Environment and Natural Resources Research; Vol. 8, No. 4; 2018
ISSN 1927-0488 E-ISSN 1927-0496
Published by Canadian Center of Science and Education

Copula Modeling of Differential Effect of Leaf Species on *Aedes albopictus* Development Time

Chang-Hyun Kim[1], Ephantus J. Muturi[2] & Seung-Hwan Lee[3]

[1] Illinois Natural History Survey, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA

[2] Crop Bioprotection Research Unit, United States Department of Agriculture, Peoria, IL 61604, USA

[3] Department of Mathematics, Illinois Wesleyan University, Bloomington, IL 61701, USA

Correspondence: Seung-Hwan Lee, Department of Mathematics, Illinois Wesleyan University, Bloomington, IL 61701, USA. E-mail: slee2@iwu.edu

Received: September 7, 2018 Accepted: September 25, 2018 Online Published: October 11, 2018

doi:10.5539/enrr.v8n4p1 URL: https://doi.org/10.5539/enrr.v8n4p1

Abstract

Decaying leaves provide a major carbon source for mosquito larvae and leaf litter of different plant species vary in their ability to support mosquito growth, survival, and development. Thus analyzing the effects of leaf species treatment on development time of *Aedes albopictus* has the potential to discover a plant-based strategy for mosquito control. Here, we employ a statistical model named copula that provides a convenient methodology for modeling multivariate dependence to determine the association between leaf litter identify and mosquito performance. A copula that best fits the association of leaf litters on mosquito performance is selected, and statistical tests are performed to check the adequacy of the copula chosen. By computer-based Monte Carlo methods, a large number of simulated development times are generated under the copula chosen. From the simulated development times, we calculate the percentiles to determine expected development time of female *Aedes albopictus* under the five different leaf species treatments, and compared the results to those when all the effects of leaf infusion are combined.

Keywords: *Aedes albopictus*, Copula, Detritus, Larval development, Tail dependence.

1. Introduction

Aedes albopictus is one of the most effective disease vectors of dengue, chikungunya and Zika viruses, which used to be concerns of the Old World, but now the public health issues of the New World (Lounibos, 2002, Juliano & Lounibos, 2005, Mlakar et al., 2016). *Aedes albopictus* was first introduced to the U.S. in early 1980s via used tire trade from Asia and commercial import of 'lucky bamboo' from China. Within a decade, *Ae. albopictus* had colonized most of southern states of U.S. and today, it occurs in at least 26 states and continues to expand northward and westward (O'Meara et al., 1995; Moore & Mitchell, 1997; Moore, 1999; Yee, 2016).

While the role of most environmental factors on the spread of *Ae. albopictus* remain elusive, studies have shown that several factors such as food availability, temperature, intra- and inter-specific competition, desiccation and predation during aquatic larval stage may affect the characteristics of resulting adult mosquitoes suggesting their pivotal roles on the spread of *Ae. albopictus* (Alto & Juliano, 2001, Alto et al., 2005, Muturi et al., 2010, Muturi & Alto, 2011, Muturi et al., 2011a). However, the effects of xenobiotics from natural sources including arborescent leaf detritus on mosquito development have received little attention despite several studies that have shown that leaf litter identity may influence mosquito performance and community structure (David et al., 2000, Murrell & Juliano, 2008, Pavela, 2008).

Small sample size and consequent overgeneralization of the experimental findings to explain the phenomena occurring in nature are some of the limitations associated with experimental studies to investigate the influence of leaf litter on mosquito performance. In addition, the effects of various leaf litter on mosquito are often complex and their interpretation is quite difficult in case multiple leaf species are involved. To tackle these problems, we utilize a statistical model that controls multiple leaf species concurrently. The statistical model serves as a major tool for drawing inferences about mosquitoes behavior under the effects of various leaf detritus. Due to ecological and medical importance, we exploit *Ae. albopictus* and its development time under various leaf litter treatments as a representative of the procedures. Mosquito development time is one of the important elements that affect mosquito ecology (Muturi et al., 2011b). Shorter development time for a mosquito species is usually advantageous as it increases the probability of completing full development processes under limited resources in a habitat. Even in the event of toxic chemical contamination, shorter development time may reduce the duration of exposure, which would consequently increase survivorship. Conversely, longer development time may be beneficial to surviving mosquito under limited food sources as mortality of less tolerant individuals may release them from competition (Barrera, 1996). Also, having long development time may increase a chance of getting accidental addition of food sources such as falling leaves and dead insects (Barrera, 1996). Less commonly, in the event of having chemical contaminants of which toxicity diminishes over time, mosquito species with longer development time is advantageous over others, too.

In the analysis of multivariate data, a commonly used measure of dependence structure is the linear correlation. However the linear correlation cannot capture a non-linear dependence relationship that may exist in the tail regions of the joint distribution (Embrechts et al., 2002). For our data where changes in the tail behavior are important, we use an alternative statistical model named copula (Sklar, 1959) which has much recent attention in the context of modeling multivariate data. A copula provides a convenient way to express multivariate distributions of two or more variables by linking a multivariate distribution and its one-dimensional (univariate) marginal distributions on which the notion of copula is based. In contrast with the linear correlation, a copula captures a wide range of dependence structures that include non-linear association between variables, accounting for how the marginal distributions of data relate to each other. We use copula models to delineate interactions among the development times under several different types of leaf species treatment and to analyze their behaviors during their development.

The most commonly used copulas are elliptical copulas. Elliptical copulas are simply the copulas of elliptical distributions such as Gaussian and Student's t-distributions (Embrecht et al., 2003; Demnarta & McNeil, 2005). The key advantages of elliptical copulas are that they are suitable in modeling symmetric tail dependence structures with multi-dimensions and specify different levels of correlations between the marginal distribution functions.

