an incremental change in time, μ is the directional component, Σ is a diffusion parameter, and $\mathbf{W}$ is assumed to be a brownian process. In [11], the diffusion parameter is given by

$$\Sigma\{\mathbf{r}(t), t\} = \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix}$$

where it is assumed $\sigma_x = \sigma_y = \sigma$. Furthermore, the drift parameter is specified by

$$\mu\{\mathbf{r}(t), t\} = v_s \begin{bmatrix} \cos\left(\theta_o(t_i) + \epsilon(t, d_i)\right) \\ \sin\left(\theta_o(t_i) + \epsilon(t, d_i)\right) \end{bmatrix}$$

where v_s is the spider's velocity, $\theta_o(t_i)$ is the optimal direction, i.e., the direction the spider would go if the spider went directly to the prey, and $\epsilon(t, d_i)$ gives the error in direction. When observing *A. studiosus*, it is noted that unlike predators with good vision, spiders obtain their directional cues through vibrations in the web similarly to *Hololena curta* [13] in which prey localization usually requires several pauses where the spider stops and feels for the prey. The term t_i. $i = 1, 2, ..., n$ represents the update time at which the spider changes direction. By substituting the above terms into Equation (1), the stochastic differential equation is given by

$$\begin{bmatrix} dX_s(t) \\ dY_s(t) \end{bmatrix} = v_s \begin{bmatrix} \cos\left(\theta_o(t_i) + \epsilon(t, d_i)\right) \\ \sin\left(\theta_o(t_i) + \epsilon(t, d_i)\right) \end{bmatrix} dt + \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} d\mathbf{W}(t). \tag{2}$$

In [11], Joyner et. al. assumed the prey was similar to a vibrational point source for the spider. The amplitude of vibrations from a point source will dampen as a function of $\frac{1}{d}$ where d is the distance from the source [7]. Given this assumption, $\epsilon(t, d_i)$ in Equation 2 can be modeled from a normal distribution with mean error given by

$$e(d) = \frac{\alpha d}{d + \beta}. \tag{3}$$

Furthermore, although the actual spiders pause several times during prey capture, the model in [11] does not incorporate pauses but instead uses a slower average velocity for the variable v_s.

The model given by Equation (2) with error function $e(d)$ (Eq. (3)) was shown to be a good fit for three of four data sets given in [11] when parameters were estimated to match video data for the time taken for a spider to reach the prey and the distance traveled to the prey. Figure 1 shows the four previous data sets and six new data sets for an individual spider (green), a prey (blue) and five simulations (red) from the model in Equation (2). The stochastic model matched well with spiders 1, 2, and 3 (data from [11]) with relative error less than 6% between the actual spider and the mean of 1000 simulations of the model when comparing time to the prey and total distance traveled (see [11] for more precise results). It was hypothesized in [11] that the fourth video didn't match well, because during the middle of the capture run, the spider took a detour in the wrong direction.

The model in [11] uses the assumption that if the spider gets closer to the prey, the error will *always* decrease with the mean error given by Equation (3). In video 4, the error in the direction increased for the actual spider even though the distance to the prey decreased. When comparing the actual spiders to the simulations in Figure 1, it can be observed that in almost all the cases, the simulated spider takes a more direct path to the prey indicating the error function needs to be modified to account for this possibility of increasing error in direction traveled. In this paper, we modify the error function as well as incorporate varying pauses and varying velocities as suggested by the analysis of the data (see Section 4). Before analyzing the data, the data is processed in order to easily extract the key features necessary for modeling.

3. **Data collection and processing.** The new data sets were obtained from spiders collected near Warriors State Park and Boone Lake in the state of Tennessee. The spiders were then placed into 35x28x7 cm plastic boxes with hair nets acting as a substrate for the spiders to build their web. The spiders were housed in a laboratory setting where the light is set to mimic the 12-hour-day-and-night cycle with one hour transition ramps simulating dusk and dawn. To record the movement of

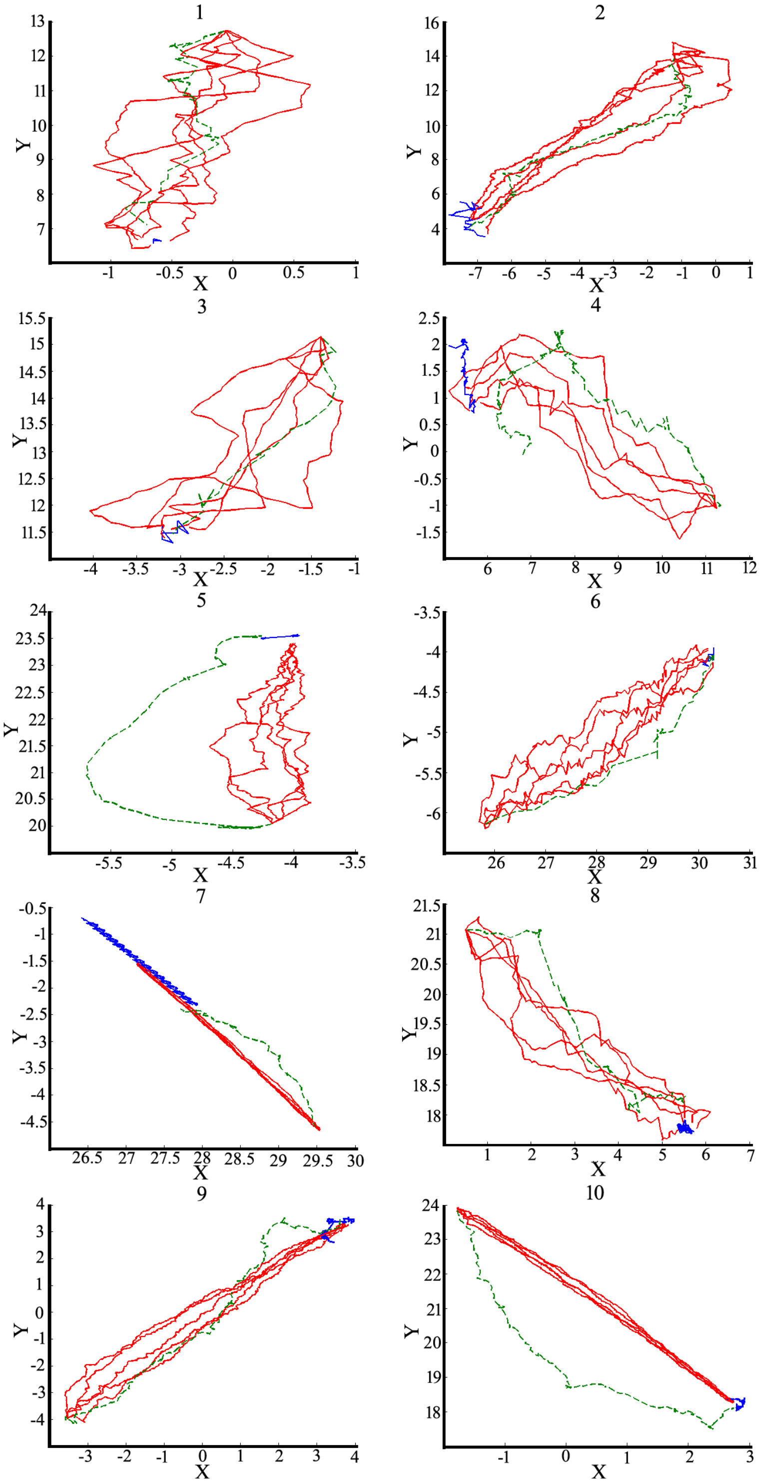

Figure 1. Ten actual spiders (Green dashed line) with five simulated spiders (Red Line) for each spider using Equation 2.

the spider during prey capture, the box was placed on a table with light underneath and a camera placed above the box facing downwards towards the box illustrated in Figure 2. For observations in the dark phase of the cycle, a red filter was used which is invisible to the spiders [5].

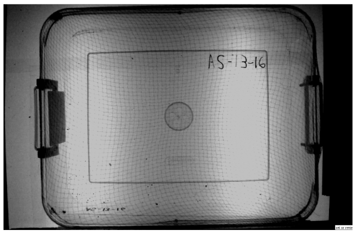

FIGURE 2. Laboratory setup during feedings.

