NETWORKS AND HETEROGENEOUS MEDIA
©American Institute of Mathematical Sciences
Volume **8**, Number **1**, March **2013**
doi:10.3934/nhm.2013.8.149
pp. **149–170**

ARCHIMEDEAN COPULA AND CONTAGION MODELING IN EPIDEMIOLOGY

Jacques Demongeot

FRE 3405, AGIM (AGeing Imaging Modeling), CNRS-UJF-EPHE-UPMF
University J. Fourier of Grenoble
Faculty of Medicine of Grenoble, 38700 La Tronche, France

Mohamad Ghassani and Mustapha Rachdi

FRE 3405, AGIM (AGeing Imaging Modeling), CNRS-UJF-EPHE-UPMF
Université Pierre Mendès France, UFR SHS, BP.47, 38040 Grenoble Cedex 09
Faculty of Medicine of Grenoble, 38700 La Tronche, France

Idir Ouassou and Carla Taramasco

FRE 3405, AGIM (AGeing Imaging Modeling), CNRS-UJF-EPHE-UPMF
Faculty of Medicine of Grenoble, 38700 La Tronche, France

Abstract. The aim of this paper is first to find interactions between compartments of hosts in the Ross-Macdonald Malaria transmission system. So, to make clearer this association we introduce the concordance measure and then the Kendall's tau and Spearman's rho. Moreover, since the population compartments are dependent, we compute their conditional distribution function using the Archimedean copula. Secondly, we get the vector population partition into several dependent parts conditionally to the fecundity and to the transmission parameters and we show that we can divide the vector population by using p-th quantiles and test the independence between the subpopulations of susceptibles and infecteds. Third, we calculate the p-th quantiles with the Poisson distribution. Fourth, we introduce the proportional risk model of Cox in the Ross-Macdonald model with the copula approach to find the relationship between survival functions of compartments.

1. **Introduction.** Advances in epidemics modelling have been done recently by introducing demographic aspects (i.e. consideration of host populations whose global size changes during the epidemic and the endemic history) as well as spatial aspects about vector or infectious agents spread.

As examples of application, malaria endemics in South Mali or recurrent seasonal influenza epidemics with irregular pandemics, are modelled by the same type of mathematical framework, coming from the Ross-Macdonald tradition. We will focus in this paper to some improvements of this classical model, in terms of:

- mechanism of contacts, supposed to be assimilated to shocks, like in stochastic chemistry.
- demography of non constant population, with fecundity and mortality parameters specific to considered populations of hosts and vectors.

2010 *Mathematics Subject Classification.* 34, 60, 62, 92.
Key words and phrases. Archimedean copula, quantile, malaria, transmission system, measure of concordance.

- differential infectious risk caused by a given population of hosts or vectors, in comparison with other populations less or more subjected to infectious risk.

These improvements come from the fact that the Ross-Macdonald model has, in the case of the malaria, the following interaction graph (cf.[9, 10, 13, 15, 21, 22, 25, 27, 28, 32, 37]):

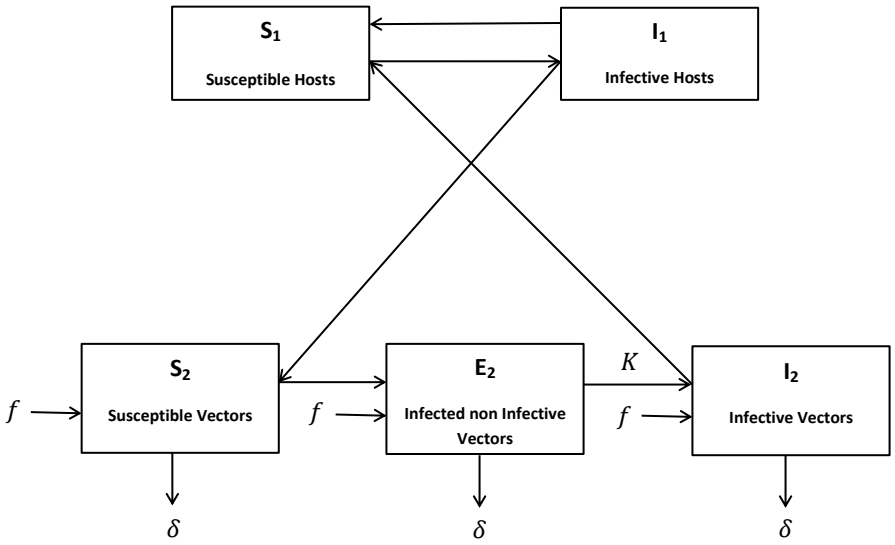

FIGURE 1. Interaction graph of the Ross-Macdonald model

The equations of the Ross-Macdonald model are:

$$\begin{aligned}
\frac{dS_1}{dt} &= \frac{-\beta_2 S_1 I_2}{N_H} + rI_1 \\
\frac{dI_1}{dt} &= \frac{\beta_2 S_1 I_2}{N_H} - rI_1 \\
\frac{dS_2}{dt} &= \omega + fS_2 - \frac{\beta_1 S_2 I_1}{N_V} - \delta S_2 \\
\frac{dI_2}{dt} &= fI_2 + KE_2 - \delta I_2 \\
\frac{dE_2}{dt} &= fE_2 + \frac{\beta_1 S_2 I_1}{N_V} - KE_2 - \delta E_2
\end{aligned} \tag{1}$$

where f (resp. δ) is the fecundity (mortality) rate of the vector population (susceptible, infected and infective vectors being supposed to have the same fecundity and mortality), β_1(resp. β_2) is the host (resp. vector) contagion parameter, N_H(resp. N_V) is the host (resp. vector) population size, the ratio $m = N_V/N_H$ is the vector/host ratio, K (resp. r) is the vector (resp. host) speed of passage from the infected/not infective (resp. infected) state to the infective (resp. susceptible) state. If $f = \delta - \mu$ (the fecundity of vectors compensating partly their mortality), the value of R_0, the mean number of secondary infected vectors for one infective

host, is equal to:

$$R_0 = \beta_1\beta_2 K/[N_H N_V \mu r(K+\mu)]$$

If $R_0 > 1$, assuming that $\omega = 1$, $N_H = 1/\mu$ and $m(0) = \frac{N_V(0)}{N_H} = 1$, the stationary state $(0,0,0,0,0)$ is unstable and the endemic stationary state is stable and reached after a transient epidemic wave for the values:

$$i_1^* = I_1^*/N_H, i_2^* = I_2^*/N_V, e_2^* = E_2^*/N_V,$$

with $i_1^* = (R_0-1)/(R_0+\beta_1/\mu), i_2^* = i_1^* r/[m\beta_2(1-i_1^*)], e_2^* = i_1^*\mu r/[Km\beta_2(1-i_1^*)]$.

If μ is small with respect to K, then $K/(K+\mu) \approx e^{-\mu/K}$, where $1/K$ is the mean sojourn time in the compartment E_2 (sporogonic cycle duration) and $R_0 = [\beta_1\beta_2/N_V N_H \mu r]e^{-\mu/K}$.

In this paper, we are interested in analyzing interactions between vector compartments in system (1), by using a copula approach. In fact, partitioning the studied population into several compartments having different statistical characteristics of interaction must be done in a rigorous way. This can be made for example by estimating the marginal distribution of the couples of sizes of susceptible and infected populations which may be done through an Archimedean copula approach. It must be said that copula functions permit to model compartments associations and then highlight their dependence structure which allows a better stratification of the population. So, to make clearer these associations we introduce the concordance measure and then the Kendall's tau and Spearman's rho. Afterward, we show that we can divide the population by using the p-th percentiles, where $p = 5\%, 25\%, 50\%, 75\%$ and 95%. Remark that it is known that the copula functions are used essentially in finance and this is the first time we use the information it brings to explain an epidemiological phenomena. We believe that this study could open a window on several future developments.

Recall that, in statistics, a copula is used as a general way of formulating a multivariate distribution in such a way that various general types of dependence can be represented. The approach to formulating a multivariate distribution using a copula is based on the idea that a simple transformation can be made on each marginal variable in such a way that each such transformed variable has a uniform distribution. Once this is done, the dependence structure can be expressed as a multivariate distribution of the obtained uniform variables, and a copula is precisely the multivariate distribution of these marginally uniform random variables. There are many families of copulas which differ in the detail of the dependence they represent. A parameterized family of distribution functions will typically depend on several parameters which relate to the strength and form of their dependence. A typical use of copulas consists in choosing such a family in order to define the adequate multivariate distribution fitting the empirical distribution observed from a sample of data, and then to derive the copula corresponding to this multivariate distribution.