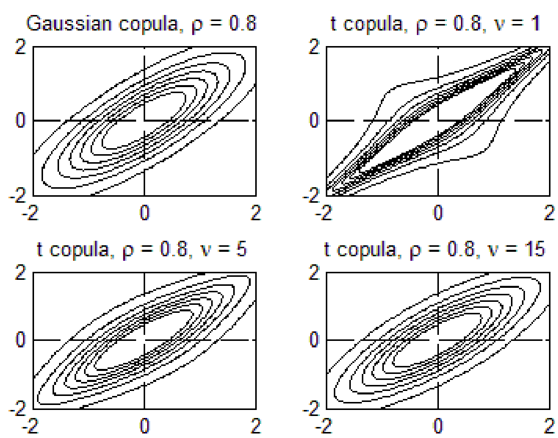

Figure 1. Contour plots of Gaussian and Student's t-copulas, $\rho = 0.8$

Variation that may lead to a misguided interpretation of data occurs when we use a small sample instead of the entire population. The goals of this paper are to create a multivariate distribution for the development time using an elliptical copula, to generate a large number of realizations from the distribution using Monte Carlo simulation, and finally, to better understand the effect of the leaf litter treatments on the development time by looking at the large simulated data points instead of small observations that may cause a variation in results. Incidentally, we discuss how to select an appropriate copula for our data. As a tool in multivariate modeling, a copula has been successfully applied in many areas including survival analysis (Zheng & Klein, 1995), risk management and financial applications (Breymann et al., 2003; Embrechts et al., 1999, 2003), all of which heavily focus on the extreme value analysis. In this paper, we demonstrate the benefit of using a copula to study *Ae. albopictus* development time under various leaf litter treatments, where extreme values of interest could occur in the tails of the distribution. For comparison, we also include the case when all the effects of leaf infusion are combined.

This work is organized as follows. Copula and its dependence structure are discussed in Section 2. In Section 3, we describe data, choose a copula that well-explains data association in the tails of the joint distribution, and check the adequacy of the copula chosen. Section 4 presents analysis and interpretation of the development times obtained from the chosen copula. Section 5 concludes this paper.

2. Methodology: Copula

2.1 Copulas

A random variable associates a numerical value with each outcome of an experiment, determined from some characteristic pertaining to the outcome. Throughout the paper it is assumed that random variables and its possible values (realizations) are denoted by the capital and lower-case letters, respectively. If $X_1,...,X_n$ are random variables, the multivariate distribution function is

$$F(x_1,...,x_n) = P(X_1 \le x_1,....,X_n \le x_n),$$

where $P(A)$ is the probability of A. This multivariate distribution completely describes the dependence between the random variables.

A copula is a function that combines univariate marginal distributions to construct a multivariate distribution with a specific dependence structure (Sklar, 1959). So, it provides a convenient way to construct the multivariate distribution functions of two or more random variables. The foundation of the use of a copula is based on Sklar's theorem (Sklar, 1959; Scheweizer & Sklar, 1983). Sklar's theorem essentially states that if F is a multivariate distribution function with marginal distribution functions $F_1,...,F_n$, then there is a copula C such that

$$F(x_1,...,x_n) = C(F_1(x_1),...,F_n(x_n)), \tag{1}$$

Where $F_i(x_i)$, $i = 1,...,n,$ are uniformly distributed on [0,1]. We can also rewrite (1) for u_i, $i = 1,...,n,$ to be uniformly distributed on [0,1], as

$$C(u_1,...,u_n) = F(F_1^{-1}(u_1),...,F_n^{-1}(u_n)), \tag{2}$$

where $F_1^{-1},...,F_n^{-1}$ denote the quantile functions of the marginal distributions $F_1,...,F_n$. The expressions (1) and (2) indicate that any choice of marginal distributions can be used in copula modeling, which is an important feature of a copula. It is also worth noting that the dependence structure depends on the type of copula and not on the choice of marginal distribution. This is most easily understood in terms of the probability density function. Letting f be the density function of a multivariate distribution F and c the density function of a copula C, Sklar's theorem postulates

$$f(x_1,...,x_n) = c(u_1,...,u_n) \times \prod_i f_i(x_i),$$

where f_i, $i = 1,...,n,$ are marginal density functions of X_i, $i = 1,...,n$. This expression indicates that a multivariate probability density function $f(x_1,....,x_n)$ can be split into the univariate marginal probability density functions $f_i(x_i)$'s and the copula density function $c(u_1,...,u_n)$ that determines a dependence structure. Hence, it is possible to separately specify standalone distribution functions and the dependence structure determined by the copula, which is the key advantage of using a copula. We demonstrate several different copulas that lead to different dependence structures once the marginal distributions are determined, but the resulting multivariate distributions have the same marginal distributions. Some examples of copulas considered in this work are presented in Section 2.2.

2.2 Elliptical Copulas

Elliptical copulas are simply the copulas of elliptical distributions such as Gaussian and Student's t-distributions, so it is useful when we look at the symmetry of data (Embrecht et al., 2003; Demnarta and McNeil, 2005; Nelson, 2006). Elliptical copulas are flexible in the sense that they account for co-movement of the random variables in the tails when modeling multi-dimensional dependence structures. This implies that the elliptical copulas capture differences in pairwise dependence structures. Gaussian copula and Student's t-copula are the most commonly used elliptical copulas for analysis of the tail dependence of data. Student's t-copula takes into account joint extreme events in both tails. Unlike Student's t-copula, a Gaussian copula does not allow for extreme events to be dependent, and thus underestimates the possible effects of extreme events in the tails.