During feeding sessions, the movement of the spider is recorded starting when the prey hits the web until the spider captures the prey. We then use a video tracking software *Tracker* [4] to track and record the path of the spider during prey capture. *Tracker* works by uploading a video into the software and by calibrating an object shown in the video with measurements of that object in real life. An x-y axis is defined onto each video with the origin centered at the lower left corner of the box. *Tracker* has a built-in program that automatically tracks and records the x and y coordinates of the spider and prey with a time stamp starting at $t = 0$. Figure 3 illustrates the path of the spider as defined by the coordinate data. The brownian motion of the spider is detected by the software and depicted in the observed "zig-zag" motion. In addition to brownian motion, the spider's pauses are also depicted by random movement about a point. For instance, if the spider pauses at position $(1, 1)$, the auto tracker randomly chooses a coordinate point (x, y) such that $0.95 \leq x, y \leq 1.05$ at time t for all t during the pause (this is seen in the second plot of Figure 3).

Using the coordinate data, we can extract the distance of the spider from the prey at all times, the velocity of the spider while in motion, how often the spider runs, the duration of a run, how often the spider pauses, duration of pauses, and the average direction in which the spider travels when it runs. To obtain accurate estimates for some of these factors, it is necessary to first cut the data and then smooth the data. We note that although collection of video data began when the prey first hit the web and ended when the spider captured the prey, these spiders exhibit a latency period [14] or a period of time in which the prey is in the web but the spider has not yet responded to the prey. We consider the latency period and a pause to be two distinct behaviors. The latency period is the period of time when the spider first detects the prey and must decide whether to pursue the prey or not. On the other hand, a pause is a period of time in which the spider has already decided to pursue the prey but has stopped motion momentarily (often to feel for the direction of the prey). In this model we are only concerned with the *movement* of the spider during prey capture, and hence, we remove all data prior to the spider's first movement towards the prey thus ignoring any latency period.

In addition, we remove a portion of the data towards the end of the capture run. Technically, a spider successfully captures its prey by either biting the prey or by wrapping the prey with silk [5]. However, the spider tends to circle around the prey

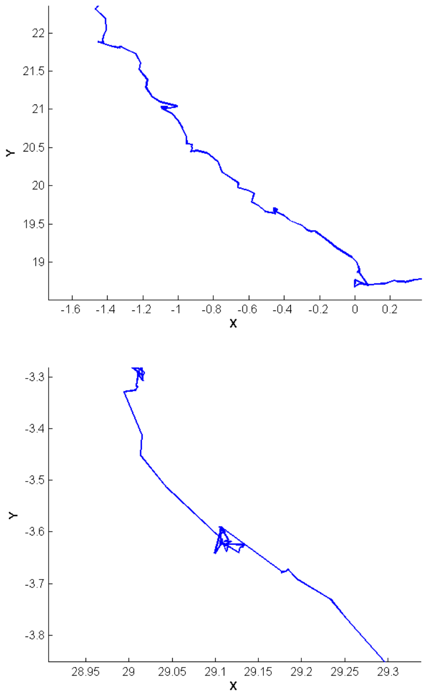

FIGURE 3. Zigzag path and pause illustration respectively.

before biting or wrapping the prey with silk. We speculate this may be a capture strategy, but further work is needed. Once the spider starts circling or gets close enough to the prey, we assume there is no way for the prey to escape and thus would be successfully captured. Therefore, we also remove the data corresponding to the circling period. The edited data then only includes the runs and the pauses in between the runs, ignoring the latency period at the beginning and the circling strategy at the end.

Once the data has been cut, we want to isolate the *general* path of the spider by minimizing the effect of brownian motion and the random positioning caused by the tracking software during pauses. In order to do this, we smooth the data using a *moving average* function with span size k, k odd [8]. Let (x_j, y_j) be the coordinate position of the spider at an arbitrary time t_j, $j = 1, ..., n$. The new value for (x_j, y_j) after smoothing is given by

$$(sx_j, sy_j) = \begin{cases} \left(\frac{1}{k} \sum\limits_{l=j-d}^{j+d} x_l, \frac{1}{k} \sum\limits_{l=j-d}^{j+d} y_l \right), & 1 + d \leq j \leq n - d, \\ \\ (x_j, y_j), & j < 1 + d \text{ or } j > n - d \end{cases}$$

where $d = (k - 1)/2$. The *smoothing function* tends to deflate or lose the realistic movements of the spider if the span is too large [8]. We chose a span of 19 for this data which maintained the major aspects of the data while smoothing out the unwanted "noise". Figure 4 shows the smoothed data; the points on the figure indicate the positions at which the spider paused. In the next section, we discuss the majors factors that influence the trajectory of the spiders during predation using the transformed data and the resulting updated model.

4. **The updated model.** The major factors which must be incorporated in a stochastic model of spider movement during predation are the error in direction, the number and duration of pauses, the velocity and duration of spider runs, and the degree of variance in the brownian motion. Prey localization is determined according to vibrational cues [13] which will also define the directional error in spider movement. The pause serves as an update in the average direction the spider will travel, since the spider pauses, reorients itself according to the vibrational cues and then continues at its run velocity towards the prey [13]. The error in direction, duration and number of pauses, the magnitude of the velocity, and the degree to which the spider exhibits brownian motion should all contribute to the total distance the spider travels to the prey and the total time it takes the spider to cover this distance.

It was previously assumed in the paper by Joyner, et. al. [11] that as the spider gets closer to the prey, the error should always decrease and hence could be modeled by Equation (3). However, simulations for spiders 4-10 in Figure 1 indicate the error may actually increase during the capture run. Initially, if the prey is far enough away, the spider will only feel a dampened vibrational cue and thus there should be larger error than when the spider is close to the prey. As the spider gets closer to the prey, the vibrations from the prey should become stronger and easier to localize which was the assumption made in the previous paper [11]. However, this is an oversimplification, because the vibrations from the spider itself are ignored. We can think of the prey's vibrational cues as source waves and the spider's own vibrations as an interfering wave. The interfering wave can then alter the vibrational cues causing more error in the direction as the spider moves. Therefore, in order to describe the error function, we need a nonlinear function which is greater for longer distances from the prey but should still allow for an increase in error even as the distance gets smaller. We choose a cubic function of the form

$$e(d) = \beta_3 d^3 + \beta_2 d^2 + \beta_1 d + \beta_0 \tag{4}$$

where β_0, β_1, β_2, and β_3 are constants and d is the distance between the spider and prey.

Let $P_1 = (x_P(t_\rho), y_P(t_\rho))$ be the position of the prey at the current pause $t = t_\rho$, $S_1 = (x_S(t_\rho), y_S(t_\rho))$ be the location of the spider at the current pause, and $S_2 = (x_S(t_{\rho+1}), y_S(t_{\rho+1}))$ be the location of the spider at the next pause $t = t_{\rho+1}$. The directional vector from the spider to the prey is denoted $\overrightarrow{S_1P_1} = [x_P(t_\rho) - x_S(t_\rho),\quad y_P(t_\rho) - y_S(t_\rho)]$ (the optimal direction) and $\overrightarrow{S_1S_2} = [x_S(t_{\rho+1}) - x_S(t_\rho),\quad y_S(t_{\rho+1}) - y_S(t_\rho)]$ is the directional vector the spider actually travels from the current pause to the next pause (illustrated in Figure 5). The angle θ_e between these two vectors gives the error in direction and can be calculated ([12]) by

$$\theta_e = \cos^{-1}\left(\frac{\overrightarrow{S_1P_1} \cdot \overrightarrow{S_1S_2}}{\| \overrightarrow{S_1P_1} \| \| \overrightarrow{S_1S_2} \|}\right). \tag{5}$$

Figure 6 displays the values for θ_e for all ten videos at each pause location plotted as a function of the distance d between the spider and prey. The error function $e(d)$ (black line in Figure 6) was then determined by fitting Equation (4) to the data using *Matlab*. The resulting fit yields $\beta_3 = 0.004$, $\beta_2 = -0.069$, $\beta_1 = 0.343$, and