A copula is a multivariate joint distribution defined on the n-dimensional unit cube $[0,1]^n$ such that every marginal distribution be uniform on the interval $[0,1]$. Let consider a copula function C for an n-dimensional random variable $X = (X_1, ..., X_n)$ defined on a probability space $(\Omega, \Sigma, \mathbb{P})$ with a joint distribution function F_X such that for any $x_1, ..., x_n \in \mathbf{R}$, (cf. [3, 6, 18, 19, 20, 24, 33, 36]), we have:

$$F_X(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n))$$

where F_i is the marginal distribution function of X_i for $i = 1, \dots, n$.

This paper is organized as follows. In Section 2, we give the definition of the Archimedean copula and provide the two examples of the Gumbel and Clayton copulas, we define and discuss the Kendall's tau, the Spearman's rho, and then, we propose a measure of concordance. At the end of this Section, we give the formulas of conditional distribution functions using Archimedean copulas. Eventually, we will discuss the conditional quantiles from copulas. In Section 3, we study an application in which we calculate the p-quantiles for any distribution function and after for the Poisson distribution. In section 4, we discuss the interaction between compartments of the system (1) (vector compartments: susceptible, infected not infectious, and infectious), by providing a link between the Kendall's tau and the regression parameter of copulas. After, we will conduct a simulation study to highlight the interaction between the distribution functions of these compartments. Eventually, we will discuss the conditional quantiles from copulas.

In the last Section, we introduce the proportional risk model of Cox in the Ross-Macdonald model (1) with the copula approach, in order to quantify the relationships existing between the survival functions of the considered compartments.

2. **Preliminaries.**

2.1. **Archimedean copula.**

The most important family of copulas is the Archimedean one [13]. This latter is defined as follows :

$$C(u_1, ..., u_n) = \begin{cases} \phi^{-1}\left(\phi(u_1) + \cdots + \phi(u_n)\right) & \text{if} \quad \phi(u_1) + ... + \phi(u_n) \leq 0 \\ \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where the generator of the copula ϕ is a twice continuously differentiable function which satisfies:

$$\phi(1) = 0, \ \phi^{(1)}(u) < 0 \ and \ \phi^{(2)}(u) > 0 \text{ for all } u \in [0, 1]$$

where $\phi^{(i)}$ denotes the ith order derivative of ϕ.

There are some popular Archimedean copulas:

- Clayton copula:

$$C(u, v) = (u^{-\alpha} + v^{-\alpha} - 1)^{\frac{-1}{\alpha}} \text{ where } \alpha > 0 \text{ and } \phi(t) = t^{-\alpha} - 1$$

- Gumbel copula:

$$C(u, v) = \exp\{-[(-\ln u)^{\alpha} + (-\ln v)^{\alpha}]^{\frac{1}{\alpha}}\} \text{ where } \alpha \geq 1 \text{ and } \phi(t) = (-\ln t)^{\alpha}$$

Let consider that U and V are two random variables defined on the probability space $(\Omega, \Sigma, \mathbb{P})$ and uniformly distributed on $[0, 1]$, and let X be the random variable $C(U, V)$ valued in $[0, 1]$.

The real α corresponds to a regression parameter.

In the case of Archimedean copula, the distribution function of X is given by:

$$F_X(t) = t - \frac{\phi(t)}{\phi^{(1)}(t)}$$

In the case of the Gumbel copula, we have $\phi(t) = (-\ln t)^\alpha$, so $\phi^{(1)}(t) = \frac{-\alpha}{t}(-\ln t)^{\alpha-1}$.
Then:

$$\begin{aligned} F_X(t) &= t - \frac{(-\ln t)^\alpha}{\frac{-\alpha}{t}(-\ln t)^{\alpha-1}} \\ &= t + \frac{t}{\alpha}(-\ln t)^{\alpha-\alpha+1} \\ &= t - \frac{t\ln t}{\alpha} \end{aligned}$$

and the probability density of X is given by:

$$f_X(t) = 1 - \frac{1}{\alpha} - \frac{\ln t}{\alpha} \tag{3}$$

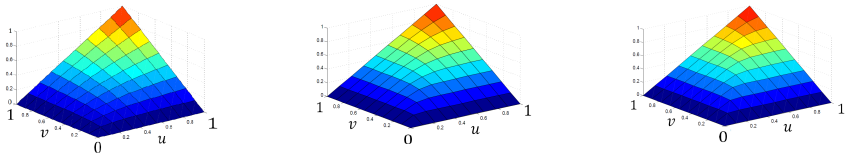

FIGURE 2. Distribution functions of Clayton bivariate copula (resp. for $\alpha = 2$, $\alpha = 5$ and $\alpha = 10$)

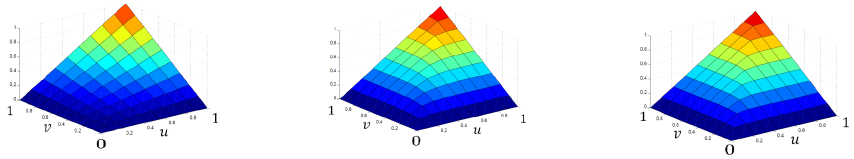

FIGURE 3. Distribution functions of Gumbel bivariate copula (resp. for $\alpha = 2$, $\alpha = 5$ and $\alpha = 10$)

2.2. **Concordance, Kendall's tau, Spearman's rho, copula.** In order to define the concordance measure, we begin by defining the concordance (cf. [24]): two observations (x_1, y_1) and (x_2, y_2) of a pair (X, Y) of continuous random variables are said to be concordant if both values of one pair are greater than the corresponding values of the other pair, that is if $x_1 < x_2$ and $y_1 < y_2$ or if $x_1 > x_2$ and $y_1 > y_2$; and they are said to be discordant if for one pair one value is greater and the other is smaller than the corresponding value of the other pair, that is if $x_1 < x_2$ and $y_1 > y_2$ or $x_1 > x_2$ and $y_1 < y_2$. The simple version of the measure of association known as Kendall's tau is defined in terms of concordance as follows (cf. [24]): let $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ denote a random sample of n observations from a continuous random vector (X, Y). The Kendall's tau for the latter random sample is defined by:

$$\tau = \frac{number\ of\ concordant\ pairs - number\ of\ discordant\ pairs}{total\ number\ of\ pairs} \tag{4}$$

Equivalently, τ may be written and interpreted as the difference between the empirical probabilities of concordance and discordance for a pair of observations (x_i, y_i) and (x_j, y_j), chosen randomly from the sample. The multivariate version of Kendall's tau will be defined similarly.
Let (X_1, Y_1) and (X_2, Y_2) be independent and identically distributed random vectors, each with joint distribution function H. So the probability that (X_1, Y_1) and (X_2, Y_2) are concordant is equal to $\mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0)$ and the probability that these two vectors are discordant is equal to $\mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0)$, and we have :

$$\mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0) = \frac{number\ of\ concordant\ pairs}{total\ number\ of\ pairs}$$
$$\mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0) = \frac{number\ of\ discordant\ pairs}{total\ number\ of\ pairs}$$

Then the multivariate generalization of Kendall's tau is defined as the difference between the probability of concordance and the probability of discordance:

$$\tau = \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0) - \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0)$$

2.2.1. *Kendall's tau and copula relation.*

Theorem 2.1. *Let X and Y be continuous random variables whose copula is C. Then the multivariate version of Kendall's tau for X and Y, denoted by $\tau_{X,Y}$ or τ_C, is given by:*

$$\tau_C = 4 \int_0^1 \int_0^1 C(u,v)c(u,v)dudv - 1, \ \textit{where}\ c(u,v) = \partial^2 C(u,v)/\partial u \partial v \qquad (5)$$

The proof is in Appendix A.

For computational purpose, there are alternative expressions for τ_C. The integral which appears in (5) can be interpreted as the expected value of the function $C(U, V)$ where U and V are random variables uniformly distributed on $[0, 1]$ for which the joint distribution function is C, i.e.,

$$\tau_C = 4\,\mathbb{E}(C(U,V)) - 1 = 4\int_0^1 t f_C(t)dt - 1 \qquad (6)$$

where $f_C(t)$ denotes the density function of C.

2.2.2. *Spearman's rho.* The multivariate version of Spearman's rho is based on the concordance and discordance values. In order to obtain the population value of this measure (cf.[24]), let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) be three independent random vectors whose components are continuous random variables with a common joint distribution function H (whose marginal distribution are F and G) and let C be a copula function. The multivariate version of Spearman's rho is three time the difference between the probabilities of concordance and discordance of the two vector pairs (X_1, Y_1) and (X_2, Y_3), i.e., two vector pairs with the same margins, but one vector has distribution function H, while the components of the other are independent:

$$\rho = 3\left[\mathbb{P}((X_1 - X_2)(Y_1 - Y_3) > 0) - \mathbb{P}((X_1 - X_2)(Y_1 - Y_3) < 0)\right]$$

The pair (X_1, X_3) could be used equally as well.