Gaussian copula

Gaussian copula is the copula of multivariate normal distribution. For $X_1,...,X_n$, from (2) it is given by

$$C_G(u_1,...,u_n) = \Phi_\rho(\Phi^{-1}(u_1),...,\Phi^{-1}(u_n)) ,$$

where Φ_ρ denotes the standard multivariate normal distribution function with correlation coefficient matrix ρ and Φ^{-1} the inverse of the standard univariate normal distribution function. In the case of imperfect linear dependence, $-1 < \rho < 1$. As the correlation (dependence) parameter approaches -1 and 1, Gaussian copula captures stronger positive and negative linear relationship, respectively, between random variables. The multivariate normal distribution has thin tailed margins, so Gaussian copula is useful when the correlated data rarely appear in the tail regions of the joint distribution. In the presence of extreme values that lead to a heavy-tailed distribution, we need to consider a copula that captures dependence between variables in the tails of the distribution. It is Student's t-copula.

Student's t- copula

Student's t-copula is based on the multivariate Student's t-distribution. For $X_1,...,X_n$, from (2) it is given by

$$C_t(u_1,...,u_n) = t_{\rho,\nu}(t_\nu^{-1}(u_1),...,t_\nu^{-1}(u_n)) ,$$

where $t_{\rho,\nu}$ denotes the multivariate Student's t distribution function with correlation coefficient matrix ρ and the degrees of freedom ν, and t_ν^{-1} is the inverse of the standard univariate Student's t distribution function. Compared with Gaussian copula, Student's t-copula includes an additional parameter, namely degrees-of-freedom. This is used to model the tendency for extreme events, located in the tails of the distribution, to jointly occur. The degrees of freedom parameter controls the heaviness of the tails of the distribution. For example, increasing value of the degrees of freedom decreases the tendency to exhibit co-movements of variables in the tails of the distribution. This implies its ability to incorporate tail dependence of variables.

3. Data and Modeling

3.1 Data

Leaf-Infusion Water

The leaf species used were black alder (*Alnus glutinusa*), black walnut (*Juglans nigra*), common bald cypress (*Taxodium distichum*), eastern white pine (*Pinus strobes*) and sugar maple (*Acer saccharum*). The leaf species were chosen because they have a widespread distributed in North America and thus most likely to affect aquatic mosquito habitats in nature. The effect of leaf species on larval mosquito development was investigated using leaf-infused water. Defoliated leaves were collected from parks in Champaign County, IL and transported to laboratory for sorting and identification. The leaf identification was carried out according to leaf morphological keys (Tekiela 2006). Bald cypress leaves were a gift from Dr. Barry Alto and were collected and transported from Florida. The leaves were sorted and ground into fine powder in order to synchronize the rate of leaf decomposition. The infusion was prepared by mixing 10 L deionized water and 20 g leaf powder in an 18.9 L plastic container with a lid and fermenting at room temperature for 8 days. The infusion water was filtered using cheese cloth to remove large leaf debris prior to adding mosquito larvae.

Mosquito development

The experiments were conducted using F_8 generation of *Ae. albopictus* originally from Florida. Microcosms consisted of 360 ml of leaf infusion in 400 ml tri-pour beakers and 40 first instar larvae (~12h old) were added to each microcosm. No additional food was provided. The microcosms were maintained at 25°C, 14:10 h light-dark cycle and 70% relative humidity. Every 10 days, the tri-pour beakers were supplemented with 0.25 g of leaf powder until all larvae in the container had either pupated or died. The microcosms were monitored for pupation every 24 hours. Pupae were collected and placed individually in a cotton sealed vial containing distilled water until eclosion. The species, sex and date of adult emergence was recorded for each individual mosquito. The experiments were replicated 3 times.

Mosquito development time is defined as the length of time in days from egg hatching to adult emergence. Designating the date for egg hatching as day 0, the development time was calculated by subtracting the date of egg hatching from the date of adult emergence. Only individual female *Ae. albopictus* which successfully emerged into live adults were included in the data while individuals which died at pupal stage or during eclosion were excluded.

Because different leaf litter showed different level of inhibition on mosquito larval development, the number of emerged adult mosquitoes varied depending on the type of leaf detritus infusion. Among 5 different leaf species tested in the study, sugar maple infusion produced 28 female *Ae. albopictus* from triplicate experiment runs. In order to obtain the same sample size across the different leaf species, the data points collected from 3 replicates within a leaf litter treatment were combined and then 28 data points were randomly selected using computer generated random integer that was pre-assigned for each data point.

We paid special interest in both upper (longer development time) and lower (shorter development time) tails of distribution of *Ae. albopictus* development time under various leaf litter treatments. Shorter development time for a mosquito species under xenobiotic chemical contaminants with inhibitory effects is usually beneficial to surviving mosquitoes as it may reduce the duration of larval exposure to toxic chemicals and the risk of desiccation (Muturi et al. 2011b). Conversely, longer development time under xenobiotic inhibitory chemicals may increase the chances of survival because mortality of individuals that are not tolerant to the chemical contaminants may release tolerant individuals from competition on shared resources (Muturi et al. 2011b). Also, mosquito larvae that are tolerant to chemical contaminants may persist long enough until the effect of chemicals diminishes by time.