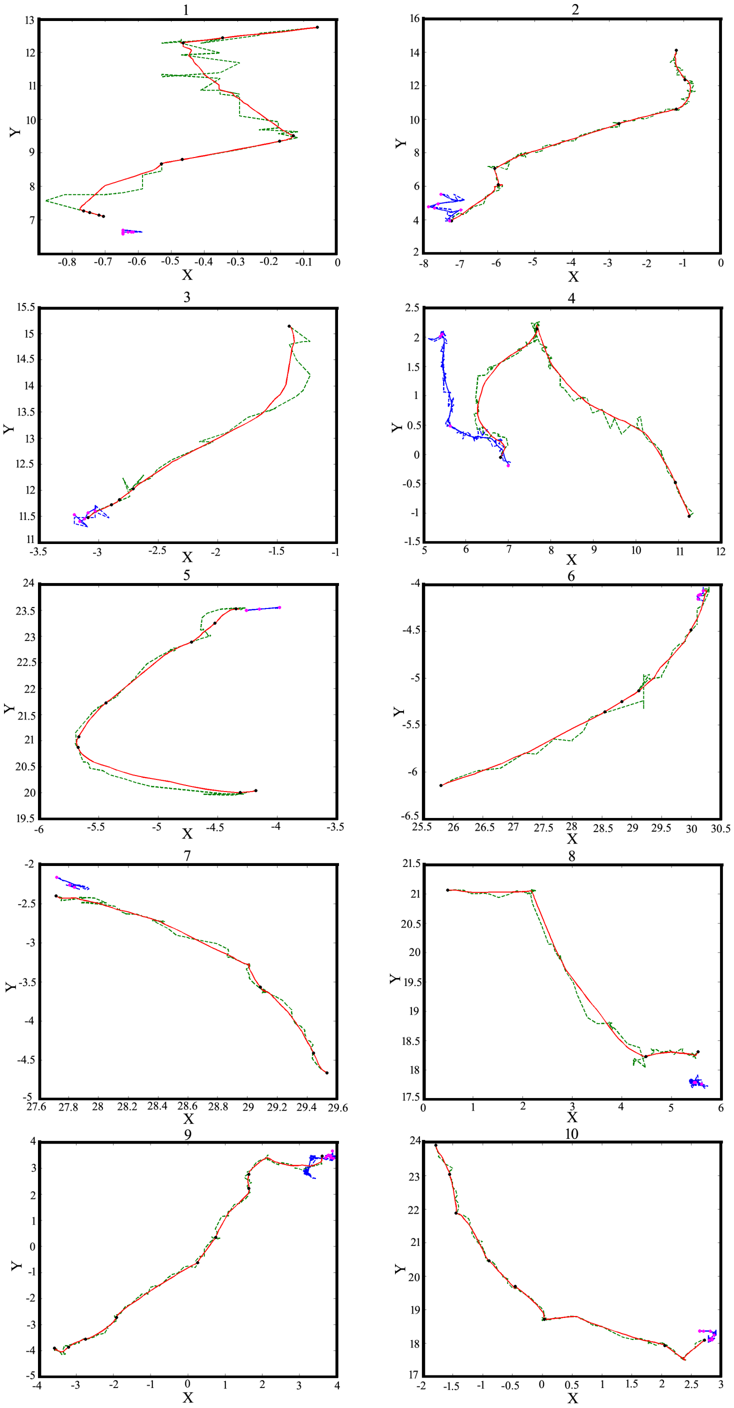

FIGURE 4. This figure shows the actual coordinate data (green dashed line), the path after smoothing (red line) and pause locations (black dots) of each of the ten spiders.

$\beta_0 = -0.074$. The error function will be used in the model to predict the direction in which the simulated spider should travel based on the current distance between the prey and the spider.

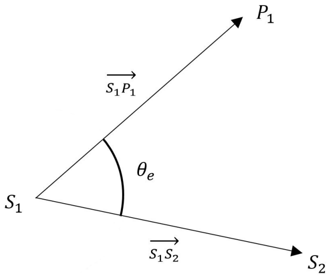

FIGURE 5. The error in direction is given by the angle between vectors $\overrightarrow{S_1P_1}$ and $\overrightarrow{S_1S_2}$.

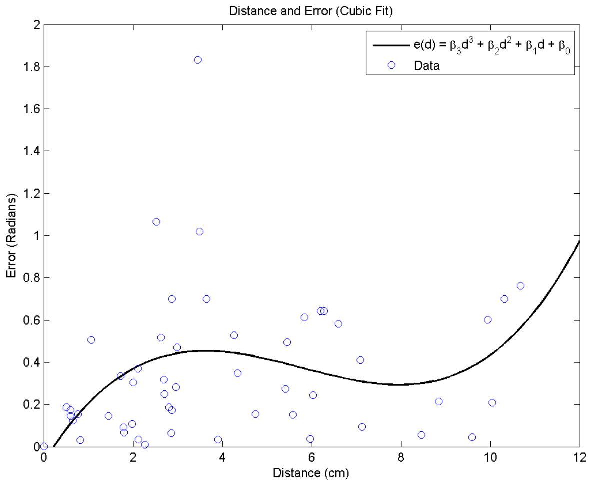

FIGURE 6. Combined directional errors for all spider data with respect to distance d in centimeters with the fitted function $e(d)$, Equation 4.

The new mean error function as given by Equation (4) is a continuously increasing function for $d > 8$ cm. This will cause the error in the direction of the simulated spider to increase indefinitely if the spider happens to go in the wrong direction initially and moves farther from the prey. This is unrealistic; therefore, we limit the mean error to 1 radian for $d > 12$ cm. We assume the error is normally distributed with mean $e(d)$ and standard deviation of q_1 for longer distances, $8\text{cm} \leq d \leq 11\text{cm}$, and a standard deviation of q_2 for shorter distances, $d < 8\text{cm}$.

As mentioned above, pauses allow the spider to calm its own vibrations and redirect itself in the perceived direction of the prey. The duration of a pause will directly affect the total time it takes for the spider to reach the prey. We determine the duration of an individual pause by a simple calculation of $DP = t_{\rho f} - t_{\rho i}$ where $t_{\rho f}$ and $t_{\rho i}$ is the final and initial time of a pause respectively and DP denotes the total duration of the pause. The duration of a pause is incorporated in the stochastic model as a parameter drawn from an exponential distribution with mean q_3.

Each spider has different personality traits [6, 14]. This includes aggression levels which are usually measured by how fast the spider moves during prey capture or the response behavior when stimulated [14]. From casual observation of the *A. studiosus*, these spiders are generally nocturnal; they retreat into the deep parts of the web during the day and come out to their webs during the night. Both aggression and time of day may have a significant effect on a spider's behavior during predation and more specifically on how we model the velocity of a spider during prey capture. We have not fully analyzed all these aspects of velocity; nonetheless, we do see differences in both the velocity of a spider run and the duration of the run across the ten spiders. Figure 7 shows the distribution of the average velocity of all single runs and their durations for all ten spiders. In the model, we let q_4 represent the mean velocity and q_5 the mean duration of the run. The final parameter in the model, q_6, is the term σ in Equation (2) representing the magnitude of brownian diffusion. Figure 8 illustrates the stochastic model algorithm where D_p is the duration of pause, D_r is the duration of run, T_p is the timer for pauses, and T_r is the timer for runs. Note that the variable i is a counter for the pauses and $\theta_o(t_i)$ is the optimal direction discussed in Section 2.