On the other hand, notice that while the joint distribution function of (X_1, Y_1) is $H(x, y)$, the joint distribution function of (X_2, Y_3) is $F(x)G(y)$ (because X_2 and Y_3 are independent), thus the copula of X_2 and Y_3 is the independence copula $(C(u_1, u_2) = u_1u_2)$.

Theorem 2.2. *Let X and Y be continuous random variables whose copula function is C. Then the multivariate version of Spearman's rho for X and Y, denoted by $\rho_{X,Y}$ or ρ_C, is given by:*

$$\rho_C = 12 \int \int_{I^2} u_1 u_2 \; dC(u_1, u_2) - 3$$

The proof is in Appendix B.

2.2.3. *Measure of concordance.*

Definition 2.3. A measure of association, denoted M, between two continuous random variables X and Y whose copula function is C, is said to be a measure of concordance, if it satisfies the following properties (in which we denote M by $M_{X,Y}$ or M_C when convenient):

- $M_{X,Y}$ is defined for every pair (X, Y) of continuous random variables
- $M_{X,Y} \in [-1, 1]$ with $M_{X,X} = 1$ and $M_{X,-X} = -1$
- $M_{X,Y} = M_{Y,X}$
- If X and Y are independent, then $M_{X,Y} = 0$
- $M_{-X,Y} = M_{X,-Y} = -M_{X,Y}$
- if C_1 and C_2 are copulas with $C_1 \leq C_2$ then $M_{C_1} \leq M_{C_2}$
- if $\{(X_n, Y_n)\}_{n\in\mathbb{N}^*}$ is a sequence of pairs of continuous random variables with copula functions C_n and if C_n converges pointwise to C then $\lim_{n\to+\infty} M_{C_n} = M_C$

Theorem 2.4. *The multivariate versions of Kendall's tau and Spearman's rho are two measures of concordance.*

The proof is in Appendix C.

2.3. **Multidimensional conditional distribution function using copulas.** Let $X_1, X_2, ..., X_k$ be random variables defined on a probability space $(\Omega, \Sigma, \mathbb{P})$, where the joint probability density function f_k of the vector $X = (X_1, X_2, ..., X_k)$ is assumed to exist. By using the Archimedean copula construction, the joint probability density function may be written as follows:

$$\begin{aligned} f_k(x_1, ..., x_k) &= \frac{\partial^k}{\partial x_1 ... \partial x_k} \phi^{-1} \left\{ \phi \left[F_1(x_1) \right] + \cdots + \phi \left[F_k(x_k) \right] \right\} \\ &= \phi^{-1(k)} \left\{ \phi \left[F_1(x_1) \right] + \cdots + \phi \left[F_k(x_k) \right] \right\} \prod_{j=1}^{k} \phi^{(1)} \left[F_j(x_j) \right] F_j^{(1)}(x_j) \end{aligned} \tag{7}$$

for all $(x_1, ..., x_k) \in \mathbb{R}^k$, where F_j denotes the marginal distribution function of X_j.

Thus, the conditional density of X_k given $X_1, ..., X_{k-1}$ is given by:

$$
\begin{aligned}
& f_k(x_k|x_1,...,x_{k-1}) \\
&= \frac{f_k(x_1,...,x_k)}{f_{k-1}(x_1,...,x_{k-1})} \\
&= \phi^{(1)}\left[F_k(x_k)\right] F^{(1)}(x_k) \frac{\phi^{-1(k)}\left\{\phi\left[F_1(x_1)\right]+\cdots+\phi\left[F_k(x_k)\right]\right\}}{\phi^{-1(k-1)}\left\{\phi\left[F_1(x_1)\right]+\cdots+\phi\left[F_{k-1}(x_{k-1})\right]\right\}}
\end{aligned} \tag{8}
$$

Further, the conditional distribution function of X_k given $X_1,...,X_{k-1}$ is also given by:

$$
\begin{aligned}
F_k(x_k|x_1,...,x_{k-1}) &= \int_{-\infty}^{x_k} f_k(x|x_1,...,x_{k-1})dx \\
&= \frac{\phi^{-1(k-1)}\left\{\phi\left[F_1(x_1)\right]+\cdots+\phi\left[F_k(x_k)\right]\right\}}{\phi^{-1(k-1)}\left\{\phi\left[F_1(x_1)\right]+\cdots+\phi\left[F_{k-1}(x_{k-1})\right]\right\}} \\
&= \frac{\phi^{-1(k-1)}\left\{c_{k-1}+\phi\left[F_k(x_k)\right]\right\}}{\phi^{-1(k-1)}(c_{k-1})}
\end{aligned} \tag{9}
$$

where $c_k = \phi[F_1(x_1)]+\cdots+\phi[F_k(x_k)]$.

3. **A case where the population sizes, S_1 and I_2, are Poissonian random variables.** A simple way to prove the Poissonian character of the distribution of the sizes S_1 and I_2, considered as random variables, is to take into account only the two compartments of random size S and I (for the sake of simplicity, we will omit in the following the indices), assuming that there is at least one event (contact, birth, death or recovering) in $(t, t+dt)$, and by denoting by f, ν, μ and ρ respectively the fecundity, contagion, mortality and recovering rates, we can write:

$$
\begin{aligned}
&\mathbb{P}\Big(S(t+dt)=k, I(t+dt)=j\Big) = \\
&\quad \mathbb{P}\Big(S(t)=k, I(t)=j\Big)\Big[1-\big(\nu kj - fk + \mu k - \rho j\big)dt\Big] \\
&\quad + \Big[\nu(k+1)(j-1)\mathbb{P}\Big(S(t)=k+1, I(t)=j-1\Big) + (f(k-1)+\rho(j+1)) \\
&\quad \mathbb{P}\Big(S(t)=k-1, I(t)=j+1\Big) + \mu(k+1)\mathbb{P}\Big(S(t)=k+1, I(t)=j-1\Big)\Big]dt
\end{aligned} \tag{10}
$$

and by multiplying equation (10) by k and summing over k and j, we obtain:

$$
\begin{aligned}
&\sum_{k,j\geq 0} \frac{k\Big[\mathbb{P}\Big(S(t+dt)=k, I(t+dt)=j\Big) - \mathbb{P}\Big(S(t)=k, I(t)=j\Big)\Big]}{dt} \\
&\quad = \frac{d\Big[\sum_{k,j\geq 0} k\mathbb{P}\Big(S(t)=k, I(t)=j\Big)\Big]}{dt}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{k,j\geq 0} \Big[-\nu k^2 j \mathbb{P}\Big(S(t)=k, I(t)=j\Big) \\
&+ \nu k(k+1)(j-1)\mathbb{P}\Big(S(t)=k+1, I(t)=j-1\Big) \\
&+ k(f(k-1)+\rho(j+1))\mathbb{P}\Big(S(t)=k-1, I(t)=j+1\Big) + (fk^2+\rho kj) \\
&\mathbb{P}\Big(S(t)=k, I(t)=j\Big) - \mu k^2 \mathbb{P}\Big(S(t)=k, I(t)=j\Big) \\
&+ \mu k(k+1)\mathbb{P}\Big(S(t)=k+1, I(t)=j-1\Big)\Big]
\end{aligned} \tag{11}
$$

Hence:

$$
\begin{aligned}
&\frac{d\Big[\sum_{k,j\geq 0} k\mathbb{P}\Big(S(t)=k, I(t)=j\Big)\Big]}{dt} \approx \\
&-\sum_{k,j\geq 0} \Big[\nu kj\mathbb{P}\Big(S(t)=k, I(t)=j\Big) + (fk+\rho j)\mathbb{P}\Big(S(t)=k, I(t)=j\Big) \\
&-\mu k\mathbb{P}\Big(S(t)=k, I(t)=j\Big)\Big]
\end{aligned} \tag{12}
$$

and we get from (12) the following expectation equation:

$$
\frac{d\mathbb{E}(S)}{dt} = f\mathbb{E}(S) - \nu\mathbb{E}(SI)dt - \mu\mathbb{E}(S) + \rho\mathbb{E}(I) \tag{13}
$$

If the random variables S and I can be considered as uncorrelated, we can obtain from (13) a deterministic differential equation ruling $E(S)$ and $E(I)$:

$$
\frac{d\mathbb{E}(S)}{dt} = \mathbb{E}(S)[f-\mu-\nu\mathbb{E}(I)] + \rho\mathbb{E}(I) \tag{14}
$$