3.2 Marginal Distribution

For development time data described in Sections 3.1, appropriate marginal distributions that are plugged in the copula function can be found via graphical and numerical methods. Let X_i $i = 1,...,n$, be a random sample of size n and x_i , $i = 1,...,n$, be the realizations of X_i 's that have the distribution function $F(x)$. This $F(x)$ is a theoretical quantity which is estimated by the empirical distribution function, defined as

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \le x) , \quad (3)$$

where $I(x_i \le x)$ is the indicator function which is defined to be 1 when $x_i \le x$ holds, and equals 0 otherwise. The (theoretical) distribution function gives the probability that a random variable is less than a given value. The empirical distribution function is quite similar, the difference being that the empirical distribution is computed by data, not the theoretical distribution function that describes the parent distribution. So, it is expected that the empirical distribution function resembles the distribution function that fits data. Consequently, if the theoretical distribution function, F , appropriately describes data, the plot of $\hat{F}$ versus F will be close to each other. For example, Figure 2 (left) shows that the empirical distribution for the development time of female *Ae. albopictus* emerged from black alder infusion treatment is fairly close to the logistic distribution. Another graphical method is called the Probability-Probability (P-P) plot which will be approximately linear if the assumed model is correct. Figure 2 (right) displays that there appears to be no significant linear departure of the graph points from the straight line that is the reference diagonal line. From both graphical methods, it seems reasonable that the logistic distribution fits the data well.

There are numerical model checking methods that lead to more objective ways to examine if the distribution function chosen is an appropriate distributional model for data. We utilize the Kolmogorov-Smirnov and Anderson-Darling tests. Both are tests based on the empirical distribution function. Specifically, the Kolmogorov-Smirnov test uses maximum difference between the empirical distribution and the theoretical distribution. The Anderson-Darling test is more sensitive in the tails of the distribution, using the average weighted squared distance between the empirical distribution and the theoretical distribution. An optimal marginal distribution is one that minimizes these distance functions.

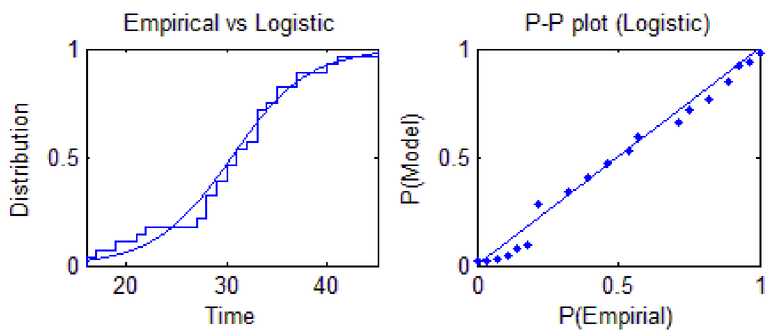

Figure 2. Empirical distribution (left) and P-P plot (right), Black Alder

Judging from the two criteria above, along with the graphical methods, we arrive at the distributions summarized in Table 1 that may best fit the data. Each distribution has the numerical descriptive measures called parameters that completely describe location and shape of a parent distribution. Table 1 also shows estimates of the distribution parameters obtained by the maximum likelihood method.

Table 1. Distribution and descriptive statistics for the development time of *Ae. albopictus* under various leaf infusion treatments

Leaf Species	Distribution	Parameter Estimation	Mean	Std. Deviation
Black Alder	Logistic	σ=3.8241, μ=30.5	30.5	6.9362
Black Walnut	Normal	σ=6.3482, μ=31.679	31.679	6.3482
Common Bald Cypress	Gumbel (Min)	σ=4.304, μ=27.27	24.786	5.52
Eastern White Pine	Gumbel (Min)	σ=10.489, μ=52.733	46.679	13.452
Sugar Maple	Gamma	α=42.283, β=1.1327	47.893	1.3919

3.3 Modeling

We use copula-based multivariate models to analyze development time of *Ae. albopictus* under 5 different leaf species treatments. Because inadequate models result in underestimating or overestimating the true development significantly, it is important to choose an appropriate copula model that will lead to reliable multivariate outcomes. This section searches a copula that may best fit the data and tests the null hypothesis that the copula chosen is statistically significant. The procedures are based on the empirical copula introduced by Deheuvels (1978).

The empirical distribution computed by data estimates the unknown parent distribution of the data points in the sample. Similar to the empirical distribution that approximates an unknown parent distribution, the empirical copula, denoted by $\hat{C}$, of the sample estimates the theoretical (or equivalently, parametric) copula. It is in fact a transformation of the empirical distribution, defined as, for a pair of data (x, y),

$$\hat{C}(x, y) = \hat{H}(\hat{F}(x), \hat{G}(y)), \tag{4}$$

where $\hat{H}$, $\hat{F}$ and $\hat{G}$ are defined in the same manner as (3). The best copula is chosen in such a way that the distance of the empirical copula $\hat{C}$ and the theoretical copula C is minimized. Since some values of the distance are positive and some negative, in practice, the squared distance of them is considered, as will be seen in (5). From the procedures, Student's copula with 10 degrees of freedom was chosen as the one that best fits the association of the development time of *Ae. albopictus* under various leaf species treatments considered.

Although Student's t-copula with ν =10 fits best among others, there is no guarantee that it actually fits the data in an absolute sense. We need to check whether the copula represents the true dependence structure. To this end, it is essential to test the null hypothesis that the copula chosen is statistically appropriate versus the alternative hypothesis that the copula does not actually fit data. The procedures are based on a statistic calculated from a sample. Its value is used to decide whether or not the null hypothesis is rejected in the hypothesis testing. Similar to Genest and Rémillard (2008) and Genest et al. (2009) that use the distance of the copula and its estimated version called the empirical copula, consider the Cramér-von Mises type statistic defined as

$$T = n\int_0^1 \left\{\hat{C} - C\right\}^2 d\,\hat{C}, \tag{5}$$

where $\hat{C}$ is the empirical copula defined in (4) and C is the assumed theoretical copula. If the copula C appropriately describes data, the behaviors of $\hat{C}$ obtained from a sample and C specified theoretically will be close to each other. Therefore, the statistic T can be used to assess the adequacy of the assumed copula. A small value of T indicates that the assumed copula is adequate for a given data set. On the other hand a large value of T shows that the specified copula is not appropriate for the data.