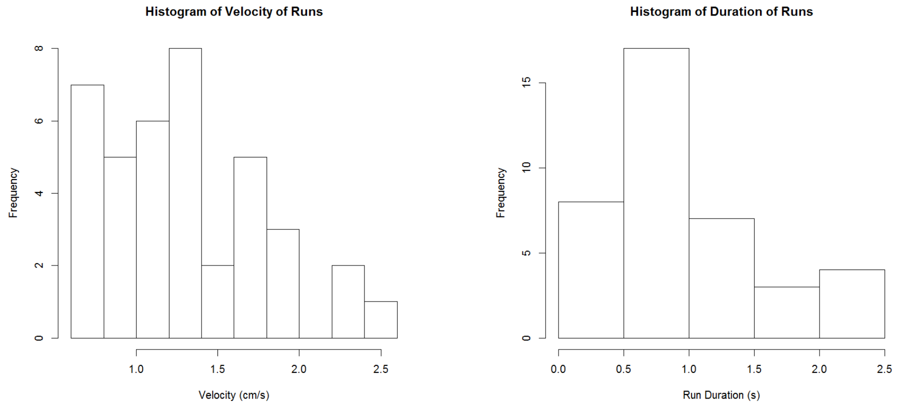

FIGURE 7. Distribution of the velocity of runs in centimeters per second and the duration of runs in seconds for all spider data.

5. **Parameter estimation.** The parameters discussed in Section 4 are estimated using a weighted least squares optimization algorithm in which we try to minimize the difference between the total distance traveled and time elapsed of the simulated spider with that of the actual spider. Assuming each spider is unique, each spider has distinct optimal parameter values $\mathbf{q}^{(k)} = [q_1^{(k)}, q_2^{(k)}, q_3^{(k)}, q_4^{(k)}, q_5^{(k)}, q_6^{(k)}]$ for each spider $k = 1, ..., 10$. For each spider, let $\mu_{ST}(\mathbf{q}^{(k)})$ be the mean simulated time elapsed of 1000 simulated spiders using the stochastic model with parameter $\mathbf{q}^{(k)}$, $\mu_{SD}(\mathbf{q}^{(k)})$ be the mean simulated distance traveled of 1000 simulated spiders, $T^{(k)}$

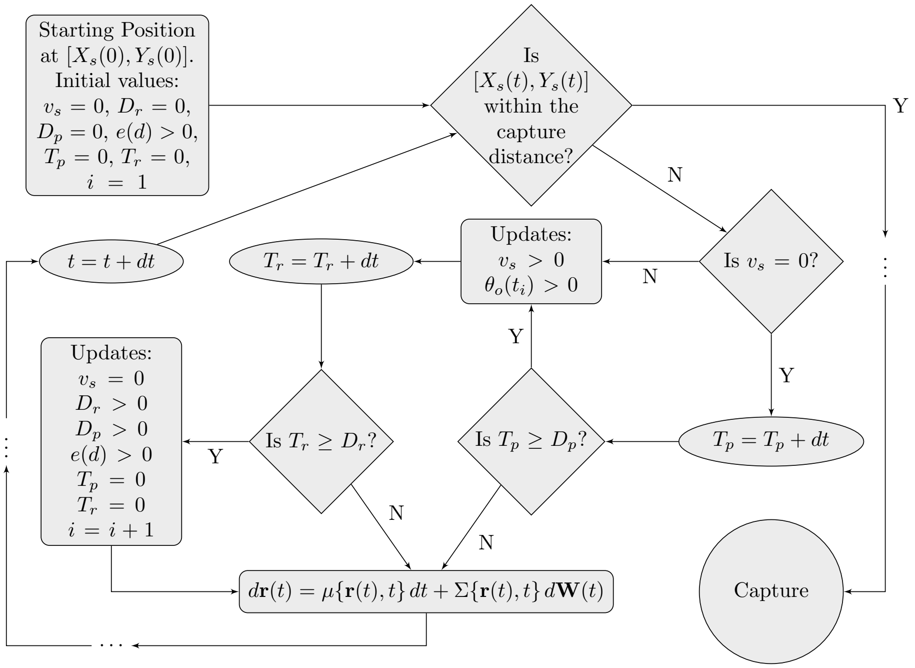

FIGURE 8. The Stochastic Model Algorithm.

be the actual time elapsed, and $D^{(k)}$ be the actual distance traveled. Using the stochastic model together with the algorithm from Section 4, we seek $\mathbf{q}^{(k)}$ which minimizes the cost function

$$J(\mathbf{q}^{(k)}) = \left(\frac{T^{(k)} - \mu_{ST}(\mathbf{q}^{(k)})}{T^{(k)}}\right)^2 + \left(\frac{D^{(k)} - \mu_{SD}(\mathbf{q}^{(k)})}{D^{(k)}}\right)^2. \tag{6}$$

We use the optimization algorithm *fminsearch* in *Matlab* to estimate the parameters for each spider. Table 1 shows the optimal parameters for each spider.

TABLE 1. Six optimal parameters for each spider estimated by minimizing the cost function shown in Equation 6.

Spider	q_1	q_2	q_3	q_4	q_5	q_6
1	1.4006	0.0001	0.0001	1.6557	0.3464	1.9026
2	1.2242	0.7872	0.0000	1.2384	0.2151	0.2345
3	1.0054	0.4882	0.0000	3.0771	0.1804	0.2965
4	0.9292	0.6199	0.0092	2.1581	0.3799	3.3740
5	1.2116	0.9569	0.0001	1.7787	0.2443	0.0116
6	0.9213	0.3765	0.0000	1.3400	0.3452	0.0148
7	1.3015	1.0240	0.0609	0.9626	0.1423	0.2365
8	1.1207	0.7167	0.0000	3.0159	0.1947	0.2665
9	0.6380	0.9371	0.0093	0.7693	0.2866	0.7852
10	1.2003	0.9761	0.0108	0.1591	0.3029	0.2353

Let $\epsilon_T^{(k)}$ be the relative error for time elapsed where

$$\epsilon_T^{(k)} = \frac{\mid T^{(k)} - \mu_{ST}(\mathbf{q}^{(k)}) \mid}{T^{(k)}}.$$

Let $\epsilon_D^{(k)}$ be the relative error for distance traveled where

$$\epsilon_D^{(k)} = \frac{\mid D^{(k)} - \mu_{SD}(\mathbf{q}^{(k)}) \mid}{D^{(k)}}.$$

Table 2 shows the relative errors of each spider comparing the actual spider and the mean of 1000 simulated spider of both time elapsed and distance traveled using 6 optimal parameters (see Table 1). A relative error of less than 0.10 or 10% indicates a good match. We note than in all cases, there is less than 3% relative error and in many cases, there is less than 1% relative error. By fitting each spider individually, the updated model appears to be a good fit for an individual spider; however, future applications on the model require us to be able to determine how a random spider might behave during predation. Therefore, we need a model which will incorporate all the individual dynamics depicted by the differences in parameter values across spiders. Therefore, we first determine which parameter values are approximately constant across all spiders. We then characterize the remaining parameter values according to their probability distributions.

TABLE 2. Goodness of fit of the current model using six optimal parameters.

Spider	T (s)	μ_{ST} (s)	ϵ_T	D (cm)	μ_{SD} (cm)	ϵ_D
1	4.8380	4.8126	0.0053	7.6733	7.6387	0.0045
2	18.7180	18.6933	0.0013	19.9008	19.8070	0.0047
3	2.2690	2.3053	0.0160	5.7477	5.8242	0.0133
4	6.2070	6.1829	0.0039	14.7958	14.6744	0.0082
5	4.5045	4.5472	0.0095	6.6750	6.7627	0.0131
6	9.4000	9.3930	0.0007	6.7534	6.9162	0.0241
7	8.5502	8.5585	0.0010	4.8273	4.8362	0.0018
8	4.1291	4.1899	0.0147	9.9694	10.0232	0.0054
9	24.2325	24.5539	0.0133	16.2675	16.4714	0.0125
10	87.1704	87.2504	0.0009	12.0595	12.0662	0.0006

From Table 1, we note that q_3, the duration of a pause, is relatively close to zero and almost constant across all spiders. Therefore, we choose q_3 to be constant and set it equal to the median value across all spiders (0.0001011). The parameter estimation problem is then repeated with q_3 held constant to obtain new estimates for the 5 remaining parameters. By fixing q_3, we loose one degree of freedom and thus expect the relative errors to increase. Table 4 shows the relative error in total distance traveled and total time elapsed for each spider when comparing the actual spider to the mean of 1000 simulated spiders using the fixed value of q_3 and 5 optimal values for the remaining parameters (see Table 3). The relative errors shown in Table 4 are still less than 0.10 indicating that fixing q_3 caused no significant change in the model fit. Figure 9 shows the actual spider with simulated spiders using the updated model with the individual optimal parameters. Figure 10 shows a distribution of the time elapsed and distance traveled of one thousand simulated spiders (blue) compared with the actual spider (red line). In each case, the actual spider time and distance fall right into the mean of the distribution which again illustrates a good fit when using individual optimal parameters.