Then, we get the Ross macroscopic equation for the deterministic variable $\bar{S}$ (resp. $\bar{I}$) which represents the size of the susceptible (resp. infected) population:

$$
\frac{d\bar{S}}{dt} = -\nu\bar{S}\bar{I} + (f-\mu)\bar{S} + \rho\bar{I} \tag{15}
$$

By replacing k by k^2 in equation (12), we obtain more, the differential system ruling the not centred moments of order 2, $\mathbb{E}(S^2(t))$, then we get the differential system ruling the variance: $V(S(t)) = \mathbb{E}(S^2(t)) - \mathbb{E}^2(S(t))$, and we can draw the confidence cylinder (or viability tube) around the expected trajectory. By replacing k by s^k in the case where $f=\mu$ (the host population is constant) and where I can be considered as independent of S (which can be tested statistically and is observed for example if the contagion rate ν is sufficiently small with respect to the whole population sizes of hosts and vectors), we get the differential system ruling the generating function $\psi(s)$ of the distribution of S:

$$
\frac{d\psi}{dt} \approx \rho\psi - \nu\frac{d\psi}{ds}
$$

this equation having a stationary solution: $\psi = \exp((\rho/\nu)(s-1))$. The generating function ψ corresponds to a Poisson distribution for S, with expectation and variance equal to ρ/ν, the last one being in general not negligible, which explains the differences observed between the random and the deterministic models. The same argument can be used for proving the Poissonian character of the distribution of I. As equation (15) is similar to the equation relative to S_2 in system (1), it

is possible to obtain all other equations of system (1) from a microscopic random contact mechanism and by using the same type of arguments, to show that all variables involved $(S_1, I_1, S_2, I_2, E_2)$ are Poissonian. Then, the copula approach can serve to show the independency between the pairs of variables (S_1, I_2) and (S_2, I_1), hypothesis necessary to derive the system (1) from the microscopic equations.

In the following, the statistical sampling of S and I will be numerically obtained from Poissonian distributions, whose parameter is fixed by their steady state value in the deterministic version of the epidemic model (equation (1)). Near this steady state solution, we can consider that the non correlation between S and I can be postulated if the host and vector populations are sufficiently numerous and the contagion rate is sufficiently small. This circumstance is checkable by the statistical test associated to the Kendall's tau, whose value is about zero near the steady state and whose absolute value is near 1 during the transient phase of the trajectories [34].

4. **Quantiles regression based on the system of transmission of malaria.** Recall that the regression function is the most widely used tool for describing multivariate relationships. Then, copulas functions can help to understand the full joint distribution and thus be used to address some important applications, which we tackle to explain in the sequel. In this part, we will use the Gumbel copula.

4.1. **Interaction between compartments of hosts.** Let X be the random variable $C(U, V)$ defined on the probability space $(\Omega, \Sigma, \mathbb{P})$,where F_X is the distribution function of X. From (3) we obtain (cf. [33]):

$$\begin{aligned} \mathbb{E}(X) &= \int_0^1 t f_X(t) dt \\ &= \left(1 - \frac{1}{\alpha}\right) \int_0^1 t dt - \frac{1}{\alpha} \int_0^1 t \ln t dt \\ &= \frac{1}{2} - \frac{1}{2\alpha} - \frac{1}{\alpha}\left(\frac{-1}{4}\right) \\ &= \frac{2\alpha - 1}{4\alpha} \end{aligned} \tag{16}$$

Then, from the equation (6) the Kendall's tau is expressed in the following manner:

$$\begin{aligned} \tau_C &= 4\left(\frac{2\alpha - 1}{4\alpha}\right) - 1 \\ &= \frac{\alpha - 1}{\alpha} \end{aligned} \tag{17}$$

where C denotes the Gumbel copula. Notice that the parameter α measures a degree of dependence. Recall that the higher it is, the strong is dependence between the two studied variables.

4.1.1. *Relationship between the Kendall's tau and the equilibrium of the system of transmission of malaria.* Let $(S_{21}, E_{21}), (S_{22}, E_{22}), ..., (S_{2n}, E_{2n})$ be calculated from n successive independent observations of samples of m individuals extracted from the part of the vector population made of the two compartments, susceptible and infected non infectious: S_{2i} (resp. E_{2i}) is the number of susceptible (resp.

infected non infectious) vectors, counted in the ith sample observed in mosquitoes population transmitting malaria [4, 12, 16, 23, 26, 35, 38].

From the equation (4), if the number of concordant pairs is equal to the number of discordant pairs, so we will have the Kendall's tau equal to 0 and the regression parameter equal to 1 ($\tau_C = 0$, $\alpha = 1$), then in this case the two compartments will be non concordant (notion similar to uncorrelated), property observed in case of stochastic independence, and for example when the system is in equilibrium (but also in case of deterministic chaos, cf. [5]). However, if the number of concordant pairs is equal to the total number of pairs, so we will have $\tau_C = 1$ and then we will have the perfect concordance and so the system is in transient state. The sampling used to calculate τ_C can also be obtained by simulating Individual Based Models (IBM) in which social networks allow to simulate all the possibilities of contact between hosts and vector, by using the same simulation techniques as in stochastic molecular kinetics [1, 2, 7, 17, 29, 30, 31]. An example of such a social network has been simulated for explaining the contagion process of a social disease, the obesity [11]. This social network has been simulated on Figure 4, in which the homophilic rule is based on the fact that individuals tend to interact with those who resemble them in terms of social behaviour.

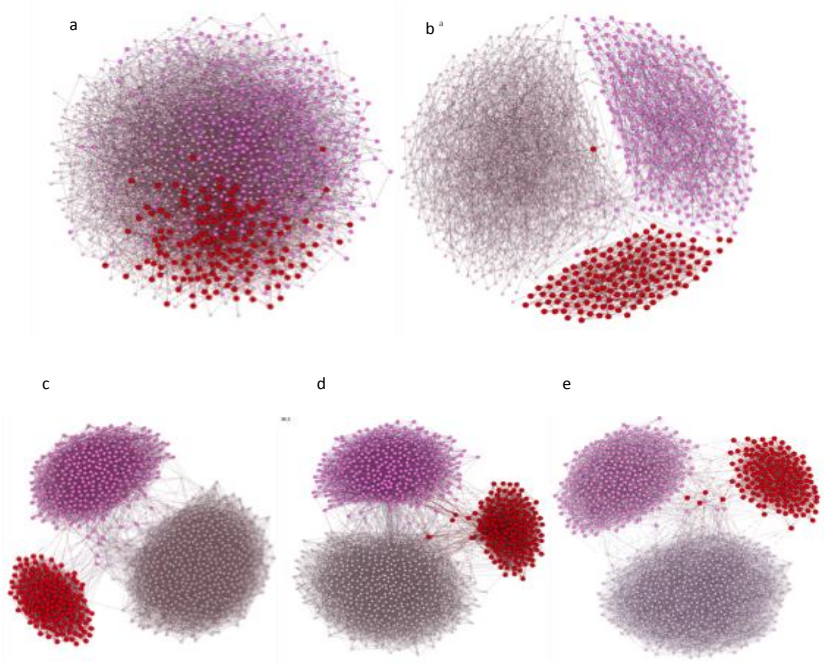

FIGURE 4. Simulation of social graphs representing a contagion network, with initial conditions (a) and asymptotic state in case of an homophilic graph (b), random graph (c), scale free graph (d) and small world graph (e).

In the system (1), when the contagion parameter β_1 or the kinetic parameter K between the two compartments S_2 and E_2 increases, then the number of concordant pairs increases, because in both cases the size of S_2 diminishes and the size of E_2 is growing, thus the Kendall's tau increases, so we could expect that the system is still in a transient state.