The statistic is converted to a probability called a $p-$value that measures the strength of evidence in support of the null hypothesis. Let t be a value as a result of T in (5). Then, the $p-$value is the conditional probability of observing a statistic as extreme as t, the observation of T, assuming that the null hypothesis is true. A small $p-$value provides evidence that the null hypothesis is rejected because the probability in terms of $p-$value says the observed data are unlikely when the null hypothesis is true. A large $p-$value provides evidence that the null hypothesis is retained. We use this $p-$value as a numerical measure of how well Student's t-copula with 10 degrees of freedom fits data. For the comparative study in the assessment, we also include Gaussian copula that does not have tail dependence and Student's t-copula with 20 degrees of freedom. The use of a Gaussian copula will observe the consequence of ignoring extreme events. Student's t-copula with 20 degrees of freedom have relatively small amount of the association in the tails of the distribution, when compared to Student's t-copula with 10 degrees of freedom.

As with Lee and Yang (2007), the $p-$values, defined as $P(T \geq t)$, are approximated through Monte Carlo method such as the Bootstrap (Efron and Tibshirani, 1993). Again, note that lower the $p-$value, the less likely the goodness of fit is. Based on the re-sampling procedures, the estimated $p-$values for Student's copulas with 10 and 20 degrees of freedom are 0.6350 and 0.3255, respectively. Gaussian copula yields the $p-$value of 0.1742. The results indicate that the three copulas considered are appropriate models for the data, and the outcomes based on the copulas are reliable. Especially, Student's t-copula with 10 degrees of freedom has the largest $p-$value, among others. That is, Gaussian copula and Student's t-copula with 20 degrees of freedom are less correctly specified, so its performance is inferior to Student's copula with 10 degrees of freedom. The $p-$values also imply that actual extreme events in the data could happen more often than forecasted by Gaussian copula.

4. Results

4.1 Dependence Structure

We first present how data are simulated by re-sampling according to copula. From a large number of simulated data points, inferences and analysis on the development time of *Ae. albopictus* can become more systematic, reliable and accurate. The following algorithm outlines the Monte Carlo simulation scheme that generates random variables, whose multivariate distributions are Gaussian and Student's t-copulas: Set $x_i = F_i^{-1}(u_i)$, $i = 1,...,n$, and generate a random sample following Student's t and Gaussian copulas. At the last stage, repeat the Monte Carlo simulation steps m times to generate m pairs of $(x_1,...,x_n)$.

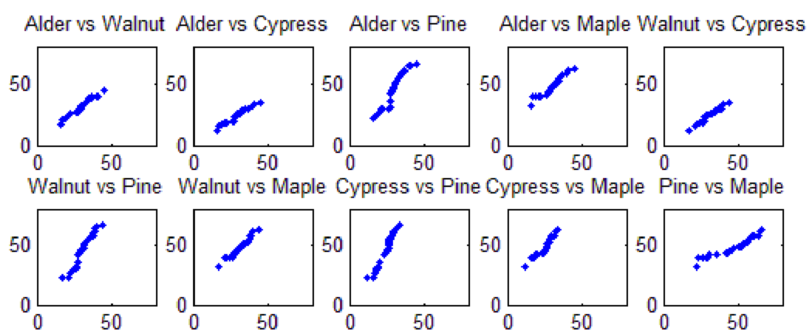

Figure 3. Scatter plot of the development times of *Ae. Albopictus*

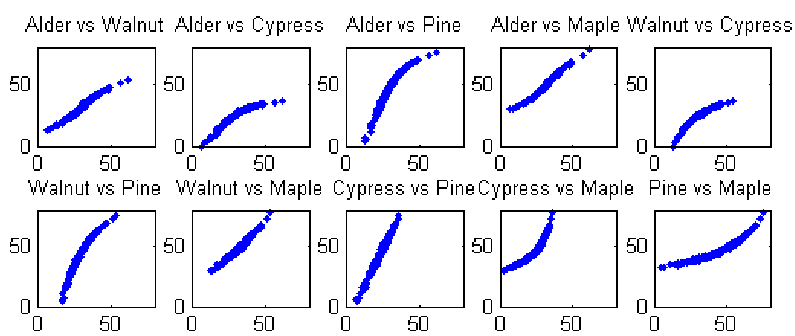

Figure 4. Student's t-copula with $\nu = 10$, 500 simulated realizations

As demonstrated in Figure 1 in Section 2, the behavior in the tails of the distribution depends solely on the type of copula and not on the choice of marginal distribution. Therefore, for copula-based Monte Carlo simulation, care must be taken to select an appropriate copula model. Poorly chosen copula may lead to some undesirable results due to a failure to correctly identify dependence structure in the sample. In Section 3, we have chosen

Student's t-copula with 10 degrees of freedom as the best copula for our data among others. Figure 3 shows the scatter plots of data, and Figure 4 displays the scatter plots of 500 simulated realizations from the Student's t-copula in association with the marginal distributions listed in Table 1. Figure 4 clearly demonstrates that there are strong positive effects between the leaf species. Unlike the usual linear correlation coefficient, Student's t-copula captures non-linear trends as well as linear trends. For example, linear relationship on mosquito development time was found from treatments of black alder vs. black walnut, black alder vs. sugar maple, black walnut vs. sugar maple, and bald cypress vs. eastern white pine. On the other hand, non-linear relationship was observed from black alder vs. bald cypress, black alder vs. eastern white pine, black walnut vs. bald cypress, black walnut vs. eastern white pine, bald cypress vs. sugar maple and eastern white pine vs. sugar maple. In the linear relationship, it can be interpreted that the effect of the two different leaf litter treatments on the development time of *Ae. albopictus* would be constant from early to late development time. In case the relationship is non-linear, the effect of two different leaf species treatment on the development time of *Ae. albopictus* would vary depending on part of development time. For example, the development time of *Ae. albopictus* under the black alder and eastern white pine treatment would be affected more by black alder than by eastern white pine at late development time. In another example, the treatments of bald cypress vs. sugar maple indicates that the development time of *Ae. albopictus* would be more affected by sugar maple than by bald cypress at late development time and the effect would be reversed at early development time.