TABLE 3. Five optimal parameter values for each spider estimated by minimizing the cost function shown in Equation 6.

Spider	q_1	q_2	q_4	q_5	q_6
1	1.4820	0.0001	1.6532	0.3410	1.9389
2	1.2171	0.7911	1.2369	0.2212	0.2429
3	1.0516	0.4806	3.0990	0.1798	0.3027
4	0.9812	0.6325	2.1612	0.3714	3.3671
5	1.2406	0.9504	1.7754	0.2419	0.0124
6	0.9488	0.3691	1.3288	0.3420	0.0158
7	1.4489	1.0425	0.7481	0.1495	0.2494
8	1.1268	0.7062	3.0180	0.1963	0.2746
9	0.6282	0.9282	0.7587	0.3037	0.8302
10	1.2385	0.9684	0.1584	0.3065	0.2472

TABLE 4. Goodness of fit of the current model using five optimal parameters (constant value for q_3).

Spider	T (s)	μ_{ST} (s)	ϵ_T	D (cm)	μ_{SD} (cm)	ϵ_D
1	4.8380	4.8181	0.0041	7.6733	7.6388	0.0045
2	18.7180	18.8120	0.0050	19.9008	20.0342	0.0067
3	2.2690	2.2731	0.0018	5.7477	5.8011	0.0093
4	6.2070	6.2852	0.0126	14.7958	15.0026	0.0140
5	4.5045	4.4576	0.0104	6.6750	6.5643	0.0166
6	9.4000	9.4754	0.0080	6.7534	6.8588	0.0156
7	8.5502	8.5025	0.0056	4.8273	4.8084	0.0039
8	4.1291	4.1287	0.0001	9.9694	10.0167	0.0047
9	24.2325	24.4133	0.0075	16.2675	16.2849	0.0011
10	87.1704	87.1462	0.0003	12.0595	12.0512	0.0007

6. **Probability distributions for the parameters.** From the six original parameter values, we only considered one parameter value to be constant across all ten spiders. Even though we have a good model fit for the ten spiders in this study, we want to develop a model that incorporates all the aspects found within these ten spiders to potentially capture the predation dynamics of a random spider not in this study. To this end, we seek to develop an aggregate model in which the five variable parameters are drawn from a chosen probability distribution which appropriately captures the variability demonstrated across the ten spiders (see Table 3).

The optimal parameter values for a given spider were determined such that the *joint* parameter values within the model provide a good fit to the data for a particular spider. Before determining appropriate probability distributions for each parameter, we first need to know if any of the parameters are correlated. If two (or more) parameters are correlated, we need to include this correlation in the aggregate model. Figure 11 visually depicts correlation between parameters using a

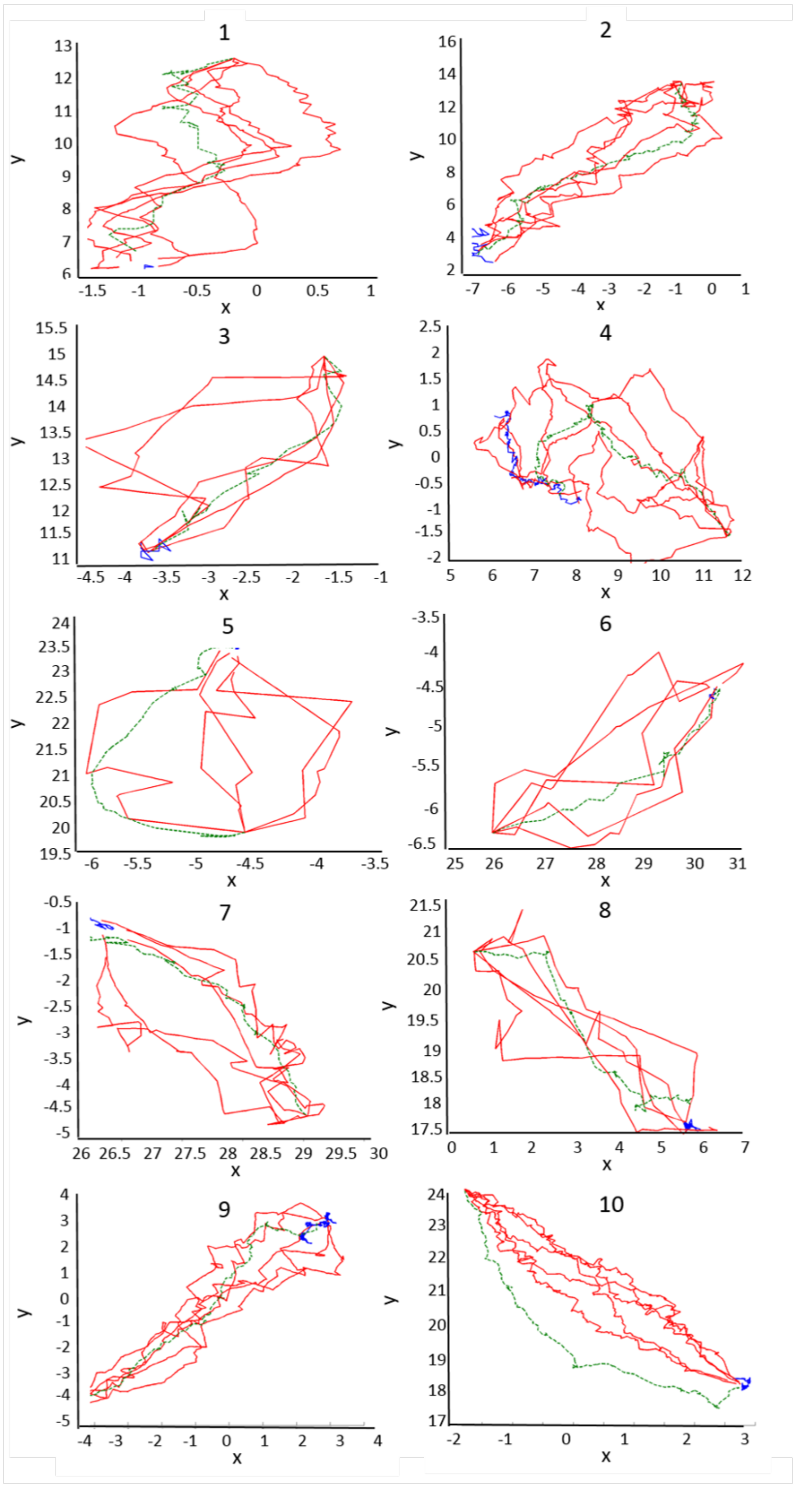

FIGURE 9. Ten actual spiders (green dashed line) with five simulated spiders (red Line) for each spider using Equation 2 with the algorithm.