Let consider a simplified version of the system (1), given by:

$$\begin{aligned} \frac{dS_1}{dt} &= -\beta_2 S_1 I_2 + rI_1 \\ \frac{dI_1}{dt} &= \beta_2 S_1 I_2 - rI_1 \\ \frac{dS_2}{dt} &= \omega - \beta_1 S_2 I_1 - \mu S_2 \\ \frac{dE_2}{dt} &= \beta_1 S_2 I_1 - (\mu + K) E_2 \\ \frac{dI_2}{dt} &= KE_2 - \mu I_2 \end{aligned} \tag{18}$$

If we calculate its Jacobian matrix J, we get :

$$J = \begin{pmatrix} -\beta_2 I_2 & r & 0 & 0 & -\beta_2 S_1 \\ \beta_2 I_2 & -r & 0 & 0 & \beta_2 S_1 \\ 0 & -\beta_1 S_2 & -\beta_1 I_1 - \mu & 0 & 0 \\ 0 & \beta_1 S_2 & \beta_1 I_1 & -(\mu + K) & 0 \\ 0 & 0 & 0 & K & -\mu \end{pmatrix}$$

0 is a trivial eigenvalue and the rest of the spectrum is given by the eigenvalues of the submatrix A:

$$A = \begin{pmatrix} -r & 0 & 0 & rI_1/I_2 \\ -\beta_1 S_2 & -\beta_1 I_1 - \mu & 0 & 0 \\ \beta_1 S_2 & \beta_1 I_1 & -(\mu + K) & 0 \\ 0 & 0 & K & -\mu \end{pmatrix}$$

If we neglect the mortality and the recovering parameter r, the eigenvalues can be calculated at the stationary state by making explicit the characteristic polynomial: $\lambda(K - \lambda)(\lambda - \beta_1(I_1^* - S_2^*)) = \lambda(K - \lambda)(\lambda - \beta_1(I_1^* - \omega/(\beta_1 I_1^*)))$. If β_1 increases, we have seen in Introduction that R_0 increases, then I_1^* tends to 1 if $K >> \beta_1 - \omega$, the rate of convergence to the stationary state being given by $\log(K/(\beta_1 - \omega))$, which causes a long transient. We observe the same behaviour, if $\beta_1 >> K - \omega$.

For example, let us take a sample of tail 1000 in each compartment. Figures 5, 6, 7 and 8 show that the interaction between some pairs of random variables coming from the set S_1, I_1, S_2, I_2, E_2 with the Gumbel and Clayton copulas. In each figure we have three graphs. In the first graph on the left we assume that there is neither a contagion nor a fertility for the two compartments so that the number of concordant pairs is equal to the number of discordant pairs: it is the case of independence between the two compartments, then the regression parameter is equal to 1 and the Kendall's tau is equal to 0, so in this case the equilibrium of the system is reached. In the graph that is in the middle, we assume that there is a dependency between the two compartments, with a regression parameter α equal to 3, so the Kendall's tau increases and we have an unbalanced system. In the graph on the right, we increased the part of dependency between the two compartments, by taking the regression parameter equal to 5, thus the Kendall's tau increases a

little more and we have a system increasingly unbalanced. To summarize, when $\tau_C = 1$, the system is totally unbalanced and when $\tau_C = 0$, the system is totally balanced.

In the graphs that are in the middle and right of Fig. 5 to 8, it is clear that we are not in the case of complete independence between variables of the studied couples, so for checking if there are sub-populations where the dependency is more or less important, we will use in the next section the quantiles to divide the population into several parts, where variables of the studied couples inside each part are more or less dependent, regarding their Kendall's tau.

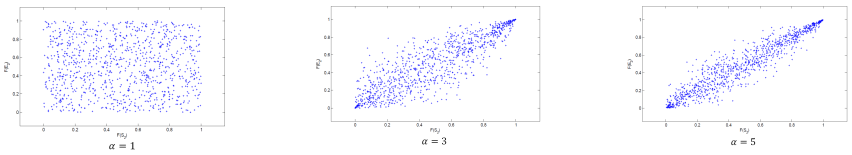

FIGURE 5. Interaction between the distribution functions of S_2 and E_2 using the Gumbel copula with the parameter of regression equal to 1, 3 and 5 respectively

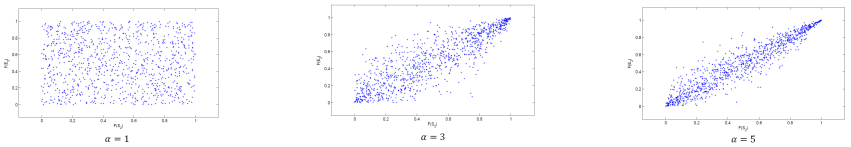

FIGURE 6. Interaction between the distribution functions of E_2 and I_2 using the Gumbel copula with the parameter of regression equal to 1, 3 and 5 respectively

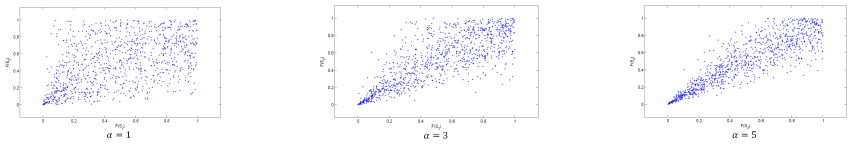

FIGURE 7. Interaction between the distribution functions of S_2 and E_2 using the Clayton copula with the parameter of regression equal to 1, 3 and 5 respectively

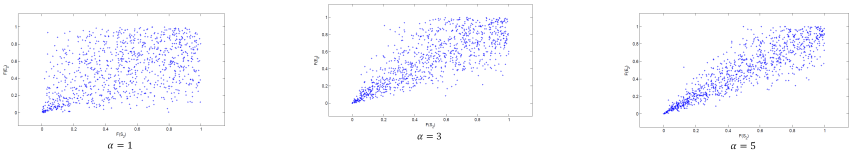

FIGURE 8. Interaction between the distribution functions of E_2 and I_2 using the Clayton copula with the parameter of regression equal to 1, 3 and 5 respectively

4.2. **Quantile regression using the bivariate Gumbel copula.** In general, it is known that the calculation of the regression function is tedious. As an alternative, copulas are well-suited to the concept of quantile regression. Instead of examining the mean of a conditional distribution, one looks at the median or some other quantiles (for instance, percentiles) of this distribution. For $p \in [0,1]$, the p-th quantile is defined as the solution t_p of the equation: $p = F_{X_k}(t_p | X_1 = x_1, ..., X_{k-1} = x_{k-1})$, that we will denote, for the sake of simplicity, in the following by:

$$p = F_k(t_p | x_1, ..., x_{k-1})$$

Let F_{S_2} and F_{E_2} be the distribution functions of the sizes of the two vector compartments, susceptible and infected respectively. The bivariate Gumbel copula for these two compartments is defined as follows:

$$C\left(F_{S_2}, F_{E_2}\right) = \exp\{-[(-\ln F_{S_2})^\alpha + (-\ln F_{E_2})^\alpha]^{\frac{1}{\alpha}}\} \tag{19}$$

where α is the parameter of regression. So, we can write the equation (19) as follows:

$$[-\ln C\left(F_{S_2}, F_{E_2}\right)]^\alpha = (-\ln F_{S_2})^\alpha + (-\ln F_{E_2})^\alpha \tag{20}$$

For simplicity, we denote $C\left(F_{S_2}, F_{E_2}\right)$ by C. Now, we take the partial derivative with respect to F_{S_2} of both sides of the equation (20) to get:

$$\frac{(-\ln C)^{\alpha-1}}{C} \frac{\partial C}{\partial F_{S_2}} = \frac{(-\ln F_{S_2})^{\alpha-1}}{F_{S_2}}$$

Then, we extract the first partial derivative denoted $C_1 = \partial C / \partial F_{S_2}$ as follows:

$$\begin{aligned} C_1 &:= C_1(F_{S_2}, F_{E_2}) \\ &= \frac{\partial C(F_{S_2}, F_{E_2})}{\partial F_{S_2}} \\ &= \frac{\partial C}{\partial F_{S_2}} \\ &= \left(\frac{\ln F_{S_2}}{\ln C}\right)^{\alpha-1} \frac{C}{F_{S_2}} \end{aligned} \tag{21}$$

By symmetry, we get also the second partial derivative as follows:

$$C_2 := \frac{\partial C}{\partial F_{E_2}} = \left(\frac{\ln F_{E_2}}{\ln C}\right)^{\alpha-1} \frac{C}{F_{E_2}} \tag{22}$$

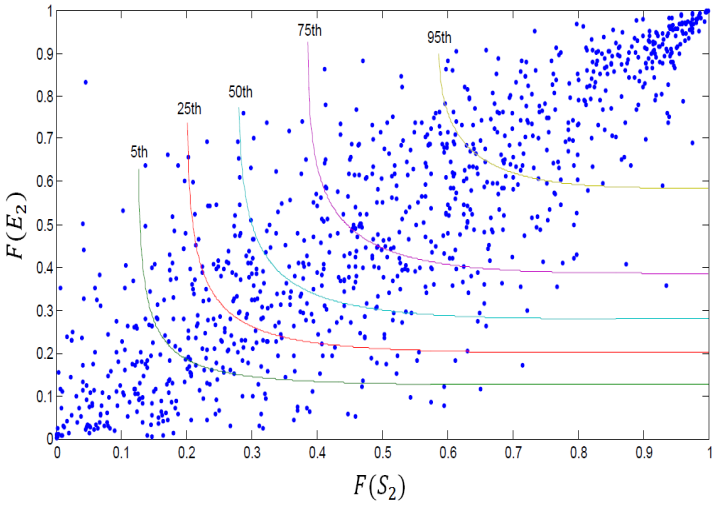

FIGURE 9. Interaction between the distribution functions of S_2 and E_2 using the Gumbel copula with $\alpha = 3$, and using the quantile regression (of E_2 on S_2) curves.