4.2 Development Time Expectancy

We identify and measure the anticipated development time through simulated development times. To this end, based on Student's t-copula with $\nu = 10$, we generate a large number of realizations for five random variables that represent the five leaf species. Specifically, letting X_1, X_2, X_3, X_4 and X_5 be random variables that represent Black Alder, Black Walnut, Common Bald Cypress, Eastern White Pine and Sugar Maple leaf infusion treatments, respectively, we generate 20,0000 realizations of the random variables. That is, 20,0000 x_i 's ($i = 1,...,5$) are simulated from 28 original data points using Student's t-copula with $\nu = 10$. Based on the simulated realizations, the averages for x_1, x_2, x_3, x_4 and x_5 are 30.5418, 31.7152, 24.8212, 46.7644 and 47.9355, respectively.

Table 2. Development time estimates (in days), Student's t-copula with ν =10

Per (%)	0.1	0.5	1	5	10	30	50	70	90	95	99	99.5	99.9
							Black Alder						
SP*	3.82	9.94	12.72	19.14	22.04	27.23	30.47	33.73	38.91	41.81	48.28	51.03	57.78
CA**	0.20	6.18	8.83	15.16	17.96	-	-	-	43.00	45.81	52.26	55.06	61.54
							Black Walnut						
SP	11.92	15.17	16.81	21.17	23.50	28.33	31.66	34.99	39.80	42.14	46.56	48.21	51.66
CA	10.21	13.17	14.61	18.48	20.46	-	-	-	42.86	44.85	48.79	50.28	53.39
							Common Bald Cypress						
SP	-2.81	4.10	7.31	14.35	17.53	22.81	25.67	28.06	30.86	32.01	33.88	34.51	35.71
CA	-6.98	-0.17	2.88	9.96	13.05	-	-	-	32.27	33.15	34.69	35.23	36.26
							Eastern White Pine						
SP	-20.6	-3.62	3.87	21.31	28.98	41.86	48.85	54.66	61.48	64.27	68.87	70.37	73.30
CA	-30.6	-14.1	-6.74	10.53	18.05	-	-	-	64.92	67.06	70.83	72.11	74.61
							Sugar Maple						
SP	28.24	30.89	32.34	36.39	38.67	43.74	47.49	51.45	57.54	60.67	66.81	69.19	74.50
CA	26.87	29.24	30.46	33.91	35.77	-	-	-	61.70	64.45	70.12	72.37	77.23
							Aggregate						
SP	4.27	11.33	14.65	22.48	26.15	32.80	36.82	40.57	45.71	48.17	52.87	54.64	58.55
CA	-0.03	6.90	10.04	17.63	21.07	-	-	-	48.94	51.05	55.32	56.99	60.59

*SP = sample percentile, **CA = conditional average.

For a sample of m observations, when the observations are ordered from small to large, the resulting ordered data are called the order statistics of the sample. The pth percentile is the value of the order statistics of m observations that exceeds p% of the observations and is less than the remaining ($100 - p$)%. For example, the 60th percentile means that 60% of all the observations are less than it and 40% are greater. The 50th percentile is called the median. To estimate the maximum feasible development time for *Ae. albopictus* that interacts with different leaf species, we use the 95th, 99th, 99.5th and 99.9th percentiles of a simulated sample of size 200,000. These imply estimated values at risk at 95, 99, 99.5 and 99.9 percent confidence levels, i.e., 95%, 99%, 99.5% and 99.9% longest development times, respectively. To measure the minimum feasible time in the lower tail, we use the 0.1th, 0.5th, 1th and 5th percentiles. For comparison, we also calculate other sample percentiles such as the 10th, 30th and 50th percentiles. This measurement technique is similar to the notion of Value at Risk often used in the area of financial risk management (Jorion, 2007) to hedge risks. Based on 200,000 simulated realizations, Table 2 reports various the percentiles of the mosquito development times under the five different leaf species treatments. For example, at the level of 95 % confidence, female *Ae. albopictus* emerged from eastern white pine infusion would have the longest development time of 64.27 days, while female mosquitoes from bald cypress infusion would have the shortest development time of 32.01 days. Some negative values are located in the lower tails in bald cypress and eastern white pine treatment groups. Development time cannot be a negative value. They just fall into negative territory due to a basic distributional property applied, and so have no practical meaning. Conditional average is also used to investigate the anticipated development time. This averages data over all levels greater (lower) than or equal to a specific percentile in the upper (lower) tail of the distribution for each sample, telling us the averaged size of the development time that excesses the percentile. For example, Sugar Maple that has a 95th percentile of 60.67 has 64.45 as the conditional average at 95% confidence level.