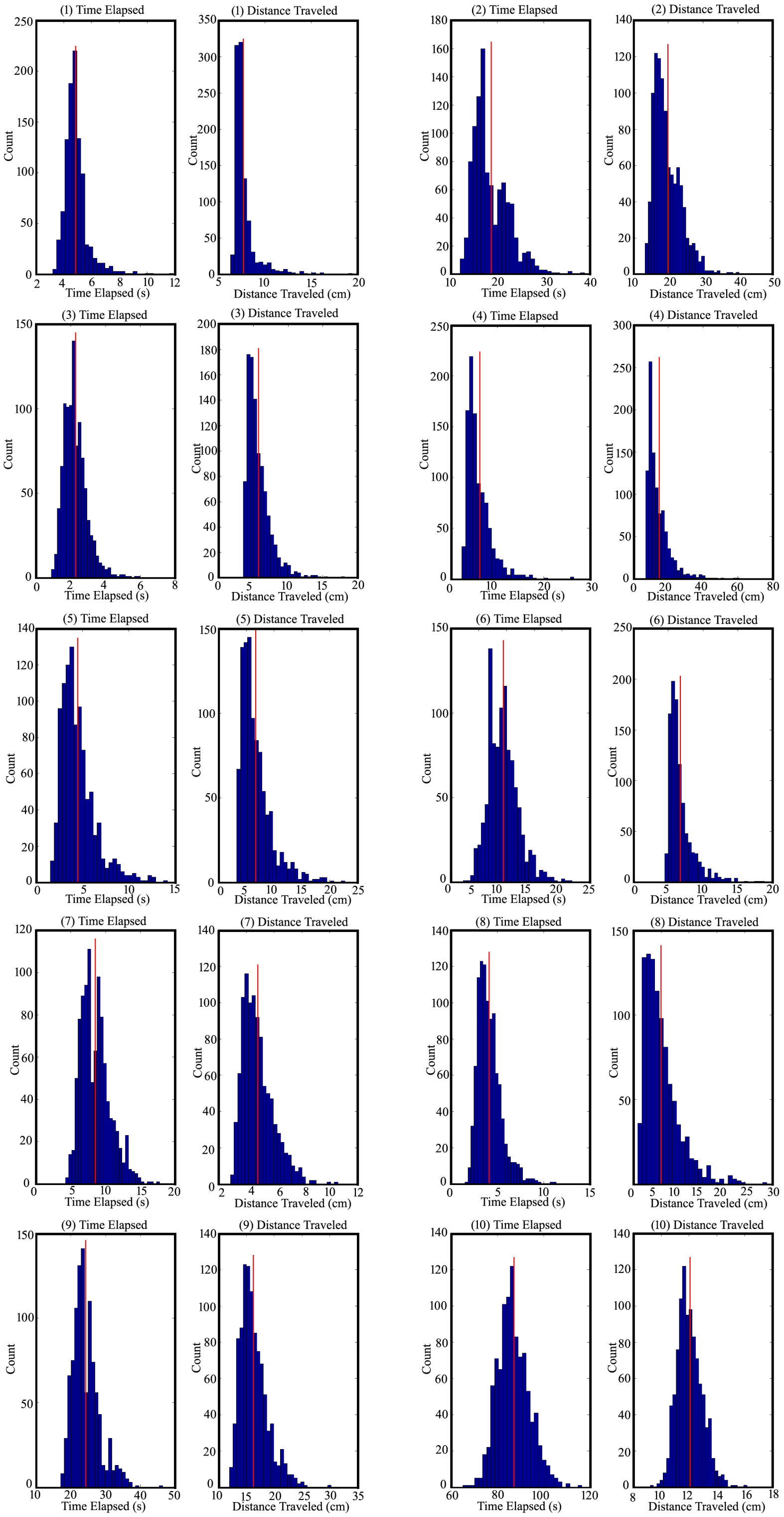

FIGURE 10. The figures show the distribution of the 1000 simulated spiders compared to actual spiders (red vertical line).

scatterplot matrix. If there is no correlation, the data should appear as a random spatial point pattern; if a correlation exists between two parameters, there will be a pattern to the data. The correlation scatterplot matrix suggests a linear relationship between q_1 and q_2 if the two outliers are ignored and a non-linear relationship between q_5 and q_6. The parameter q_4 shows no correlation to the other parameters. The parameters q_1 and q_2 represent the standard deviation in the directional error for long and short distances to the prey respectively, q_4 is the mean velocity during a run, q_5 is the mean duration of a run, and q_6 is the degree of variance in brownian motion. For the parameters in which a correlation exists, we define an appropriate probability distribution for one parameter and then the correlated parameter is defined according to the relationship.

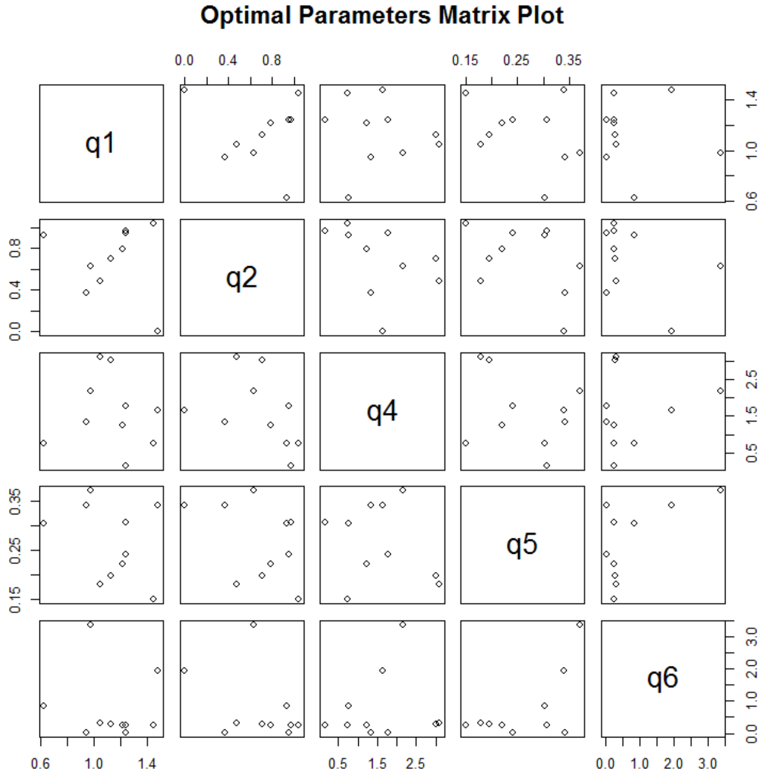

FIGURE 11. Scatter-plot matrix of the optimal parameters.

The scatterplot in Figure 12 suggests a linear relationship between parameters q_1 and q_2 if the outliers are ignored. We note that the outliers result from the optimal parameter estimations for spiders 1 and 9 (see Table 3) although there is no clear indication in the data for why these spiders are outliers. To find a linear model

$$\hat{q_2} = \alpha_0 + \alpha_1 q_1 \tag{7}$$

for q_1 and q_2 where $\hat{q_2}$ is the estimate of q_2, we let q_1 be the explanatory variable and q_2 be the response. We then fit a linear model using R software [15] and find

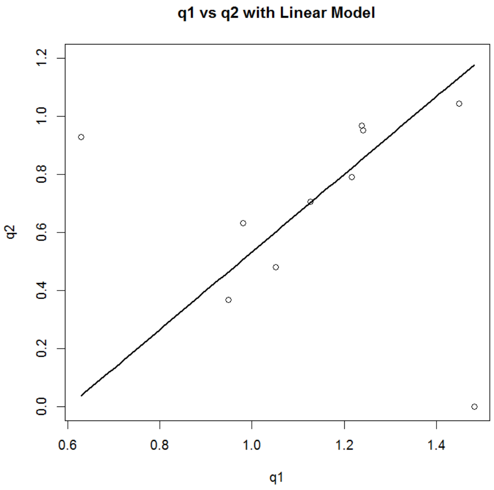

FIGURE 12. Linear model of q_1 and q_2 (Outliers ignored).

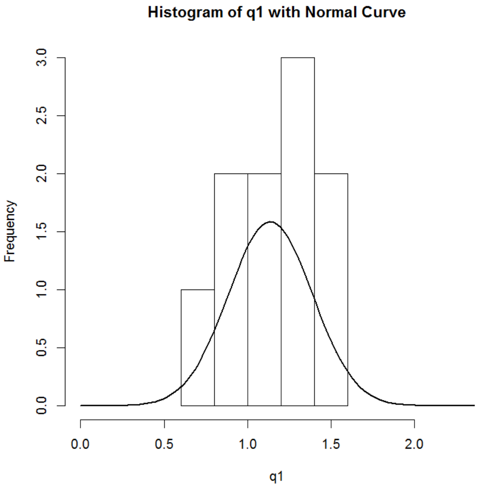

FIGURE 13. Histogram of q_1 with theoretical normal curve.