Using the conditional distribution function from (9) and (21), we obtain the p-th quantile, $e_{2,p}(s)$ as follows:

$$\begin{aligned} p &= F_{E_2}(e_{2,p}(s)|S_2 = s) \\ &= \frac{\phi^{-1}\left\{c_1(s, e_{2,p}(s)) + \phi\left[F_{E_2}(e_{2,p}(s))\right]\right\}}{\phi^{-1}(c_1(s, e_{2,p}(s)))} \\ &= \frac{\phi^{-1}\left\{\phi\left[F_{S_2}(s)\right] + \phi\left[F_{E_2}(e_{2,p}(s))\right]\right\}}{\phi^{-1}\left[\phi(F_{S_2}(s))\right]} \\ &= C_1[F_{S_2}(s), F_{E_2}(e_{2,p}(s))] \end{aligned} \tag{23}$$

where C_1 is the first partial derivative of the Archimedean copula, ϕ is the generator of the Gumbel copula ($\phi(t) = (-\ln t)^\alpha$), and $F_{E_2}(e_{2,p}(s))$ is the distribution function of the infected compartment where $e_{2,p}(s)$ denotes the p-th conditional quantile conditioned by $S_2 = s$.

Notice that, in the case of the Gumbel-Hougaard copula, one can use Equation (21) to get the distribution function $F_{E_2}(e_{2,p})$:

$$p = \left[\frac{\ln(F_{S_2})}{\ln C(F_{S_2}, F_{E_2}(e_{2,p}))}\right]^{\alpha-1} \frac{C(F_{S_2}, F_{E_2}(e_{2,p}))}{F_{S_2}} \tag{24}$$

Finding the p-th quantile $e_{2,p}$ permits to divide the population into several parts where the individuals in each part are dependent. In Figure 9, there are 500 individuals in each compartment and then there are 1000 individuals in the population. Taking into account the interaction between the two distribution functions F_{S_2} and F_{E_2} of the two compartments susceptible and infected respectively, we can divide the population, with the conditional quantiles $e_{2,p}$'s, taken for different p's, for instance for $p \in \{0.25, 0.5, 0.75\}$. For tracing the separation lines of the Figure 9, we must calculate the former conditional quantiles related to the two compartments

susceptible and infected. Then, by computing these quantiles, we can analyze each part to find the dependence between the two vector compartments.

4.3. **Quantiles from the distribution functions.** The purpose of this Section is to calculate the p-th percentile, as in equation (24), from the distribution functions of the susceptible and infective host populations sizes F_{S_1} and F_{I_1}, by writing:

$$F_{S_1}(s) = 1 - R_\gamma(s) \tag{25}$$

$$F_{I_1}(k) = 1 - R_\nu(k) \tag{26}$$

where $R_\gamma(s)$ and $R_\nu(k)$ are the rests at order k of the distribution functions F_{S_1} and F_{I_1} respectively.

Then:

$$\ln\left[F_{S_1}(s)\right] = \ln\left[1 - R_\gamma(s)\right]$$
$$\ln\left[F_{I_1}(k)\right] = \ln\left[1 - R_\nu(k)\right]$$

Consequently, using the Gumbel copula in (19), we will have :

$$C\Big(F_{S_1}(s), F_{I_1}(k)\Big) = \exp\left\{-\left[\left[-\ln\left(1 - R_\gamma(s)\right)\right]^\alpha + \left[-\ln\left(1 - R_\nu(k)\right)\right]^\alpha\right]^{\frac{1}{\alpha}}\right\} \tag{27}$$

or with the Development Limited at Order 2 $DL_2(+\infty)$ we will have when k tends to infinity:

$$\ln\left(1 - R_\gamma(s)\right) = -R_\gamma(s) - \frac{1}{2}R_\gamma^2(s) + o(R_\gamma^2(s))$$

and

$$\ln\left(1 - R_\nu(k)\right) = -R_\nu(k) - \frac{1}{2}R_\nu^2(k) + o(R_\mu^2(k))$$

Then :

$$\begin{aligned}
C\left(F_{S_1}(s), F_{I_1}(k)\right) = \exp\Bigg\{-\Bigg[&\left[R_\gamma(s) + \frac{1}{2}R_\gamma^2(s) + o(R_\gamma^2(s))\right]^\alpha \\
&+ \left[R_\nu(k) + \frac{1}{2}R_\nu^2(k) + o\left(R_\nu^2(k)\right)\right]^\alpha\Bigg]^{\frac{1}{\alpha}}\Bigg\} \\
= \exp\Bigg\{-\Bigg[&R_\gamma^\alpha(s)\left(1 + \frac{\alpha}{2}R_\gamma(s) + o(R_\gamma(s))\right) \\
&+ R_\nu^\alpha(k)\left(1 + \frac{\alpha}{2}R_\nu(k) + o(R_\nu(k))\right)\Bigg]^{\frac{1}{\alpha}}\Bigg\} \\
= \exp\Bigg\{-\left[R_\gamma^\alpha(s) + R_\nu^\alpha(k)\right]^{\frac{1}{\alpha}} \cdot \Bigg[&1 + \frac{\alpha}{2}\frac{R_\gamma^{\alpha+1}(s) + R_\nu^{\alpha+1}(k)}{R_\gamma^\alpha(s) + R_\nu^\alpha(k)} \\
&+ \frac{R_\gamma^\alpha(s)}{R_\gamma^\alpha(s) + R_\nu^\alpha(k)}o\left(R_\gamma^\alpha(s)\right) + \frac{R_\nu^\alpha(k)}{R_\gamma^\alpha(s) + R_\nu^\alpha(k)}o(R_\nu^\alpha(k))\Bigg]^{\frac{1}{\alpha}}\Bigg\} \\
= \exp\Bigg\{-\left[R_\gamma^\alpha(s) + R_\nu^\alpha(k)\right]^{\frac{1}{\alpha}} \cdot \Bigg[&1 + \frac{1}{2}\frac{R_\gamma^{\alpha+1}(s) + R_\nu^{\alpha+1}(k)}{R_\gamma^\alpha(s) + R_\nu^\alpha(k)} \\
&+ o\left(R_\gamma(s) + R_\nu(k)\right)\Bigg]\Bigg\}
\end{aligned}$$

$$= \exp\left\{ -\left[R_\gamma^\alpha(s) + R_\nu^\alpha(k)\right]^{\frac{1}{\alpha}} \right\}. \left[1 - \frac{1}{2}\frac{R_\gamma^{\alpha+1}(s) + R_\nu^{\alpha+1}(k)}{\left(R_\gamma^\alpha(s) + R_\nu^\alpha(k)\right)^{1-\frac{1}{\alpha}}} - \left(R_\gamma^\alpha(s) + R_\nu^\alpha(k)\right)^{\frac{1}{\alpha}} o\left(R_\gamma(s) + R_\nu(k)\right)\right]$$

We denote by: $L_\alpha(s,k) = R_\gamma^\alpha(s) + R_\nu^\alpha(k)$. For simplicity, we denote in the following $L_\alpha(s,k)$ by L_α.

Because $o\left(R_\gamma(s) + R_\nu(k)\right)$ tends to 0 when $s \to \infty$ and $k \to \infty$ then:

$$C\left(F_{S_1}(s), F_{I_1}(k)\right) \approx \left[1 - \frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}\right].e^{-L_\alpha^{\frac{1}{\alpha}}} \tag{28}$$

Thus, the p-th percentile of the conditional distribution of the random variable I_1 given the random variable S_1 is obtained from a formula similar to (24) by:

$$p \approx \frac{\left[\ln\left(1 - R_\gamma(s)\right)\right]^{\alpha-1}}{\left[-L_\alpha^{\frac{1}{\alpha}} + \ln\left(1 - \frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}\right)\right]^{\alpha-1}} \times \frac{\left[1 - \frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}\right].e^{-L_\alpha^{\frac{1}{\alpha}}}}{1 - R_\gamma(s)} \tag{29}$$

where $L_\alpha = R_\gamma^\alpha(s) + R_\nu^\alpha(k_\nu(p))$, and $k_\nu(p) = \inf\left\{m \in \mathbb{N}, R_\nu(m) \leq 1 - p\right\}$.