In study of development time under several leaf species, it is worth using an aggregate distribution of the five different leaf species used. Consider a linear combination of $X_1,...,X_5$,

$$G = w_1X_1 + w_2X_2 + w_3X_3 + w_4X_4 + w_5X_5 ,$$

where $w_1,...,w_5$ are the weights corresponding to each X_i taken on the real number such that $\sum_{i=1}^{5} w_i = 1$. In this work, we assume that it is equally likely to avoid a possible bias that may cause erroneous percentiles when no information about the weight is available. That is, the weight used for each variable is 0.2. Let g be a realization of G . Then, we have 200,000 g 's by generating 200,000 x_i 's. Based on the distribution of G with 200,000 g 's, Table 2 also shows the percentiles of the aggregate distribution for which the average is 36.36. The results indicate that the expected development time for a female *Ae. albopictus* would be 48.17 days at the level of 95 % confidence when the effects of all leaf litter treatments are combined, which would be longer than the time for a mosquito emerging from bald cypress infusion and shorter than the time for a mosquito emerging from eastern white pine.

Table 3. Development time estimates (in days), Gaussian copula

Per (%)	0.1	0.5	1	5	10	30	50	70	90	95	99	99.5	99.9
							Black Alder						
SP*	3.71	10.28	12.98	19.27	22.14	27.14	30.50	33.76	38.92	41.74	47.92	50.67	57.00
CA**	0.31	6.34	9.06	15.31	18.08	-	-	-	42.91	45.64	51.84	54.55	60.67
							Black Walnut						
SP	11.92	15.32	16.92	21.24	23.55	28.34	31.68	35.03	39.81	42.10	46.36	47.97	51.27
CA	10.35	13.31	14.76	18.58	20.54	-	-	-	42.80	44.74	48.56	50.03	53.00
							Common Bald Cypress						
SP	-2.67	4.51	7.54	14.49	17.59	22.82	25.69	28.08	30.86	31.99	33.81	34.43	35.59
CA	-6.66	0.08	3.13	10.11	13.17	-	-	-	32.25	33.11	34.61	35.14	36.12
							Eastern White Pine						
SP	-20.2	-2.80	4.55	21.61	29.20	41.89	48.87	54.72	61.48	64.23	68.68	70.16	73.00
CA	-30.0	-13.4	-6.10	10.93	18.37	-	-	-	64.86	66.97	70.63	71.92	74.32
							Sugar Maple						
SP	28.12	31.04	32.45	36.46	38.75	43.76	47.51	51.49	57.55	60.61	66.58	68.93	73.91
CA	26.96	29.34	30.59	33.99	35.85	-	-	-	61.62	64.31	69.81	72.02	76.59
							Aggregate						
SP	4.18	11.70	14.92	22.62	26.26	32.82	36.85	40.61	45.71	48.13	52.64	54.41	58.14
CA	0.25	7.16	10.32	17.82	21.22	-	-	-	48.88	50.94	55.07	56.71	60.11

*SP = sample percentile, **CA = conditional average

We include the results from Gaussian copula which underestimates the effects of extreme events. The use of Gaussian copula will observe the consequence of ignoring heavy tails. The observed discrepancy of development time between the two copulas would demonstrate the importance of choosing appropriate dependence structure. Using Gaussian copula, Table 3 presents various the percentiles of the mosquito development times under the five different leaf species treatments from 200,000 simulated realizations. As expected, Gaussian copula generates lower development time at all percentile levels when compared to Student's t-copula. The same phenomena are observed for the conditional average case. These outcomes are reasonable due to the fact that Gaussian copula does not allow for extreme events to be dependent, while Student's t-copula has both lower and upper tail dependence and captures dependence between extreme events. From Table 3, the development time estimates with the assumption of Gaussian copula are likely to be underestimated. However, it appears that the difference between the two copulas is not significant. So, although there is a different impact of choice of copula, the impact of the extreme events on the development time seem to be modest.

5. Concluding Remarks

Analyzing relationships between development time of *Ae. albopictus* and the effects of different types of leaf species treatment help us in developing mosquito control strategy such as determining a critical time and priority to treat mosquito habitats based on the information regarding surrounding vegetations. We used copula, as a major vehicle of drawing inferences about the development time, which fully captures the association structure of leaf litters on mosquito performance. An appropriately chosen copula model can delineate the effects of different leaf species treatments on development time of *Ae. albopictus*, and so help predict expected development time for the pest. For our data, Student's t-copula with 10 degrees of freedom was chosen as the best copula within the class of elliptical copulas. We also carried out some goodness-of-fit tests for the copula model. By performing Monte Carlo simulation, a large number of simulated development times are generated. From the simulated development time, we calculated the percentiles to determine expected development time of female *Ae. albopictus* under the five individual leaf species treatments, as well as the combination of them. We found that among different leaf species treatments, eastern white pine treatment would produce *Ae. albopictus* with the longest expected development time and bald cypress infusion would produce *Ae. albopictus* with the shortest development time at a given confidence level. Also, when all the effects of leaf infusion was combined, the expected development time would be longer than the time of mosquito under bald cypress treatment and shorter that the time of mosquito under eastern white pine treatment at a given confidence level. Our study also suggests some future research topics such as the use of other copulas. For example, Archimedean copulas address asymmetries, and the grouped t copula allows heterogeneity of data. The use of these copulas could be useful when a lack of fit in elliptical copulas is detected.

Acknowledgement

Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

References

Alto, B. W., & Juliano, S. A. (2001). Temperature effects on the dynamics of *Aedes albopictus* (Diptera: Culicidae) populations in the laboratory. *J. Med. Entomol, 38*, 548-556.

Alto, B. W., Lounibos, L. P., Higgs, S., & Juliano, S. A. (2005). Larval competition differentially affects arbovirus infection in *Aedes* mosquitoes. *Ecology, 86,* 3279-3288.