$\alpha_0 = -0.8023$ and $\alpha_1 = 1.3356$ in Equation (7). The parameter q_1 exhibits a truncated normal distribution as seen in Figure 13 (truncated because q_1 cannot be negative). A mean of 1.1364 and standard deviation of 0.2516 for the distribution are found using the parameter data for q_1 in Table 3. The parameter q_2 is then found from a truncated normal distribution with $\hat{q}_2$ as the mean and using the standard deviation of the absolute value of the residuals of the linear model. Since the parameter q_2 cannot be negative the lower bound of the distribution is 0.

Figure 14 suggests that q_4 exhibits a uniform distribution with lower and upper bounds of 0.1584 and 3.0990. The bounds are taken from the maximum and minimum values of the data in Table 3 for the parameter q_4 across all spiders. Figure

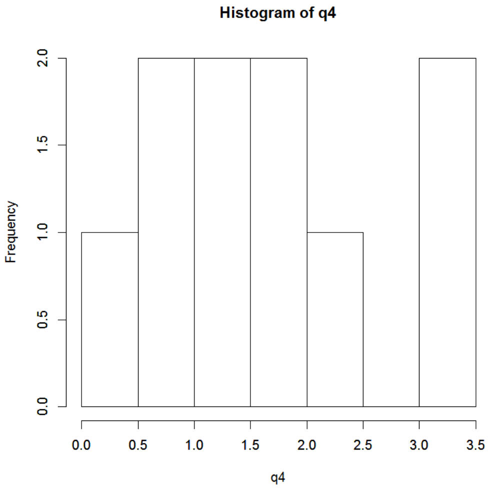

FIGURE 14. Histogram of q_4.

15 suggest a exponentially increasing relationship between q_5 and q_6 with q_5 as the explanatory variable and q_6 as the response. We propose a non-linear model of the form

$$\hat{q_6} = \gamma e^{\lambda q_5} \tag{8}$$

where γ and λ are constant and $\hat{q_6}$ is the q_6 estimate. Fitting the exponential model using R gives us the values $\gamma = 5.244 * 10^{-6}$ and $\lambda = 35.9500$. Figure 16 suggests that q_5 also exhibits a normal distribution, so the parameter q_5 is drawn from a truncated normal distribution with mean 0.2653 and standard deviation of 0.0775. We truncate the normal distribution because a value of q_5 greater than 0.35 gives us extremely large values for q_6 that would result in the stochastic model running indefinitely. We also want to avoid a negative value or extremely low value for q_5 since q_5 is the duration of a run. The mean and the standard deviation is taken from the q_5 parameter data across all spiders from Table 3. We set the truncated distribution to have a lower bound of 0.1584 and upper bound of 0.3714 and given by the maximum and minimum values q_5 across the ten spiders in Table 3. We have already defined q_6 as a function of a given value of q_5 in Equation (8) and hence know that the parameter q_6 behaves according to a log-normal distribution. To generate the specific parameter value q_6 from the distribution which corresponds to the chosen value of q_5, we first choose q_5 as discussed above from a normal distribution and then use q_5 and Equation (8) to define the mean for parameter q_6, namely $\hat{q_6}$. We note that we could stop at this point and use this value for q_6, assuming an exact exponential relationship between q_5 and q_6; however, in examining Figure 15, the values of q_6 appear to be normally distributed about the exponential curve. Therefore, once we determine $\hat{q_6}$ given the current value of q_5, we draw q_6 from a truncated normal distribution with $\hat{q_6}$ as the mean (to have the exponential relationship with q_5) and standard deviation set as the absolute value of the residuals of the exponential model (to account for the variation about the curve shown in Figure 15). Since the parameter q_6 cannot be negative the lower bound of the distribution is 0.

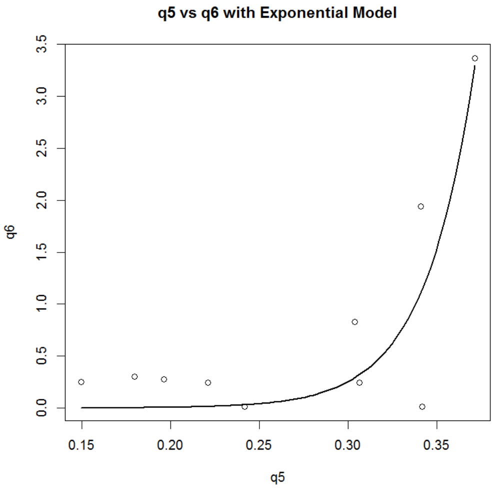

FIGURE 15. Exponential model of q_5 and q_6.

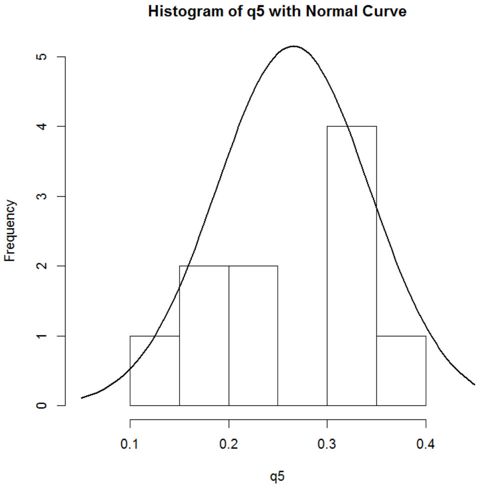

FIGURE 16. Histogram of q_5 with theoretical normal curve.

7. **Simulations and results.** To test the aggregate model, we set a fixed initial location and fixed final destination point 10 cm away. We then determine the mean time elapsed and mean distance traveled for each of the ten spiders to travel from the initial point to the final point using the updated model with the *individual* optimal parameter values in Table 3 using 1000 simulations. We then produce 5000 simulations with the same initial location and fixed final destination using the aggregate model with parameters chosen using the probability models as discussed in Section 6. Figures 17 and 18 give the distribution of the time elapsed and distance traveled for the 5000 simulations of the aggregate model. The red line in each of

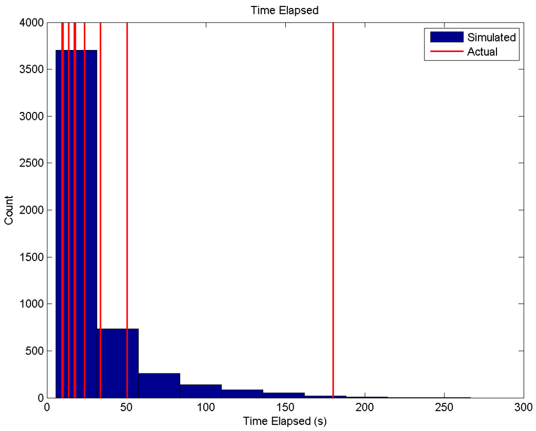

Figure 17. Time elapsed histogram of 5000 simulations with the time elapsed of the simulations using the parameters of actual spiders.

these figures indicates the mean time and mean distance traveled for each of our spiders using their individual optimal values. Each of our individual spiders fall within the distribution of the aggregate model. Furthermore, Figures 17 and 18 show that the majority of the individual spider values lie in the largest area of the distribution suggesting that the aggregate model also captures the proportions of spiders which should lie in each of these intervals. We examine this concept further in Figures 19 and 20 using a bar graph for the proportions of the 10 spiders and 5000 simulations which lie within set intervals of time elapsed and distance traveled. We note that the proportion of the 10 spiders can only increase or decrease in increments of 10% or 0.10 which indicate that the aggregate model also fairly accurately captures the proportions inherent in the data set. Figures 17 - 20 illustrate that the aggregate model successfully incorporated the individual dynamics suggested by the sample data. If we assume the given data sets comprise an appropriate random sample of the behavior of *A. studiosus* during predation, the aggregate model could be used to accurately predict predation strategies.