Because $\ln\left(1 - R_\gamma(s)\right) \approx -R_\gamma(s)$ and $\ln\left(1 - \frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}\right) \approx -\frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}$, then:

$$\left[\ln\left(1 - R_\gamma(s)\right)\right]^{\alpha-1} \approx \left(-R_\gamma(s)\right)^{\alpha-1} \tag{30}$$

and

$$\begin{aligned}\left[-L_\alpha^{\frac{1}{\alpha}} + \ln\left(1 - \frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}\right)\right]^{\alpha-1} &\approx \left[-L_\alpha^{\frac{1}{\alpha}} - \frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}\right]^{\alpha-1} \\ &\approx \left(-L_\alpha\right)^{\frac{\alpha-1}{\alpha}}\left(1 + \frac{1}{2}\frac{L_{\alpha+1}}{L_\alpha}\right)^{\alpha-1} \\ &\approx \left(-L_\alpha\right)^{\frac{\alpha-1}{\alpha}}\left(1 + \frac{\alpha-1}{2}\frac{L_{\alpha+1}}{L_\alpha}\right)\end{aligned} \tag{31}$$

Therefore :

$$p \approx (-1)^{\alpha+\frac{1}{\alpha}}\frac{\left(R_\gamma(s)\right)^{\alpha-1}}{1 - R_\gamma(s)} \times \frac{\left(1 - \frac{1}{2}L_{\alpha+1}L_\alpha^{\frac{1-\alpha}{\alpha}}\right)e^{-L_\alpha^{\frac{1}{\alpha}}}}{L_\alpha^{\frac{\alpha-1}{\alpha}}\left(1 + \frac{\alpha-1}{2}\frac{L_{\alpha+1}}{L_\alpha}\right)} \tag{32}$$

In the case of the independence between I_1 and S_1 in the Gumbel copula, $\alpha = 1$, and we have:

$$p \approx \frac{\left(1 - \frac{1}{2}L_2\right)e^{-L_1}}{1 - R_\gamma(s)}$$

This formula is general and can be applied to any distribution function. In the next section we will apply it on the Poisson distribution.

4.4. **Quantile regression with Poisson distributions.** In this Section we will assume that F_{S_1} and F_{I_1} follow Poisson distributions, whose parameters γ and ν are respectively the expected values of the susceptible and infective host populations sizes.

The distribution functions $F_{S_1}(s)$ and $F_{I_1}(k)$ of the Poisson distributions are defined as:

$$\begin{aligned} F_{S_1}(s) &= \sum_{i=0}^{s} \frac{\gamma^i}{i!} e^{-\gamma} \\ &= 1 - R_\gamma(s) \end{aligned} \tag{33}$$

with: $R_\gamma(s) = \sum_{i=s+1}^{+\infty} \frac{\gamma^i}{i!} e^{-\gamma}$

$$\begin{aligned} F_{I_1}(k) &= \sum_{i=0}^{k} \frac{\nu^i}{i!} e^{-\nu} \\ &= 1 - R_\nu(k) \end{aligned} \tag{34}$$

with: $R_\nu(k) = \sum_{i=k+1}^{+\infty} \frac{\nu^i}{i!} e^{-\nu}$

Then the p-th quantile is as in equation (32) with $L_\alpha = R_\gamma^\alpha(s) + R_\nu^\alpha(k(p))$, and $k(p) = \inf\left\{m \in \mathbb{N}, \sum_{i=m+1}^{\infty} \frac{\nu^i}{i!} e^{-\nu} \leq 1 - p\right\}$.

5. **The proportional risk model and the copula approach.** Let us consider now the Cox model with proportional risk [6] and suppose that the risk function would be given by $h(t,z) = e^{\rho z} b(t)$, where ρ is a regression parameter and $b(t)$ the baseline risk function. Then, by denoting $q = e^{\rho z}$, the survival function $T(t,q)$ (i.e., the probability to survive until the age t with a risk q) is given by:

$$T(t,q) = \exp\left[-\int_0^t h(s,z)ds\right] = B(t)^q \tag{35}$$

where $B(t) = \exp[-\int_0^t b(s)ds]$.

In the Macdonald model (1), for calculating the survival function $T_2(t,q)$ of the subpopulation I_2, it is possible to identify $z = \log(\beta_1\beta_2/N_v N_H \mu r)$, $\rho = -K/\mu$, $t = 1/K$, $b(s) = cste = \mu$, $B(1/K) = e^{-\mu/K}$ and $R_0 = [\beta_1\beta_2/N_V N_H \mu r]e^{-\mu/K} = \exp[\log(\beta_1\beta_2/N_V N_H \mu r)] \ . \exp[-\mu/K] \approx T_2(1/K, (\beta_1\beta_2/N_V N_H \mu r)^{-K/\mu})$, if $\beta_1\beta_2/N_V N_H \mu r$ is close to 1.

If there exist n age classes into the vector subpopulation E_2 whose sojourn times $U_i (i = 1, ..., n)$ are independent random variables related to the survival function T_i, we have:

$$\mathbb{P}(U_i > t_i, i = 1, ..., n|q) = \prod_{i=1}^{n} T_i(t_i, q) = \prod_{i=1}^{n} B_i(t_i)^q \tag{36}$$

If z is a random variable, then $q = e^{\rho z}$ is also a random variable and we define the mean survival function as $T(t) = \mathbb{E}_q[B(t)^q]$.

If we consider now the Laplace transform defined by: $\mathbb{E}_q[e^{-vq}] = \exp(-v^p) = L(v)$, where p is a parameter depending on the probability distribution of q, we can write:

$$\begin{aligned} \mathbb{P}(U_i > t_i, i = 1, ..., n) &= \mathbb{E}_q[\prod_{i=1}^{n} B_i(t_i)^q] \\ &= \mathbb{E}_q[\exp(q \sum_{i=1}^{n} \ln B_i(t_i))] \end{aligned} \tag{37}$$

$$
\begin{aligned}
&= \exp\left(-\sum_{i=1}^{n}\left[-\ln B_i(t_i)\right]^p\right)\\
&= \exp\left(-\left[\sum_{i=1}^{n}\left(-\ln T_i(t_i)\right)^{\frac{1}{p}}\right]^p\right)\\
&= C(S_1,...,S_n)
\end{aligned}
$$

where C is an Archimedean copula, precisely the Gumbel Copula.

6. **Conclusion.** In this paper, we have given some definitions of Archimedean copulas used to define conditional quantiles for analysing interactions between compartments of vectors and hosts in a system of transmission of malaria. By using the bivariate Gumbel copula, we have calculated explicitely conditional quantiles and applied it when the compartments sizes are supposed to be random and Poissonian.

Two of the direct applications of this work are:

- If the calculation of the regression parameter α of the Gumbel copula gives the conclusion that α is close to 1, then the sizes of the concerned populations (e.g., susceptible and infected) are not concordant or uncorrelated, which is in favour of a total population size non constant, hypothesis rarely done in the epidemiologic studies, especially in malaria spread modelling [5, 23], which leads to incorporate a refined demographic part in the epidemic model.
- The calculation of conditional quantiles with the general formula from the empiric distribution functions of the observed population sizes allows to reconstruct their density, hence allows to test a posteriori their Poissonian character, which only authorizes the use of the specific simplified formulae. The Poisson hypothesis has the interest to link the epidemic interaction mechanism to the stochastic molecular kinetics [1, 31], which is a way to model the contagious contacts.

7. **Appendix.**

7.1. **Appendix A (after [24]).**

$$
\begin{aligned}
\tau &= \mathbb{P}\left\{(X_1-X_2)(Y_1-Y_2)>0\right\} - \mathbb{P}\left\{(X_1-X_2)(Y_1-Y_2)<0\right\}\\
&= 2\times\mathbb{P}\left\{(X_1-X_2)(Y_1-Y_2)>0\right\}-1\\
&= 2\times\mathbb{P}\left\{(X_1>X_2\ ;\ Y_1>Y_2)\cup(X_1<X_2\ ;\ Y_1<Y_2)\right\}-1\\
&= 2\times\left[\mathbb{P}\left\{(X_1>X_2\ ;\ Y_1>Y_2)\right\}+\mathbb{P}\left\{(X_1<X_2\ ;\ Y_1<Y_2)\right\}\right]-1\\
&= 4\times\mathbb{P}\left\{(X_1>X_2\ ;\ Y_1>Y_2)\right\}-1\\
&= 4\int_x\int_y \mathbb{P}\left\{X_2\le x\ ;\ Y_2\le y\mid X_1=x\ ;\ Y_1=y\right\} f_{XY}(x,y)dxdy-1\\
&= 4\int_x\int_y F_{XY}(x,y)f_{XY}(x,y)dxdy-1\\
&= 4\int_x\int_y C\left(F_X(x),F_Y(y)\right) f_{XY}(x,y)dxdy-1
\end{aligned}
\tag{38}
$$