Barrera, R. (1996). Competition and resistance to starvation in larvae of container-inhabiting *Aedes* mosquitoes. *Ecol. Entomol, 21*, 117-127.

Breymann, W., Dias, A., & Embrechts P. (2003). Dependence Structures for Multivariate High-Frequency Data in Finance. *Quantitative Finance, 3,* 1-14.

David, J. P., Delphine, R., Gerard, M., & Meyran, J. C. (2000). Larvicidal effect of a cell-wall fraction isolated from alder decaying leaves. *Journal of Chemical Ecology, 26*, 901-913.

Deheuvels, P. (1978). Caracterisation complete des lois extremes multivariees et de la convergence des types extremes. *Publications de l'Institut de Statistique de l'Universite de Paris, 23,* 1-36.

Demarta, S., & McNeil, A. (2005). The *t* Copula and Related Copulas. *International Statistical Review*, *73*, 111-129.

Efron, B., & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall/CRC.

Embrechts, P., Lindskog, F., & McNeil, A. (2003). *Modelling Dependence with Copulas and Applications to Risk Management, Handbook of Heavy Tailed Distributions* (Chapter 8, pp. 329-384). Elsevier.

Embrechts, P., McNeil, A., & Straumann, D. (1999). *Correlation and Dependence in Risk Management: Properties and Pitfall, Risk Management: Value at Risk and Beyond*. Cambridge.

Genest, C., & Rémillard, B. (2008). Validity of the Parametric Bootstrap for Goodness-of-Fit Testing in Semiparametric Models. *Annales l'Institut Henri-Poincare, 44*, 1096-1127.

Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-Fit Tests for Copulas: A Review and a Power Study. *Insurance: Mathematics and Economics, 44*, 199-213.

Jorion, P. (2007). Value at Risk: The New Benchmark for Managing Financial Risk. McGraw-Hill

Juliano, S. A., & Lounibos, L. P. (2005). Ecology of invasive mosquitoes: effects on resident species and on human health. *Ecol. Lett., 8,* 558-574.

Kumar, P., & Shoukri, M. (2007). Evaluating Aortic Stenosis Using the Archimedean Copula Methodology. *Journal of Data Science*, *6*, 173-187.

Lee, S., & Yang, S. (2007). Checking the Censored Two-Sample Accelerated Life Model using Integrated Cumulative Hazard Difference. *Lifetime Data Analysis*, *13*, 371-380.

Lounibos, L. P. (2002). Invasions by insect vectors of human disease. *Annu Rev Entomol, 47,* 233-266.

Mlakar, J., Korva, M., Tul, N., Popovic, M., Poljsak-Prijatelj, M., Mraz, J., … Avsic Zupanc, T. (2016). Zika Virus Associated with Microcephaly. *The New England journal of medicine, 374*, 951-958.

Moore, C. G. (1999). Aedes albopictus in the United States: current status and prospects for further spread. *Journal of the American Mosquito Control Association, 15,* 221-227.

Moore, C. G., & Mitchell, C. J. (1997). *Aedes albopictus* in the United States: ten-year presence and public health implications. *Emerg. Infect. Dis., 3*, 329-334.

Murrell, E. G., & Juliano, S. A. (2008). Detritus type alters the outcome of interspecific competition between *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae). *J. Med. Entomol., 45*, 375-383.

Muturi, E. J., & Alto, B. W. (2011). Larval environmental temperature and insecticide exposure alter *Aedes aegypti* competence for arboviruses. *Vector Borne Zoonotic Dis., 11*, 1157-1163.

Muturi, E. J., Costanzo, K., Kesavaraju, B., Lampman, R., & Alto, B. W. (2010). Interaction of a pesticide and larval competition on life history traits of *Culex pipiens*. *Acta Tropica, 116*, 141-146.

Muturi, E. J., Kim, C. H., Alto, B. W., Berenbaum, M. R., & Schuler, M. A. (2011b). Larval environmental stress alters *Aedes aegypti* competence for Sindbis virus. *Tropical Medicine and International Health, 16*, 955-964.

Muturi, E. J., Lampman, R., Costanzo, K., & Alto, B. W. (2011a). Effect of temperature and insecticide stress on life-history traits of *Culex restuans* and *Aedes albopictus* (Diptera: Culicidae). *Journal of Medical Entomology, 48*, 243-250.

Nelsen, R. (1999). *An Introduction to Copulas*. New York, Springer Verlag.

O'Meara, G. F., Evans, L. F., Gettman Jr., A. D., & Cuda, J. P. (1995). Spread of *Aedes albopictus* and decline of *Ae. aegypti* (Diptera: Culicidae) in Florida. *J. Med. Entomol, 32*, 554-562.

Pavela, R. (2008). Larvicidal effects of various Euro-Asiatic plants against *Culex quinquefasciatus* Say larvae (Diptera: Culicidae). *Parasitol. Res., 102,* 555-559.

Schweizer, B., & Sklar, A. (1983). *Probabilistic Metric Spaces*. New York: North-Holland.

Sklar, A. (1959). Functions de Repartition a n Dimensions et leurs Merges, Publication of the Institute of Statistics, University of Paris.

Tekiela, S. (2006). Trees of Illinois Field Guide. Adventure Publications, Inc., Cambridge, Minnesota.

Yee, D. A. (2016). Thirty Years of Aedes albopictus (Diptera: Culicidae) in America: An Introduction to Current Perspectives and Future Challenges. *Journal of Medical Entomology, 53*, 989-991.

Zheng, M. & Klein, J. (1995). Estimates of Marginal Survival for Dependent Competing Risks Based on an Assumed Copula. *Biometrika, 82*, 127-138.