8. Conclusions and outlook.

In conclusion, we developed a methodology to incorporate individual dynamics of the spider *A. studiosus* during predation into an aggregate model which accurately captures the dynamics of all individuals within our data set. We first refined the model for an individual spider to more accurately capture the error in movement exhibited during the capture run. We then added two important characteristics of a spider run during predation: varying run velocities and pauses. A parameter estimation problem was performed for each individual data set, and then an aggregate model was developed using probability distributions fitted to each parameter across the individuals. The important aspects of the data for this study were the time elapsed during the capture run and the total duration of the capture run. The aggregate model was tested for these two outcomes and compared to the individual models using the optimal parameter set found for each spider. Not only did the probability distributions for the two key outcomes of the aggregate model successfully capture all the individuals, the aggregate model also

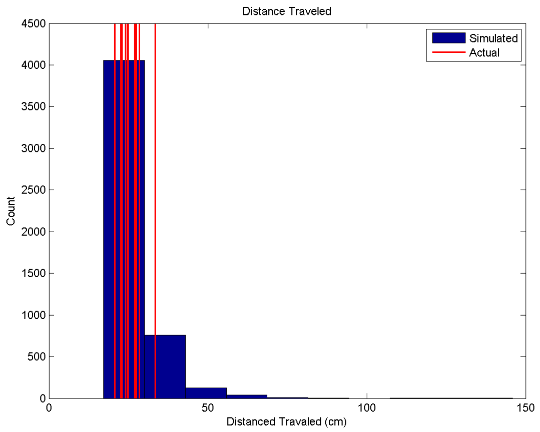

FIGURE 18. Distance traveled histogram of 5000 simulations with the distance traveled of the simulations using the parameters of actual spiders.

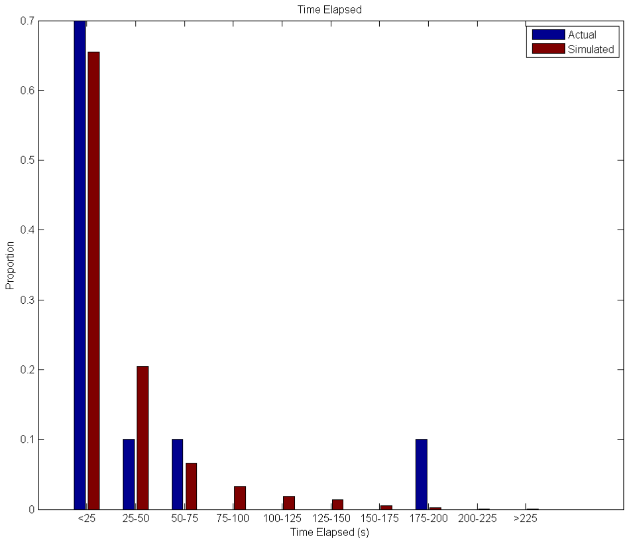

FIGURE 19. The proportions of a set interval of time elapsed of the 5000 simulations compared to the proportions of a set interval of the 10 simulated spiders using the parameters of actual spiders.

was able to fairly accurately capture the proportions of all the data that should lie in fixed intervals of these outcomes. In future work, the aggregate model will be used to test various spacing formations to determine the best spatial point pattern in communal webs during prey capture.

Moreover, the process developed for creating this model can be used in a variety of applications. Whenever one is developing a model in which there are significant

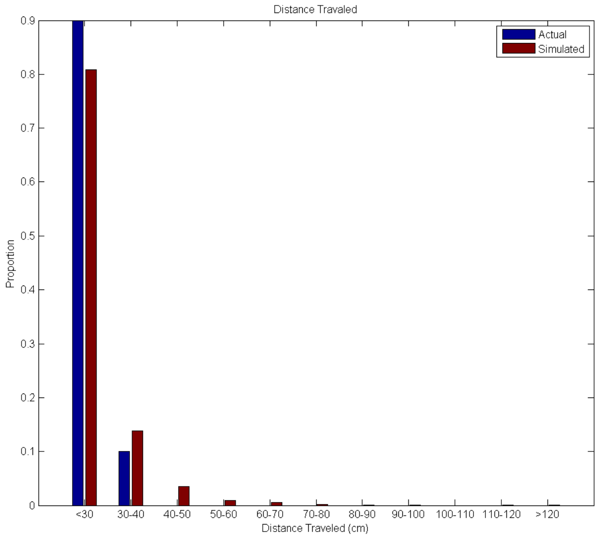

FIGURE 20. The proportions of a set interval of distance traveled of the 5000 simulations compared to the proportions of a set interval of the 10 simulated spiders using the parameters of actual spiders.

differences in individual dynamics, the techniques discussed in this paper can be applied to incorporate the individuality into one encompassing aggregate model. If a representative sample is chosen, the aggregate model could be used to determine the model outcomes for a random individual of the given population.

Acknowledgments. This research was partially supported by the National Science Foundation grant DMS-1128954. We would like to also especially acknowledge Chelsea Ross and J. Colton Watts for all their work on this project.

REFERENCES

[1] H. Banks, S. Hu and W. Clayton, *Modeling and Inverse Problems in the Presence of Uncertainty*, CRC Press, Boca Raton, 2014.

[2] D. R. Billinger, H. K. Preisler, A. A. Ager, J. G. Kie and B. S. Stewart, *Modelling Movements of Free-Ranging Animals*, Technical Report 610, Department of Statistics, University of California, Berkeley, 2001.

[3] M. Davidian and D. M. Giltinan, Nonlinear models for repeated measurement data: An overview and update, *Journal of Agricultural, Biological, and Environmental Statistics*, **8** (2003), 387–419.

[4] B. Douglas, *Tracker: Video Analysis and Modeling Tool*, Tracker version 4.80, Copyright(c), 2013, URL `http://www.cabrillo.edu/~dbrown/tracker`.

[5] R. F. Foelix, *The Biology of Spiders*, 3rd edition, Oxford University Press, 2011.

[6] L. Grinstead, J. N. Pruitt, V. Settepani and T. Bilde, Individual personalities shape task differentiation in a social spider, *Proceedings of the Royal Society B*, **280** (2013).

[7] D. Halliday and R. Resnick, *Fundamentals of Physics*, John Wiley & Sons, Inc., New York, 1988.

[8] P. Hoel and R. Jessen, *Basic Statistics for Business and Economics*, John Wiley & Sons, Inc., New York, 1971.

[9] T. C. Jones and P. G. Parker, Costs and benefits of foraging associated with delayed dispersal in the spider anelosimus studiosus (araneae: Theridiidae), *Journal of Arachnology*, **28** (2000), 61–69.

[10] T. C. Jones and P. G. Parker, Delayed dispersal benefits both mother and offspring in the cooperative spider anelosimus studiosus (araneae: Theridiidae), *Behavioral Ecology*, **13** (2002), 142–148.

[11] M. Joyner, C. Ross, C. Watts and T. Jones, A stochastic simulation model for anelosimus studiosus during prey capture: A case study for determination of optimal spacing, *Mathematical Biosciences and Engineering*, **11** (2014), 1411–1429.

[12] R. Larson and D. Falvo, *Elementary Linear Algebra*, 6th edition, Brooks/Cole, Belmont CA, 2010.

[13] S. A. Naftilan, Transmission of vibrations in funnel and sheet spider webs, *Biological Macromolecules*, **24** (1999), 289–293.

[14] J. N. Pruitt, S. E. Riechert and T. C. Jones, Behavioural syndromes and their fitness consequences in a socially polymorphic spider, anelosimus studiosus, *Animal Behaviour*, **76** (2008), 871–879.

[15] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013, URL `http://www.R-project.org/`.

[16] P. E. Smouse, S. Focardi, P. R. Moorcroft, J. G. Kie, J. D. Forester and J. M. Morales, Stochastic modelling of animal movement, *Phi.l Trans.R. Soc. B.*, **365** (2010), 2201–2211.

Received July 21, 2014; Accepted December 31, 2014.

E-mail address: `quijano@goldmail.etsu.edu`
E-mail address: `joynerm@etsu.edu`
E-mail address: `seier@etsu.edu`
E-mail address: `hancockn@goldmail.etsu.edu`
E-mail address: `largentm@goldmail.etsu.edu`
E-mail address: `jonestc@etsu.edu`