In making the change of variables $u=F_X(x)$ and $v=F_Y(y)$, we get:

$$\tau = 4\int_0^1\int_0^1 C(u,v)c(u,v)dudv-1$$

where $c(u,v)=\frac{\partial^2 C(u,v)}{\partial u\partial v}$

7.2. **Appendix B (after [24]).**

$$\rho = 3\left[\mathbb{P}\left((X_1 - X_2)(Y_1 - Y_3) > 0\right) - \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_3) < 0\right)\right]$$

So according to Appendix A, we can write:

$$\rho = 3\left[4\int_x\int_y C\left(F_X(x), F_Y(y)\right) dC\left(F_X(x), F_Y(y)\right) - 1\right]$$

In making the change of variables $u_1 = F_X(x)$ and $u_2 = F_Y(y)$, we get:

$$\rho = 12\int_0^1\int_0^1 C(u_1, u_2)dC(u_1, u_2) - 3$$

where $dC(u_1, u_2) = \partial^2 C(u, v)/\partial u_1 \partial u_2 \; du_1 du_2$
Since X_2 and Y_3 are independent, then $C(u_1, u_2) = u_1 u_2$
So

$$\rho = 12\int_0^1\int_0^1 u_1 u_2 dC(u_1, u_2) - 3$$

7.3. **Appendix C (after [24]).** For both tau and rho, the first six properties in Definition (2.3) are obvious from the properties of the Kendall's tau and the Spearman's rho. For the seventh property, we note that the Lipschitz condition implies that any family of copulas is equicontinuous, thus the convergence of C_n to C is uniform.

Lipschitz condition:
Let C' be a copula. If for every $(u_1, u_2), (v_1, v_2)$ in $DomC'$

$$\mid C'(u_2, v_2) - C'(u_1, v_1) \mid \leq \mid u_2 - u_1 \mid + \mid v_2 - v_1 \mid$$

then C' is uniformly continuous on its domain.

REFERENCES

[1] A. F. Bartholomay, *Stochastic models for chemical reactions: I. Theory of the uni-molecular reaction process*, Bull. Math. Biophys., **20** (1958), 175–190.

[2] A. F. Bartholomay, *Stochastic models for chemical reactions: II. The unimolecular rate constant*, Bull. Math. Biophys., **21** (1959), 363–373.

[3] D. Beaudoin and L. Lakhal-Chaieb, *Archimedean copula model selection under dependent truncation*, Stat. Med., **27** (2008), 4440–4454.

[4] M. F. Boni, C. O. Buckee and N. J. White, *Mathematical models for a new era of malaria eradication*, PLoS Medicine, **5** (2008), e231.

[5] W. A. Brock, J. A. Scheinkman, W. D. Dechert and B. LeBaron, *A test for independence based on the correlation dimension*, Econometric Reviews, **15** (1996), 197–235.

[6] R. M. Cooke and O. Morales-Napoles, *Competing risk and the Cox proportional hazard model*, J. Stat. Plan. Inference, **136** (2006), 1621–1637.

[7] M. Delbrück, *Statistical fluctuations in autocatalytic reactions*, J. Chem. Phys., **8** (1940), 120–124.

[8] J. Demongeot, *Biological boundaries and biological age*, Acta Biotheoretica, **57** (2009), 397–419.

[9] J. Demongeot, J. Gaudart, J. Mintsa and M. Rachdi, *Demography in epidemics modelling*, Communications on Pure and Applied Analysis, **11** (2012), 61–82.

[10] J. Demongeot and J. Waku, *Counter-examples for the population size growth in demography*, Math. Pop. Studies, **12** (2005), 199–210.

[11] J. Demongeot, O. Hansen, A. S. Jannot and C. Taramasco, *Random modelling of contagious (social and infectious) diseases: Examples of obesity and HIV and perspectives using social networks*, IEEE Advanced Information Networking and Application (AINA'12, Fukuoka March 2012), IEEE Proceedings, Piscataway, (2012), 101–108.

[12] A. Ducrot, S. B. Sirima, B. Som and P. Zongo, *A mathematical model for malaria involving differential susceptibility, exposedness and infectivity of human host*, J. Biol. Dynamics, **3** (2009), 574–598.

[13] W. E. Frees and E. A. Valdez, *Understanding relationships using copulas*, Actuarial Research Conference, at the University of Calgary, Alberta, Canada, (1997).

[14] J. Gaudart, M. Ghassani, J. Mintsa, J. Waku, M. Rachdi, O. K. Doumbo and J. Demongeot, *Demographic and spatial factors as causes of an epidemic spread, the copula approach*, IEEE Advanced Information Networking and Application (AINA'10, Perth April 2010), IEEE Proceedings, Piscataway, (2010), 751–758.

[15] J. Gaudart, M. Ghassani, J. Mintsa, J. Waku, M. Rachdi and J. Demongeot, *Demography and diffusion in epidemics: malaria and black death spread*, Acta Biotheoretica, **58** (2010), 277–305.

[16] J. Gaudart, O. Tour, N. Dessay, A. L. Dicko, S. Ranque, L. Forest, J. Demongeot and O. K. Doumbo, *Modelling malaria incidence with environmental dependency in a locality of Sudanese savannah area, Mali*, Malaria J., **8** (2009), 61.

[17] D. T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, Journal of Computational Physics, **22** (1976), 403–434.

[18] W. Hoeffding, *On the distribution of the rank correlation coefficient τ when the variates are not independent*, Biometrika, **34** (1947), 183–196.

[19] P. Hougaard, *Modelling multivariate survival*, Scand. J. Statist., **14** (1987), 291–304.

[20] P. Hougaard, *A class of multivariate failure time distributions*, Biometrika, **73** (1986), 671–678.

[21] W. O. Kermack and A. G. McKendrick, *Contributions to the mathematical theory of epidemics. II. The problem of endemicity*, Proceedings of the Royal Society of London Series A, **138** (1932), 834–841.

[22] W. O. Kermack and A. G. McKendrick, *Contributions to the mathematical theory of epidemics. III. Further studies of the problem of endemicity*, Proceedings of the Royal Society of London Series A, **141** (1933), 94–122.

[23] J. C. Koella and R. Antia, *Epidemiological models for the spread of anti-malarial resistance*, Malaria J., **2** (2003), 3.

[24] W. Kruskal, *Ordinal measures of association*, Journal of the American Statistical Association, **53** (1958), 814–861.

[25] G. Macdonald, "The Epidemiology and Control of Malaria," Oxford University Press, London, 1957.

[26] S. Mandal, R. R. Sarkar and S. Sinha, *Mathematical models of malaria -a review*, Malaria J., **10** (2011), 202.

[27] A. W. Marshall and I. Olkin, *Families of multivariate distribution*, Journal of the Amercian Statistical Association, **83** (1988), 199–210.

[28] A. G. McKendrick, *Applications of mathematics to medical problems*, Proc. Edinburgh Mathematical Society, **44** (1925), 1–34.

[29] D. A. McQuarrie, *Kinetics of small systems. I.*, J. Chem. Phys., **38** (1963), 433–436.

[30] D. A. McQuarrie, C. J. Jachimowski and M. E. Russell, *Kinetics of small systems. II.*, J. Chem. Phys., **40** (1964), 2914–2921.

[31] D. A. McQuarrie, *Stochastic approach to chemical kinetics*, J. Appl. Prob., **4** (1967), 413–478.

[32] R. B. Nelsen, *Copulas and association*, in "Advances in Probability Distributions With Given Marginals. Beyond the Copulas," Kluwer-Dortrecht,Amsterdam, (1991), 51–74.

[33] R. B. Nelsen, "An Introduction to Copulas," Springer-Verlag, New York, 1999.

[34] F. Plasschaert, K. Jones and M. Forward, *Energy cost of walking: Solving the paradox of steady state in the presence of variable walking speed*, Gait and Posture, **29** (2009), 311–316.

[35] P. Pongsumpun and I. M. Tang, *Mathematical model for the transmission of plasmodium vivax malaria*, Int. J. Math. Models Methods in Appl. Sciences, **1** (2007), 117–121.

[36] T. Roncalli, "La Gestion des Risques Financiers," Economica, Paris, 2004.

[37] R. Ross, *An application of the theory of probabilities to the study of a priori pathometry. Part I*, Proceedings of the Royal Society of London Series A, **92** (1916), 204–230.

[38] D. L. Smith and F. E. McKenzie, *Statics and dynamics of malaria infection in Anopheles mosquitoes*, Malaria J., **3** (2004), 13.

Received April 2012; revised February 2013.

E-mail address: Jacques.Demongeot@agim.eu
E-mail address: Mohamad.Ghassani@agim.eu
E-mail address: Mustapha.Rachdi@agim.eu
E-mail address: iouassou@yahoo.fr
E-mail address: Carla.Taramasco@polytechnique.edu